ABSTRACT

Title of Document:	A RISK-INFORMED DECISION-MAKING METHODOLOGY TO IMPROVE LIQUID ROCKET ENGINE PROGRAM TRADEOFFS
	Richard Strunz, Doctor of Philosophy, 2013
Directed By:	Associate Professor, Jeffrey W. Herrmann, Mechanical Engineering and Institute for Systems Research

This work provides a risk-informed decision-making methodology to improve liquid rocket engine program tradeoffs with the conflicting areas of concern affordability, reliability, and initial operational capability (IOC) by taking into account psychological and economic theories in combination with reliability engineering. Technical program risks are associated with the number of predicted failures of the test-analyze-and-fix (TAAF) cycle that is based on the maturity of the engine components. Financial and schedule program risks are associated with the epistemic uncertainty of the models that determine the measures of effectiveness in the three areas of concern. The affordability and IOC models' inputs reflect non-technical and technical factors such as team experience, design scope, technology readiness level, and manufacturing readiness level. The reliability model introduces the Reliability-As-an-Independent-Variable (RAIV) strategy that aggregates fictitious or actual hot-fire tests of testing profiles that differ from the actual mission profile to estimate the

system reliability. The main RAIV strategy inputs are the physical or functional architecture of the system, the principal test plan strategy, a stated reliability-by-credibility requirement, and the failure mechanisms that define the reliable life of the system components. The results of the RAIV strategy, which are the number of hardware sets and number of hot-fire tests, are used as inputs to the affordability and the IOC models. Satisficing within each tradeoff is attained by maximizing the weighted sum of the normalized areas of concern subject to constraints that are based on the decision-maker's targets and uncertainty about the affordability, reliability, and IOC using genetic algorithms. In the planning stage of an engine program, the decision variables of the genetic algorithm correspond to fictitious hot-fire tests that include TAAF cycle failures. In the program execution stage, the RAIV strategy is used as reliability growth planning, tracking, and projection model.

The main contributions of this work are the development of a comprehensible and consistent risk-informed tradeoff framework, the RAIV strategy that links affordability and reliability, a strategy to define an industry or government standard or guideline for liquid rocket engine hot-fire test plans, and an alternative to the U.S. Crow/AMSAA reliability growth model applying the RAIV strategy.

A RISK-INFORMED DECISION-MAKING METHODOLOGY TO IMPROVE
LIQUID ROCKET ENGINE PROGRAM TRADEOFFS

By

Richard Strunz

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
2013

Advisory Committee:
Associate Professor Jeffrey W. Herrmann, Chair
Professor Mohammad Modarres
Professor Ali Mosleh
Professor Peter Sandborn
Associate Professor David Akin, Dean's Representative

Acknowledgements

Table of Contents

List of Acronyms

ACF	Autocorrelation Function
AF	Acceleration Factor
AHP	Analytic Hierarchy Process
AMSAA	Army Materiel Systems Analysis Activity
ANOVA	Analysis of Variance
BRDT	Bayesian Reliability Demonstration Testing
BTDA	Bayesian Test Data Aggregation
CCCG	Common-Cause Component Group
CER	Cost Estimation Relationship
DD	Design and Development
DoD	Department of Defense
DVS	Design Verification Specification
EQL	Equivalent Life
EQM	Equivalent Mission
ESA	European Space Agency
FLPP	Future Launchers Preparatory Programme
FMEA	Failure Mode Effects Analysis
FMECA	Failure Mode Effects and Criticality Analysis
FT	Fault Tree
FTA	Fault Tree Analysis
GG	Gas Generator
HFTD	Hot-Fire Test Duration
IOC	Initial Operational Capability
IRL	Integration Readiness Level
ISTB	Integrated Subsystem Test Bed
LCC	Life Cycle Cost
LH2	Liquid Hydrogen
LOx	Liquid Oxygen
LRECM	Liquid Rocket Engine Cost Model
MCDM	Multiple Criteria Decision-Making

MCMC	Markov chain Monte Carlo
MCTPP	Multiple Criteria Test Planning Problem
MFV	Main Fuel Valve
MH	Metropolis-Hastings
MoE	Measure of Effectiveness
MOV	Main Oxygen Valve
MRL	Manufacturing Readiness Level
MTBF	Mean Time Between Failure
NAFCOM	NASA/Air Force Cost Model
NASA	National Aeronautics and Space Administration
NHPP	Non-Homogeneous Poisson Process
O&S	Operations & Support
POF	Physics-Of-Failure
POV	Preburner Oxygen Valve
PP	Powerpack
PPP	Purchasing Power Parity
PRA	Probabilistic Risk Assessment
PRICE-H	Parametric Review of Information for Cost and Evaluation – Hardware
RAIV	Reliability-As-an-Independent-Variable
RBD	Reliability Block Diagram
R-by-C	Reliability-by-Credibility (Confidence)
RISDM	Risk-informed Satisficed Decision-Making
ROCOF	Rate of OCcurrence Of Failure
SEER-H	System Evaluation & Estimation of Resources – Hardware
SER	Schedule Estimating Relationship
SMART	Simple Multi-Attribute Rating Technique
SOGA	Single-objective Genetic Algorithm
SSME	Space Shuttle Main Engine
STH	System Test Hardware
STME	Space Transportation Main Engine
TAAF	Test-Analyze-And-Fix
TBCI	Two-sided Credibility Interval

TBPI	Two-sided Bayes Probability Interval
TCA	Thrust Chamber Assembly
TCO	Total Cost of Ownership
TFU	Theoretical First Unit
TRL	Technology Readiness Level
WinBUGS	Windows Bayesian inference Using Gibbs Sampling
YADAS	Yet Another Data Analysis System

List of Tables

List of Figures

Chapter 1:   Introduction

National prestige and military requirements previously dominated the decisions made during the development of new space transportation systems. Design choices for various subsystems were driven by the need to maximize performance, to minimize weight, and to master new technologies. This paradigm has changed, and affordability, reliability, and Initial Operational Capability (IOC), now elevated to the same level of importance with a lesser focus on performance optimization, have become the prime areas of concern in the decision-making process because they drive the overall operational effectiveness of any future space transportation system. However, these three areas of concern create a conflict because decision-makers must make tradeoffs between them.

In that context, liquid rocket engines play a dominant role for the following three reasons: (1) the engine's development and production prices are roughly 50 percent of the overall affordability of expendable space transportation systems [1], (2) the mission success is dominated by the component reliabilities of the propulsion system (i.e., more than 60 percent of all launch failures are associated to propulsion system failures) [2], and (3) the overall space transportation system performance is restricted by the maturity level of the component technologies that generate the required propulsive power levels (i.e., mainly thrust level and vacuum specific impulse) [3]. The lack of required maturity levels of enabling technologies is, however, directly linked to the IOC of the space transportation system. Therefore, the

areas of concern on liquid rocket engine level correspond directly with the areas of concern on space transportation system level, i.e., affordability, reliability, and IOC. The reliability of liquid rocket engines is generally obtained by both using the highest quality materials and conducting costly and lengthy Test-Analyze-And-Fix (TAAF) hot-fire test cycles that depend on the maturity levels of the component technologies. Therefore, it is obvious that these three areas of concern are not only interrelated but also in conflict and that the selection of the best liquid rocket engine system configuration, which meets the minimum performance requirements, becomes a Risk-Informed Satisficed Decision-Making (RISDM) problem. This dissertation presents a strategy for solving this problem.

1.1   Problem Statement

The state-of-the-art modeling approaches for the three dominating decision-making areas of concern affordability, reliability, and IOC are incomplete. Therefore, the manufacturers and agencies in the space industry lack a comprehensible and consistent solution strategy for the selection of the best liquid rocket engine system configuration [4-9].

Modeling *affordability* has been advancing since the 1980s due to the introduction of parametric cost models for the development and production cost for liquid rocket engines [10-12]. One of these models is the Rocketdyne developed Liquid Propellant Rocket Engine Cost Model (LRECM) [6] that is implemented in the contractor version of the NASA/Air Force Cost Model (NAFCOM®) [13]. However, the *main shortcoming* of the LRECM *is the lack of a quantitative link*

between the areas of concern *affordability and reliability* according to Hunt [7] and his experience on the development costing work for the RS-84 and J2-X liquid rocket engines. He also mentioned the difficulties he had about the *TAAF cycle assumptions* which are strongly related to the reliability modeling and the impact on both the development cost/price and the IOC.

Modeling *reliability* includes two aspects: (1) the inherent reliability assurance modeling techniques and (2) the inherent reliability verification. The inherent reliability assurance modeling techniques are well advanced and include, for example, reliability planning and specification, allocation, prediction, Failure Mode Effects and Criticality Analysis (FMECA), Fault Tree Analysis (FTA) [14] or Probabilistic Risk Assessment (PRA) for safety related issues [15]. The inherent reliability verification is based on both analyses and hot-fire tests. However, the confidence build-up of liquid rocket engine reliability is really obtained by means of component, subsystem, system development and finally through system qualification or certification hot-fire tests that feature different testing profiles, i.e., different hot-fire test durations and operational load points, that include also extreme testing loads in order to demonstrate design maturity/robustness [16]. Modern multilevel attribute data aggregation techniques exist to estimate the system level reliability [17-19] but they *lack the capability of aggregating different testing profiles* that trigger *multiple failure mechanisms* in system components. *But how to scope,* by means of a hot-fire test plan, *these testing profiles to attain a stated system reliability requirement that may include TAAF cycle assumptions*? *No liquid rocket engine hot-fire test plan standard/guideline exists* [4, 5, 20-22], but there is complete agreement about the

strong relationship between mission success (reliability) and the amount of hot-fire testing [4, 5, 8, 16, 20, 21].

Modeling *IOC* seems to be straightforward, but it is not because of the *dependency of the development schedule on* the other two areas of concern (*affordability* and *reliability*). In addition, the liquid rocket engine performance requirements drive the complexity of the thermodynamic cycle, the maturity levels of the enabling technologies, and consequently the scope of the hot-fire test plan.

1.2   Objectives

The main objective of the research described in this dissertation, motivated by the European initiative to prepare the development of the Next Generation Launcher [23] and the lack of a hot-fire test plan standard/guideline [20], is the development and testing of a RISDM methodology that includes quantitative links between the areas of concern affordability, reliability, and IOC, takes into account technical and non-technical factors, bases the TAAF cycle assumptions on the Technology Readiness Level (TRL) or similarly the novelty and maturity of the components, aggregates testing profiles that are different from the mission profile, accounts for multiple failure mechanisms, and scopes hot-fire test plans taking into account a stated Reliability-by-Credibility (R-by-C) requirement in order to equip decision-makers with a comprehensible and consistent solution strategy for the selection of the best liquid rocket engine system configuration in early project/program life cycle phases.

1.3   Significance of Dissertation

The RISDM methodology provides a comprehensible framework for tradeoffs that combines deterministic and probabilistic modeling of the three conflicting areas of concern (affordability, reliability, and IOC) using technical and non-technical factors and using the bounded rationality theory as reference framework [24-26]. In that context, the Reliability-As-an-Independent-Variable (RAIV) strategy is developed [27] that is also used in a Bayesian alternative to the Crow/AMSAA reliability growth model [28]. The RISDM methodology is also used to define satisficed hot-fire test plans given a stated R-by-C requirement [29].

1.3.1   Risk-informed Satisfied Decision-Making Methodology

The RISDM methodology combines psychological and economic theories and is formulated as a multiobjective satisficing problem that is solved using genetic algorithms in which the fitness function is defined by a weighted sum of truncated exponential utility functions that reflect the risk attitude of the decision-maker for each of the three areas of concern (affordability, reliability, and IOC). The risk attitude, defined by the effective risk aversion coefficient using the normative target-based decision theory, determines the shape of the utility functions. The measures of effectiveness for each of the three utility functions are determined by the interdependent affordability model, the RAIV strategy (see Section 1.3.2), and the IOC model, which depend on the decision variables, the number of hot-fire tests. Risks are expressed as TAAF cycle assumptions, i.e., number of hot-fire test failures,

which are estimated using the novelty and maturity of the system component technologies and the level of severity of the failure-inducing agents.

1.3.2   Reliability as an independent Variable Strategy

The RAIV strategy addresses the lack of an existing multilevel attribute data aggregation technique that estimates the system level reliability if both different testing profiles and multiple failure mechanisms are present. The solution approach to the RAIV strategy is based on the Bayesian estimation using a blockwise Metropolis-Hastings algorithm. The likelihood function, in view of the competing risks theory, is a function of component level reliabilities that reflects the multilevel hot-fire test strategy for which the data is defined as Equivalent Mission (EQM) in order to account for the different testing profiles and failure mechanisms. The priors for the component level reliabilities are based on two-component mixture distributions, i.e., a composite of a Jeffreys' prior and a Beta distribution in which the mix parameters reflect the knowledge transfer factor to account for the novelty and maturity levels of the component technologies. The validation of the RAIV strategy uses hot-fire test data from the U.S. liquid rocket engines F-1 and SSME. In addition, it was applied to the U.S. liquid rocket engine RS-68 and the European liquid rocket engine Vulcain 1.

1.3.3   Reliability Growth Model: a Bayesian Approach

The well-known empirical Duane and analytical Crow/AMSAA models are no longer best practice approaches to model reliability growth for systems, such as liquid rocket engines, if different hot-fire testing profiles are used to verify the inherent reliability [30, 31]. The RAIV strategy is applied to the reliability growth model taking

6

advantage of the Bayesian updating property. The modeling of the TAAF cycle accounts also for the inclusion of hot-fire test failures that is typically in reliability growth testing.

1.4 Overview of Dissertation

This dissertation introduces the RISDM methodology to perform comprehensive and consistent tradeoffs in early project/program life cycle phases. The RISDM methodology combines psychological and economic theories and is formulated as a multiobjective satisficing problem that is solved using genetic algorithms. A central pillar of the RISDM methodology is the RAIV strategy because it establishes a quantitative relation between a system level reliability and affordability using the Bayesian estimation framework. The RAIV strategy is also applied to reliability growth modeling taking into account the differences between testing profiles and the mission profile. The application of the RISDM methodology and RAIV strategy is limited to liquid rocket engines in this research, but these approaches may also be applied to any other complex decision-making problem that involves conflicting areas of concern.

This Chapter 1 introduces the decision-making environment for liquid rocket engines, highlights gaps in the state-of-the-art modeling for the main three areas of concern (affordability, reliability, and IOC), and discusses the significance of this dissertation. Chapter 2 reviews previous work on psychological and economic theories that is relevant for the RISDM methodology. The implemented mathematical solution techniques of the RISDM methodology require a review of satisficing using

genetic algorithms, computational Bayesian estimation, and the normative target-based utility-probability duality. The specific decision-making environment of liquid rocket engines requires also some discussion. Chapter 2 concludes with a brief review of reliability growth model because the RAIV strategy is also applied in that context. Chapter 3 describes in detail the mathematical formulation of the RISDM methodology. It also provides sensitivity analyses for the epistemic uncertainty and variables of the affordability, reliability, and IOC models. Chapter 4 consists of three different problems that were solved with the general RISDM methodology and one discussion on the satisficing approach by comparing single-objective genetic algorithms with the well-known and frequently used elitist multiobjective non-dominated sorting genetic algorithms NSGA-II. Each of the problems or the discussion on satisficing can be read independently from one another; therefore, some repetition of material from Chapter 3 is inevitable. Section 4.1 describes the RAIV strategy applied to liquid rocket engine [27], Section 4.2 uses the RAIV strategy to optimize test plans of liquid rocket engines [29], Section 4.3 applies the RAIV strategy to reliability growth modeling [28], and Section 4.4 discusses the satisficing aspect of the RISDM methodology. Chapter 5 concludes this dissertation and identifies further research directions.

Chapter 2: Literature Review

The RISDM methodology combines various research areas and solution strategies into a single simulation framework. The literature review is, therefore, centered on these areas and strategies. It starts with applied decision theory because it is essential to understand the psychological and economic aspects of decision-making. This also includes the normative target-based decision-making approach. The implemented mathematical solution techniques of the RISDM methodology require a review of satisficing using genetic algorithms and computational Bayesian estimation. The specific decision-making environment of liquid rocket engines necessitates some discussion to acquaint the reader with this specific field of engineering. Interested readers about the theoretical foundations of liquid rocket engines are referred to [3, 32-34]. The Chapter concludes with a brief review of reliability growth modeling because the RAIV strategy is applied in that context.

2.1 Applied Decision Theory

Howard [35] argues that practical management decision-making problems are far from novel theoretical theorems or specific models but he defines a structured formal process for the analysis of decision-making under uncertainty. He stresses the point that a good decision is a comprehensible decision that includes uncertainties, areas of concern or objectives, and measures of effectiveness which should result in a good outcome; one with high value to the decision-maker. However, he also notes that a good decision may not always result in a good outcome. To provide some theoretical

background or specific models, Section 2.1.1 provides organization decision-making frameworks, and Section 2.1.2 discusses aspects of normative decision theory based selection.

2.1.1 Organization Decision-Making

Daft [36] frames decision-making into several organization decision-making processes such as the Management Science Approach, the Carnegie Model, the Incremental Decision Process Model, and the Garbage Can Model.

The Management Science Approach is based on mathematical and statistical techniques for decision problems with well-defined and measureable variables; however, if the main variables cannot be quantified then even the most sophisticated model fails.

The Carnegie Model is based on the bounded rationality approach postulated by Simon [24] and the problemistic search introduced by Cyert and March [37]. The bounded rationality approach or the rational choice features common decision-making constraints such as a limited set of alternatives, a relationship that determines the measure of effectiveness (satisfaction or goal attainment), and the preference-orderings among the measures of effectiveness. Therefore, the bounded rationality approach includes the key characteristics of good decisions that were defined by Howard [35]. The problemistic search tries to quickly find a solution but it does not search for a perfect solution. The implication of the bounded rationality approach and problemistic search on organization–level decisions with ambiguous and inconsistent goals is that the final selection of the best alternative is based on a coalition of the main stakeholders. The main stakeholders could include internal and even external

groups. However, the process of coalition implies that decisions will be made to satisfy with suffice rather than to search for the solution that maximizes the measure of effectiveness. Based on the process of coalition, Simon [25] introduced the word "satisfice" to describe this type of decision-making. Manktelow [38] describes satisficing as portmanteau that combines the sound and the meanings of the two words "satisfy and suffice."

Simon [24, 26], Manktelow [38], and Gilboa [39] discuss the differences between satisficing and maximizing (optimizing) by looking at psychological and economic theories. In the classical economic theory the notion of satiation is not accounted for, but it is in psychology theory, which defines the motivation to act as long as no satisfaction is obtained. In addition, the motivation to attain a certain level of satisfaction is not fixed, but it is usually specified by an aspiration level that is based on past experience. If this motivation for satisfaction is reflected against the business behavior of a company, the main objective of that company would be to try to satisfice rather than to maximize (optimize) by attaining a certain level of market share, profit, or sales. The level of attainment is associated with the attained level of the measure of effectiveness of a particular area of concern, i.e., affordability, reliability, and IOC in the context of this research.

Mintzberg et al. [40] develop the incremental decision process model (see Figure 2-1) that is based on empirical evidences from 25 strategic decisions wherefore no predefined set of alternatives existed, i.e., custom-designed solutions were found for each decision. The model consists of three phases: identification, development, and selection and features the main two elements recognition and

evaluation-choice that are accompanied with the elements diagnosis, search/screen, design, and authorization. The element recognition refers to opportunities, problems, and crises. The evaluation-choice utilizes three modes: judgment, bargaining, and analysis. Despite the fact that the normative literature focuses on the analytic models that are based on maximizing predetermined utility functions, it is the least applied approach in the strategic decision-making process because of the inclusion of a large number of soft factors which are not easily quantitatively modeled. Soelberg [41] discusses the approaches maximizing and satisficing in that context. In cases where political considerations with contentious goals are key elements in the strategic decisions-making, the bargaining selection is, however, applied most often. The diagnosis element is concerned with the understanding of the cause-effect relationship and the need to perform the decision-making process. The search/screen and design elements are the heart of the overall decision-making process because they seek for ready-made (purchased item or furnished items) or custom-made solutions which are found in a complex, iterative procedure. The authorization element completes the decision-making process by selecting the best alternative that was found in the evaluation-choice element. The incremental decision process model features also interrupts that are either caused by internal or external forces as well as by new options for the ready-made or custom made solutions. In that context, Meisl [42] proposes a space transportation booster engine selection methodology that matches the main principles of the incremental decision process model approach as depicted with red marking in Figure 2-1. Mintzberg et al. [40] provide further examples in which only specific elements of the incremental decision process model were used.

In addition, Krevor [43] presents a methodology that links cost/price and reliability for early conceptual design work of space transportation systems. The methodology follows the incremental decision process model, i.e., the recognition for the need to design a new space transportation system was declared by the US president George W. Bush [44] and enforced by NASA. The conceptual design determines top-level performance requirements and the physical architecture for each space transportation system configuration. Based on the system configurations, the reliability models and Cost Breakdown Structures are established and the optimal configuration selected. One of the problems of Krevor's methodology is, however, linked to the cost modeling of the liquid rocket engines. Krevor uses a fixed reliability figure that is independent from the planned hot-fire test program despite the agreement about the strong relation as mentioned in Section 1.3.1.
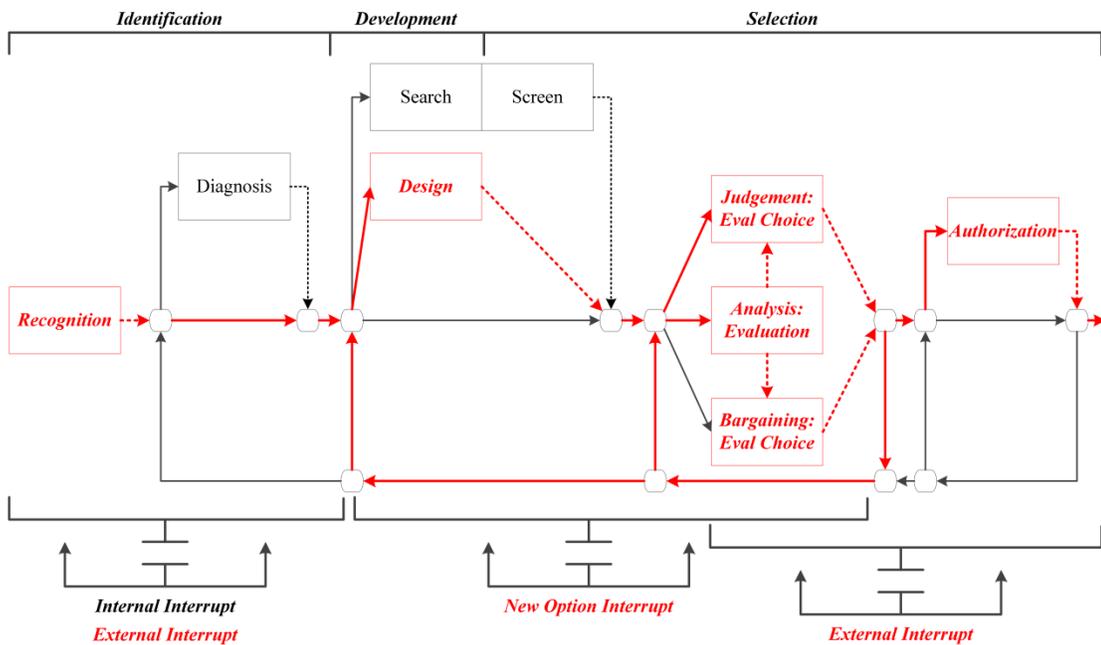
Figure 2-1:   Incremental Decision Process Model

The last organizational decision-making process model, the garbage can model, is the most recent model, that is described in Daft [36], which is not comparable with any of the above described models because it covers multiple flows of organizational decisions. Only a single flow of organizational decision is of interest in this research; therefore, the garbage can model is not further discussed.

2.1.2 Normative Decision Theory based Selection

Normative decision theory is a broad field of active research. The early work on satisficing problems was based on normative decision theory using expected utility theory for example [26]. In that context, expected utility theory was first addressed by Bernoulli [45] and then by von Neumann and Morgenstern [46] as well as by Savage [47]. However, expected utility theory was shortly criticized thereafter as descriptive model of decision-making under risk by Kahneman and Tversky [48] because empirical studies indicated the presence of a value function that is concave for gains, commonly convex for losses, and flatter for gains than for losses. Based on these empirical studies, Kahneman and Tversky [48] introduce their prospect theory with a new class of utility function. However, the prospect theory could not describe the classic Allais paradoxes [49], so an update was needed for the prospect theory, and that was named cumulative prospect theory [50]. Although the cumulative prospect theory could account for the Allais paradoxes, 11 new paradoxes arose for which the cumulative prospect theory led to contradiction or to erroneous predictions [51]. In order to overcome the identified paradoxes, recent research initiatives by Sewell [52] and Harrison and Rutström [53] focus on the combination of the expected utility

theory and the prospect theory. Unfortunately, no concluding prescriptive model has yet been published.

Bordley and LiCalzi [54] and Abbas and Matheson [55] work on another research direction using the utility–probability duality that was first discussed in detail by Abbas and Matheson [56]. The important result of the duality approach is the relation of a target, which is set by the decision-maker, to a unique effective risk aversion coefficient that is mathematically defined as

$$F(\hat{g}) = \int_{LB}^{UB} uf(g) F(g) dg = \int_{LB}^{UB} \frac{\gamma^{Eff} e^{-\gamma^{Eff} g}}{e^{-\gamma^{Eff} LB} - e^{-\gamma^{Eff} UB}} F(g) dg \qquad (2.1)$$

where $\hat{g}$ is the aspiration equivalent, $F(g)$ is the cumulative density function of the given lottery, γ^{Eff} is the effective risk aversion coefficient, and the integrands $\{LB, UB\}$ correspond to the lower and upper bound of the utility function. The effective risk aversion coefficient γ^{Eff}, which reflects the decision-maker's risk attitudes (risk-neutral, risk-averse, or risk-seeking), is fully determined given the decision maker's uncertainty bounds and a target for the specific area of concern. The utility–probability duality is appealing in the context of this research not only because of a continuous instead of a zero-one utility scale but also due to the normalization of different dimensions and ranges of the contradicting areas of concern. Note that Wilson [57] proposes a specific utility for reliability and survival that is based on expert elicitation. It features also different risk attitudes but does not elicit a reliability target or an R-by-C requirement.

The continuous behavior of the normalized areas of concern becomes practical in the satisficing formulation that is reviewed next.

15

2.2 Constrained Multiobjective Satisficing

Decision-making, as just outlined, is based on a satisficing strategy among conflicting areas of concern in which the satisficing strategy utilizes the classical constrained multiobjective optimizing using genetic algorithms. So why is then decision-making not just an optimization problem? Because of a subtle difference between satisficing and optimizing that is reviewed in Section 2.2.1 followed by the evolutionary computation in Section 2.2.2.

2.2.1 Satisficing versus Optimizing

Are we optimizers or satisficers? In that context, Odhnoff [58] discusses on the differences of optimizing and satisficing and concludes:

> *"...In my opinion there is room for both optimizing and satisficing models in business economics. Unfortunately, the difference between 'optimizing' and 'satisficing' is often referred to as a difference in the quality of a certain choice. It is a triviality that an optimal result in an optimizing model can be an unsatisfactory result in a satisficing model. The best thing would therefore be to avoid a general use of these two words."*

According to Odhnoff [58], there is, however, a subtle difference between the optimization and satisficing formulation. Optimization uses only what he calls a base model whereas the satisficing model uses three submodels: a base model that is equivalent to the model used in optimization, a seeking process, and an adaptation

process. Therefore, the main difference is linked to the seeking process that generates the alternatives and the adaptation process to select the best decision alternative.

Eilon [59] also compares managerial problem solving approaches that are based on optimizing and satisficing and concludes as follows:

> *"...True enough, the optimizing philosophy is the one that prevails in the literature, but experience and observation suggest that satisficing is the approach that prevails in practice. There is far more to be gained from scrutinizing and ranking constraints than in constructing a super utility function to delight the heart of the optimizer. ..."*

Whether we are optimizer or satisficer is not finally concluded in this research; however, the mathematical formulation of the RISDM methodology includes the characteristics of a base model and submodels that are used as inputs to a fitness function. Therefore, the approach should be satisficing if the classification of Odhnoff [58] is used. In addition, target values and ranges for the three areas of concern are expressed to include uncertainty which ranges are transferred into bounds of the measures of effectiveness imposing as a result the constraints of the feasible solutions. Eilon's [59] norm setting requirements would classify such a problem formulation also as satisficing rather than optimizing formulation.

Wierzbicki [60] discusses also the mathematical basis for satisficing decision-making models and introduces achievement scalarizing functions. These scalarizing functions feature order preservation and order approximation properties under the

limited rationality of choice of decision-makers as it is the case for the utility–probability duality derived utility functions.

2.2.2 Evolutionary Computation

The set of feasible solutions of the satisficing problem is found by applying classical optimization formulation which, according to Rao [61], may be characterized with regards to the methods of operation research: mathematical programming techniques, stochastic process techniques, statistical methods, and modern optimization techniques.

The nature of the RISDM methodology formulation rules out already the mathematical programming techniques because these search methods are calculus-based or enumerative. The calculus-based methods are either indirect or direct that seek local extrema or local optima using hill climbing techniques. Since the search is local in scope and requires continuous, unimodal, and easy derivatives of the objective functions, the application is rather limited. The enumerative schemes, such as dynamic programming, lack efficiency. The statistical methods transfer the stochastic programming problem into an equivalent deterministic problem which is then treated with the classical mathematical programming techniques.

One of the most widely used statistical methods for optimization is based on the Response Surface Methodology known from the design and analysis of experiments or robust design approaches. A major drawback of the Response Surface Methodology approach is that, because the mathematical formulation of the response surface is based on polynomials, it may not capture multimodal behavior (Kriging metamodels may be used as remedy to capture multimodal behavior). In addition, the

fitted surface should be as precise as possible, and this is a function of the experimental error, the experimental design, and the points located in the design space. The fraction of design space can be used as combined metrics. However, enough design points must be planned in order to obtain a constant fraction of design space level throughout the design space. The number of planned experiments could easily become as large as 350 to 500 for the problem of this research. Therefore, the Response Surface Methodology approach is not the most promising solution strategy because of the inherent experimental error and the high number of required experiments in comparison to modern optimization techniques that require similar number of searches. However, the Response Surface Methodology may be used to estimate metamodels for the RAIV strategy in order to improve the computational efficiency.

The modern optimization techniques, which are Monte Carlo based algorithms such as simulated annealing or evolutionary algorithms [62], are the last resort. In this research, the genetic algorithm, as one of the members of the evolutionary algorithms, is used to generate the sets of solutions for the satisficing problem. It is, however, reported that the simulated annealing or the combination of both genetic algorithm and simulated annealing offer advantages in terms of solution quality and number of iterations according to Gandomkar and Vakilian [63].

Kuo and Wan [64] discuss on optimal reliability design algorithms and current research directions. Noteworthy are the hybrid genetic algorithms and the ant colony optimization method that is a subset of evolutionary computation [62]. A simple genetic algorithm is a robust population-based direct random search method that can

be applied to almost all complex reliability problems but lacks computational efficiency. The remedy of the low computational efficiency is to combine simple genetic algorithms with heuristic algorithms, simulated annealing or simulated quenching, steepest ascent/descent methods, or any other local search method assuming that the local search is only unimodal. Combining the genetic algorithm with a more efficient algorithm is called a hybrid genetic algorithm.

More recent related applications are the ones by Nahas and Nourelfath [65], who use a problem-specific ant colony optimization method for optimizing the reliability of a series system with budget constraints, and Graves and Hamada [19], who assess the influence of test allocations on the system reliability uncertainty with multilevel data using a simple genetic algorithm. In addition, Tao et al. [66] apply a constrained multiobjective satisficing model that featured a linear weighted objective function and the classical optimization model formulation to an engineering design optimization problem using a genetic algorithm to generate the Pareto-optimal solution set. Tamura et al. [67] use a satisficing tradeoff method to solve a multiobjective combinatorial optimization problem with application to flow shop scheduling using also a genetic algorithm to generate the Pareto-optimal solution set.

It should be noted that the focus of this research is not on the improvement of the computational efficiency of optimization algorithms. Therefore, the selected solution strategy to solve the satisficing problem is based on the implementation of a genetic algorithm that optimizes a fitness function.

2.3 Computational Bayesian Statistics

The RAIV strategy is used to determine the measure of effectiveness of the area of concern reliability. The Bayesian estimation framework provides the proper mathematical implementation. Section 2.3.1 reviews Bayesian estimation in general, Section 2.3.2 the prior distribution, Section 2.3.3 the Bayesian aggregation of multilevel test data, and Section 2.3.4 computational impediments of Bayesian statistics, respectively.

2.3.1 Bayesian Estimation

Bayesian estimation refers to a statistical framework that looks upon parameters as random variables that have prior distributions. It is based on Bayes' Theorem [68] but extended to the continuous case; therefore, the expression Bayesian estimation is used according to Miller and Miller [69]. The combination of the prior distribution, which reflects the a priori information, with the sampling distribution, which models the evidence, results in the unscaled posterior distribution. This unscaled posterior provides the shape but does not feature the required properties of a random variable in order to find probabilities or moments. Therefore, no inference can be made. In order to obtain a posterior distribution, the unscaled posterior distribution must be scaled so that it integrates to one. The scaling factor is found by integrating the product of the prior distribution and sampling distribution. Mathematically, the posterior distribution is defined by

$$\pi\left(\underline{\theta}\mid Data\right) = \frac{L\left(Data\mid\underline{\theta}\right)\pi^0\left(\underline{\theta}\right)}{\int_{\Theta}L\left(Data\mid\underline{\theta}\right)\pi^0\left(\underline{\theta}\right)d\Theta} \tag{2.2}$$

where $\pi(\cdot)$ is the posterior distribution of the parameter vector $\underline{\theta}$ given the evidence $Data$, $L(\cdot)$ is the likelihood or sampling distribution, and $\pi^0(\cdot)$ is set of prior distributions for the parameters θ_i in the parameter vector $\underline{\theta}$ that defines the parameter space Θ.

In the context of this research, the evidence can be either pseudo or actual hot-fire test results. Note that the word "pseudo" was coined by Martz and Waller [70] and should mean "pretended" whereas Modarres et al. [71] use the word "fictitious." In the planning stage of a project/program, actual hot-fire test results are not available; therefore, pseudo/fictitious evidence is used to pretend hot-fire test results (successes or failures). The solution of the p-dimensional integral is usually found by numerical integration because closed form solutions exist only for sampling distributions that belong to the exponential family with conjugate prior distributions. The exponential family includes the continuous distributions Normal, Gamma, and Beta and discrete distributions Binomial, Poisson, and negative Binomial [72]. The sampling distributions used in this research are not members of the exponential family; therefore, numerical integration methods are needed.

2.3.2 Prior Distributions – The Criticism of the Bayesian Approach

Wasserman and Kass [73] and Robert [74] recall that the Bayesian estimation approach is often criticized because of the subjectivity involved in the generation of the prior distribution. The influence on the parameter estimation can be negligible, moderate, or enormous. Prior distributions should reflect the prior information, including the level of uncertainty, of the values of the parameters of interest.

In cases where little information is known a priori, the prior distribution may be dispersed naming the prior diffuse, noninformative, or vague. Depending on the sampling distribution, certain noninformative prior distributions are common choices due to their conjugacy, e.g. the binomial sampling distribution is used with a Beta prior distribution and the Multinomial sampling distribution with a Dirichlet prior distribution. Weiler [75] studies the sensitivity of different prior distribution shapes on the posterior distribution and concludes that the impact is negligible unless the prior distribution dominates the sampling distribution. Pham [76] substantiates the conclusion of Weiler and further argues as correspondence to the main findings by Duran and Booker [77] that the sensitivity depends also on the precision of the numerical code, i.e., round-off errors are very important. He further argues that the impact on the mean and standard deviation of the posterior distribution are much less affected from large variations of prior distribution parameters but rather emphasizes that the prior distribution dominates the sampling distribution for cases with a small amount of evidence.

In cases where the certainty about a parameter value is high, the prior distribution is concentrated around that value. Such prior distributions are then called informative. Information about the parameter values can be found by physical/chemical theory, computational analysis, previous test results, industry-wide generic reliability data, past experience, or expert opinions [18]. Siu and Kelly [78] provide some general advice on developing informative prior distributions in that context.

Waterman et al. [79], Martz and Waller [70], and Modarres et al. [71] present possible ways to obtain the beta prior distribution parameters based on actual

evidence, i.e., the liquid rocket engine mission reliability figures in the context of this research. McFadden and Shen [80] provide the relevant data for various liquid rocket engine systems.

Krolo [81] and Kleyner [82] present two approaches that allow the inclusion of computational analyses and past experience with similar products using a transformation factor or a knowledge and innovation factor. Note that the inclusion of computational analyses applies only to the method introduced by Krolo. Krolo's transformation factor approach is directly applied to the Beta distribution parameters and is derived from a Failure Mode and Effects Analysis (FMEA) in which risk priority numbers are used to calculate the transformation factor values that range from zero to one. Hitziger [83] enhances the work of Krolo and describes a qualitative approach, using fuzzy logic, and a quantitative approach, using the Kolmogorov-Smirnov test, to define the transformation factor. Kleyner's knowledge and innovation factor approach is used in a two-component mixture of Uniform and Beta distributions and is found subjectively. Note that the presented methods are applied only on system level and not on subsystem or component levels.

Kleyner [82] further argues that Krolo's [81] approach is more adequate for medium reliability targets $\left(0.90 \leq R < 0.98\right)$ using previous test results and computational results whereas his method suits better high reliability demonstration targets $\left(R \geq 0.98\right)$ using field data with only low failure rates. This research work proves that his argument is too limited.

2.3.3 Bayesian Test Data Aggregation

Bayesian test data aggregation refers to a statistical approach that combines multilevel data. The data, attribute or continuous, can be pseudo (fictitious), actual, or a mixture of pseudo and actual. The Mellin transform and the Monte Carlo methods have been successfully applied in that context.

The Mellin transform method belongs to a class of transform techniques for probability modeling. According to Giffin [84], the Mellin transform is useful for quotients and products of random variables. The latter one is applicable for reliability estimations. Once the system level random variable is found by means of Mellin convolution, the two moments mean and variance can be found easily by replacing the s-argument of the Mellin transform with constants. In that context, Mastran [85] studies a three component series system using Mellin transform in conjunction with the Bayesian estimation for component and system level attribute data. Springer and Thompson [86] apply the Mellin transform for a series system with exponential failure time distributions and Springer and Byers [87] modeled a mixture of a series system with exponential and attribute data. More complex systems such as parallel, r-out-of-k, and combinations of series-parallel components may also be modeled with the Mellin transform technique. Note that the applications are limited to a two level structure, i.e., component and system level. The interest in the Mellin transform method is nowadays limited in favor of the Monte Carlo methods due to the advancements made in computational Bayesian statistics and the availability of commercial software packages.

Modarres et al. [71] discuss three types of Monte Carlo methods: classical simulation, Bayesian simulation, and Bootstrap. The classical simulation uses the mathematical formulation of coherent systems and then simulates samples from the component reliabilities to obtain a system level probability sample from which any percentile can be calculated. The Bayesian method is similar to the classical except that the component reliabilities are estimated from posteriors that were generated from likelihood functions and prior distributions for the model parameters. The bootstrap, like the jackknife, is a nonparametric resampling method. The approach to estimate the system level reliability is, however, similar to the classical and Bayesian approach.

Martz and Duran [88] compare the Maximus (a frequentist method not reviewed in this research), the bootstrap, and the Bayes Monte Carlo simulation methods using binomial component level data for various complex systems. Based on the analyzed systems, none of the three methods was outstanding and no conclusive statement is made. Note that the applications are limited to a two level structure, i.e., component and system.

Martz and Waller [89] present a method to analyze the system reliability of series-parallel systems using a Bayesian procedure that aggregates either pseudo or actual data at system, subsystem, and component levels. They noted that a prior paper by Martz et al. [90] introduces the basics that is, however, limited to series systems. Martz and Waller claim that the introduced method is the first Bayesian method that integrates component, subsystem, and system pseudo or actual attribute test data.

In the context of multilevel data aggregation just described, Bier [91] and Azaiez and Bier [92] address the concern of aggregation errors in reliability models with Bayesian updating. They suggest two approaches to overcome this concern: (1) to update the component priors with component data and propagate up to obtain the system level posterior or (2) to propagate component priors up to the system prior and use the system level data to obtain the system level posterior. The problem is, however, that the two approaches result in different solutions. In order to overcome this discrepancy a new approach was developed that is discussed next.

Johnson et al. [93] introduce a Bayesian hierarchical estimation approach for complex multilevel systems that remedy the concerns raised by Bier [91] and Azaiez and Bier [92] by combining simultaneously all available attribute data and prior knowledge. The estimation approach expresses the higher system levels in terms of component reliabilities but maintains the coherent structure of the complex multilevel system; therefore, the posterior up to the normalization constant becomes a nested function which can only be solved with a Markov chain Monte Carlo method such as the Metropolis-Hastings algorithm.

Hamada et al. [17] or Graves and Hamada [19] apply the Bayesian hierarchical estimation approach for the assessment of system reliability with multilevel attribute data and the allocation of resources (additional attribute data collection) in order to minimize the uncertainty of the system reliability within a fixed budget. The optimal allocation of additional tests was found using a genetic algorithm.

The methods reviewed so far are all applied to attribute data. However, they can be applied to non-binomial data as well. The interested reader is referred to Thompson and Chang [94], Chang and Thompson [95] or Martz and Baggerly [96].

2.3.4 Computational Impediments of Bayesian Statistics

The p-dimensional integral in the divisor of the posterior distribution becomes the main impediment of Bayesian estimation because difficult numerical integrations, in particular if the parameter space is large, need to be performed. Two types of algorithms are used to draw samples from the posterior distribution: direct methods and Markov chain Monte Carlo methods [97].

According to Robert and Casella [98], the most common direct methods are the accept-reject methods, importance-resampling, and envelope/adaptive-rejection-sampling from log-concave distributions. The direct methods are, however, limited in application for posteriors with large parameter space because the acceptance proportion reduces significantly as the number of parameters increases [97]. The remedy is the Markov chain Monte Carlo method that provides an efficient algorithm for sampling from posteriors with large parameter space.

Metropolis et al. [99] introduce the Monte Carlo method that was significantly improved and extended by Hastings [100]; hence, the name Metropolis-Hastings algorithm.

In a Markov chain, random numbers are simulated from more or less arbitrary distribution with density $h\left(\underline{\theta}\,|\,\underline{\theta}^{(m)}\right)$ in which m corresponds to the iteration index of

the chain. The Metropolis-Hastings algorithm either accepts the proposed random

number $\underline{\theta}^*$ that is drawn from $h\left(\underline{\theta} \mid \underline{\theta}^{(m)}\right)$ with acceptance probability

$$\alpha\left(\underline{\theta}^{(m)}, \underline{\theta}^*\right) = \min\left[1, \frac{\pi\left(\underline{\theta}^* \mid Data\right)}{\pi\left(\underline{\theta}^{(m)} \mid Data\right)} \cdot \frac{h\left(\underline{\theta}^{(m)} \mid \underline{\theta}^*\right)}{h\left(\underline{\theta}^* \mid \underline{\theta}^{(m)}\right)}\right],$$

i.e., $\underline{\theta}^{(m+1)}$ equals $\underline{\theta}^*$ or rejects otherwise the candidate $\underline{\theta}^*$, i.e., $\underline{\theta}^{(m+1)} = \underline{\theta}^{(m)}$. It can

be shown that the resulting Markov chain converges to the posterior distribution

$\pi\left(\underline{\theta} \mid Data\right)$ given certain regularity conditions [101]. Note that the posterior

distribution up to the normalization constant is also called unscaled target whereas the

proposal distribution is sometimes called candidate density [97].

Robert [74] defines the regularity condition of an irreducible, aperiodic, and

ergodic chain which property is the detailed balance condition that satisfies the kernel

$$K\left(\underline{\theta}^{(m)}, \underline{\theta}^*\right) = \alpha\left(\underline{\theta}^{(m)}, \underline{\theta}^*\right) h\left(\underline{\theta}^* \mid \underline{\theta}^{(m)}\right) + \left[1 - r\left(\underline{\theta}^{(m)}\right)\right] \delta_{\underline{\theta}^{(m)}}\left(\underline{\theta}^*\right)$$

of the Metropolis-Hastings algorithm where $r\left(\underline{\theta}^{(m)}\right) = \int \alpha\left(\underline{\theta}^{(m)}, \underline{\theta}^*\right) h\left(\underline{\theta}^* \mid \underline{\theta}^{(m)}\right) d\underline{\theta}^*$

and $\delta_{\underline{\theta}^{(m)}}$ denotes the Dirac mass in $\underline{\theta}^{(m)}$ [98].

The detailed balance condition that satisfies the kernel of a Metropolis-

Hastings algorithm does not provide practical guidance on how to decide if the

simulated Markov chain provides an adequate approximation to the posterior

distribution in order to perform statistical inference. In that context, Robert and

Casella [98] discuss three (increasingly stringent) convergence criteria: convergence

to the stationary distribution, convergence of averages, and convergence to iid

sampling.

The approaches used for monitoring of the convergence to the stationary distribution are trace plots of the Markov chain Monte Carlo simulations against the iterations or standard nonparametric tests such as the Kolmogorov-Smirnov or Kuiper. Robert and Casella [98] argue that drawing a picture is only adequate for strong non-stationarities of the analyzed Markov chain but emphasize the use of standard nonparametric tests. One may wonder what happened to the independence assumption of statistical tests and call this "statistical terrorism." Trace plots are, however, used to estimate the length of the burn-in period of the Markov chain as pointed by Albert [102]. The convergence of averages is monitored but not limited to cumulative sums charting according to Yu and Mykland [103] and an analysis of variance based within and between variance statistics according to Gelman and Rubin [104]. As for the standard nonparametric tests, statistical terrorism prevails because the cumulative sums and analysis of variance have statistical assumptions: mainly independence and to a much lesser extent the underlying distribution of the samples. The convergence to iid sampling is assessed through the degree of autocorrelation as a scale-free measure of the strength of statistical dependence using an autocorrelation function (ACF) plot that depicts the ACF and a $100(1-\alpha)\%$ confidence interval for the sample ACF. The proper thinning of the Markov chain using the lag of the ACF at which the sample ACF is below the confidence interval ensures convergence to iid sampling. Note that the assessment of the convergence criteria remains an active area of research [105, 106].

Albert [102] addresses also the issue of estimating the standard errors of the Markov chain in relation to the lack of independence of the samples. As a remedy, he

describes the batch means method. In this method, the accepted draws are subdivided into b batches for which the sample mean is calculated and the standard error approximated.

Last but not least, the autocorrelation and the acceptance rate of the simulated draws are also closely related, i.e., too low and too high acceptance rates lead to a high autocorrelation [97]. Based on empirical studies, Gregory [107], Liu [108], and Graves and Hamada [19] recommend an acceptance rate of 0.35. The YADAS software features a method to tune the acceptance rate automatically during the burn-in period by adjusting the standard deviation of the candidate density [109].

Despite the impediments of the Markov chain Monte Carlo based methods, the Metropolis-Hastings algorithm is attractive for its universal application but may be detrimental to the convergence properties of the Markov chain. Therefore, several specific samplers were derived from the very general Metropolis-Hastings algorithm such as the Metropolis algorithm, the random-walk, the Gibbs sampling, and the Slice sampler (that is actually a special case of the Gibbs sampling) [74, 97, 98, 108].

The Metropolis algorithm utilizes a symmetric candidate density, i.e., $h\left(\underline{\theta}^{(m)} \mid \underline{\theta}^*\right) = h\left(\underline{\theta}^* \mid \underline{\theta}^{(m)}\right)$ with acceptance probability

$$\alpha\left(\underline{\theta}^{(m)}, \underline{\theta}^*\right) = \min\left[1, \frac{\pi\left(\underline{\theta}^* \mid Data\right)}{\pi\left(\underline{\theta}^{(m)} \mid Data\right)}\right].$$

A special case of the Metropolis algorithm is the random-walk, the one which was actually considered by Metropolis et al. [99], in which a function symmetric around zero is used to generate a random number that is added to the most recent

31

value of the Markov chain $\underline{\theta}^{(m)}$. However, the major drawback of the random-walk sampling algorithm is the slow movement around the whole parameter space [97].

The Gibbs sampling and the related Slice sampler are special cases of the Metropolis-Hastings algorithm and result in an acceptance probability of exactly one. However, the Gibbs sampler requires the full conditional distribution for each of the blocks that contain the parameter vector. The full set of all conditional distributions may be very difficult to derive in complex system reliability models. However, the strength of the Gibbs sampling and the Slice sampler algorithm is certainly given for data analysis using regression and multilevel/hierarchical models as given in Robert and Casella [98] and Gelman and Hill [110] for simple one dimensional models.

MH algorithm or Gibbs sampling? The number of possible implementations of the Gibbs sampling or the Slice sampler is small compared to the very general Metropolis-Hastings algorithm. Gibbs sampling is claimed to converge faster but the differences are often minor or even negligible. More generally speaking, the choice depends on the problem at hand, i.e., the proposals/hierarchical decomposition. In addition, people often prefer a method that comes along with a software package such as WinBUGS (GS for Gibbs sampling) but this sampler may not be necessarily always the best implementation.

Robert [74] classifies the Gibbs sampling as local and the Metropolis-Hastings algorithm as global in the sense that the Gibbs sampling provides a better coverage of the neighborhood of the starting point and the Metropolis-Hastings algorithm explores better the complete solution domain. He recommends taking advantage of

both approaches by combining the Metropolis-Hastings algorithm and the Gibbs sampling into a hybrid sampler.

Hastings [100] discusses also a blockwise Metropolis-Hastings algorithm, the Gibbs sampling is actually a special case of a blockwise Metropolis-Hastings algorithm, that sequentially applies the algorithm to each block of parameters conditional on knowing the values of all remaining parameters that are not in that block using the transition kernel

$$P(\underline{\theta}, A) = \prod_{j=1}^{J} P_j\left(\underline{\theta}_j, A_j \mid \underline{\theta}_{-j}\right).$$

Hastings [100] also discusses the generation of random numbers from independent candidate distribution, i.e., $h\left(\underline{\theta}^{(m)} \mid \underline{\theta}^*\right) = q\left(\underline{\theta}^*\right)$. The acceptance probability of such a Markov chain shortens then to

$$\alpha\left(\underline{\theta}^{(m)}, \underline{\theta}^*\right) = \min\left[1, \frac{\pi\left(\underline{\theta}^* \mid Data\right)}{\pi\left(\underline{\theta}^{(m)} \mid Data\right)} \cdot \frac{q\left(\underline{\theta}^*\right)}{q\left(\underline{\theta}^{(m)}\right)}\right].$$

Graves and Hamada [19] apply successfully such a blockwise Metropolis-Hastings algorithm with independent candidate densities for the parameter vector in a Bayesian hierarchical estimation for the assessment of system reliability with multilevel attribute data and the allocation of resources (additional attribute data collection) in order to minimize the uncertainty of the system reliability within a fixed budget.

2.4 Liquid Rocket Engine Programmatic Metrics

Liquid rocket engine projects/programs, as for any products, are usually divided into phases such as development, production, and utilization or operation. The entirety of all phases defines the system life cycle that may be divided into the six main stages: system planning, design and development, verification and validation, production, field deployment, and disposal [14]. Similar stages are used in the space industry [111, 112].

The system planning stage is concerned about the mission operational concept that is based on customer needs and market competition analyses. In this stage, an incremental decision-making process, as introduced in Section 2.1.1, is followed to define the key project drivers such as the performance that is first order related to enabling technologies and the time (schedule or consequently the IOC) required for technology maturation as well as their reliability and affordability. The scopes of the remaining system life cycle stages are now consequences and include the following activities. The design and development stage addresses the design of the product, matures the required technologies, and establishes manufacturing capabilities. The main two project milestones (the preliminary and critical design reviews) are part of the design and development stage. The verification and validation stage includes design verification and process validation. The design verification is based on a test plan that includes the number and types of tests, the number of hardware sets foreseen, the test operational conditions, the acceptance criteria, the explicit definitions of failures, and any other related elements in order to verify the inherent reliability of the system. The process validation checks the capability of the selected

manufacturing processes, the adequacy of the defined integration steps, and the effectiveness and efficiency of the implemented control plans in order to assure that the inherent reliability of the system does not degrade during the production phase. Note that the hardware sample size requirements may be higher for the process validation than for the design verification. The production stage is started once the qualification or certification is announced and include the classical activities manufacturing, assembly, integration, and test. In addition, lean production initiatives are usually started. The field deployment stage utilizes the system in which preplanned product improvement (P3I) may be started. The disposal stage terminates the system life cycle.

Throughout the system life cycle, program managers are concerned about the balance of the project management trilemma elements performance (quality), schedule, and cost at acceptable level of project risks. The performance is associated with enabling technologies that must be matured if not available at the beginning of the program. The technology maturation not only determines the final system reliability but also drives mainly the development schedule (IOC) and as a consequence the development and production cost. The customer view of the project management trilemma is, however, not on the required performance levels because they are expected to be met or even to be exceeded. The customer is rather concerned about the IOC, which constraints the schedule, the reliability, and the affordability which can be easily deduced from the project management trilemma elements. As already stated above, the decisions about these customer concerns are made in the system planning stage and are of utmost importance for the overall program success.

Therefore, Sections 2.4.1, 2.4.2, and 2.4.3 provide further insights on the key liquid rocket engine project/program metrics.

2.4.1 Performance, Technology, and Development Duration

The key performance metrics – thrust level, specific impulse, run duration, propellant mixture ratio, weight of the engine system at burnout, geometric envelope – are determined by space transportation system optimization subject to trajectory and minimum payload capability constraints [3]. Note that space transportation system optimization is not within the scope of this research. Interested readers are referred to Krevor [43]. The derived liquid rocket engine performance requirements are, however, closely related to the enabling technologies of the piece parts and subassemblies that must withstand the operational challenges of the selected thermodynamic cycle, i.e., high specific impulse requirements promote staged combustion cycles whereas medium specific impulse and high thrust requirements endorse gas generator cycles. The envelope (geometric size) may also impact the choice. In any case, the availability or maturation of enabling technologies, independent from the thermodynamic cycle, must be assessed in the decision-making because of the impacts on affordability, reliability, and IOC that drive as a consequence the operational effectiveness [113].

A study performed by Emdee [4] provides typical development durations of cryogenic booster / main stage and upper stage liquid rocket engines which range from nine to 11 years for booster / main stage engines and six to eight years for upper stage engines, respectively. He also includes amelioration programs for upper stage engines that range from one to five years. The figures provided by Emdee can be

36

substantiated with an assessment performed by Meisl [10], who provides a range of eight to ten years for the booster engines. In another study, Emdee [5] assesses the development durations for LOx/kerosene booster / main stage and upper stage liquid rocket engines ranging from three to ten years for booster / main stage engines and four to ten years for upper stage engines, respectively. Table 2-1 summarizes typical liquid rocket engine performance and schedule metrics. Emdee's concluding statement on both studies is that the development durations (schedule) have not significantly reduced over the last 40 years.

Table 2-1:   Liquid Rocket Engine Performance and Schedule Metrics

Engine name	SSME	F-1	J-2	RL10	LR87	LR91
Vacuum thrust, kN	2174	7643	1023	73	2353	460
Specific impulse, s	452.9	304.1	425.0	444.4	298.0	314.0
Chamber pressure, MPa	21.55	6.77	5.38	3.21	5.70	5.70
Weight, kg	3177	8444	1567	138	2055	572
Duration, y	9	8	6	3	4	4

In order to put Emdee's concluding statement about the development durations (schedule) into the performance and technology maturation perspective, Meisl [10] discusses a typical test program that is required to mature the enabling technologies for high performance system. The test program follows the classical TAAF cycle with the three distinct phases for eliminating failure modes: fundamental modes, repeat modes, and quality control modes. Meisl points out that significant development duration reductions may be feasible if a technology maturation program is preceding the actual flight engine development which eliminates the fundamental failure modes. Another development duration reduction approach is to use extensively existing technologies; the RS-27 is a prominent example. The development duration of the RS-27 liquid rocket engine was only one year [5] but one needs to note that the RS-27

engine is a derivative from the H-1 and MB-3 Block III engines, whereupon the H-1 used existing technologies of the MB-1, MB-3, and X-1 engines [114].

Besides the performance that is closely linked with the development duration (schedule) due to the technology maturation and technology program drivers, Hamaker [8] identifies also non-technical variables that impact the development duration. These are requirement stability, funding stability, team experience, number of prime contractors, number of customers, and international involvement.

2.4.2 Reliability, Test Plans, and the Lack of Guidance

Wasserman [115] defines reliability as the probability of a product performing its intended function over its specified period of usage, and under specified operating conditions, in a manner that meets or exceeds customer expectations. The probabilistic aspect of reliability is assured through modeling techniques such as reliability planning and specification, allocation, prediction, Failure Mode Effects and Criticality Analysis (FMECA), Fault Tree Analysis (FTA) [14] or Probabilistic Risk Assessment (PRA) for safety related issues [15]. The main intended functions of a liquid rocket engine are to provide thrust and to generate specific impulse. The period of usage is specified in terms of design starts and design life. Table 2-2 lists initial engine design and mission requirements of realized liquid rocket engine systems [9]. The operating conditions at piece parts and subassembly level are determined by the thermodynamic cycle that is selected in the space transportation system optimization. Typical thermodynamic cycles are pressure-fed, expander, gas generator or staged combustion. Note that the thermodynamic cycle with the system induced internal load levels is closely related to the performance levels and the design requirements.

Table 2-2:   Liquid Rocket Engine Design and Mission Requirements

Engine name	SSME	F-1	J-2	RL10	LR87	LR91
Design starts	55	20	30	20	12	12
Design Life, s	22700	2250	3750	4500	1980	2700
Missions w/o Overhaul	55	1	1	1	1	1
Mission Starts	1	1	1 [1] 2 [2]	2	1	1
Mission nom. time, s	520	165	380 [1] 150 [2] 350 [2]	700	165	225

[1] first hot firing
[2] restart

Reliability engineers associate the design starts and design life with the notion of a reliable life requirement which is typically not the case for rocket scientists because the classical safety factor approach is used in the design process. Therefore, the data given in Table 2-2 are only of qualitative use and they cannot be associated to an inherent reliability requirement. Advanced probabilistic engineering analysis codes and physics-of-failure models exist to evaluate the reliable life, but both require design details that are not available during early design tradeoffs. In addition, some of the failure mechanisms as well as their combination are still subject for further research, e.g. accumulated cyclic strain (ratcheting) superimposed with creep and reduction-oxidation of the materials during operation [116].

At least there is a common understanding among reliability engineers and rocket scientists that the reliability must be built into the design of liquid rocket engine piece parts and subassemblies using modern reliability engineering methods, but as pointed out by Sackheim [16] the reliability confidence-building game is to test, test, test, and then do more testing. Koelle [117] provides the empirical evidence of the confidence-building game by relating the number of hot-fire tests with the

mission reliability, i.e., the larger the test scope is the higher the mission reliability becomes.

Unfortunately, the general trend is to reduce the number of tests due to lack of funds and possibly overconfidence of the decision-maker. Emdee [5] identifies even a negative trend in the flight success rates (mission reliability) as a consequence of the test scope reduction. But how many tests are then enough? No industry or government standard exists [20].

Emdee [5] suggests a test program of 400 tests and 40,000 seconds of accumulated hot-fire time spread over 15 engine hardware sets. Pempie and Vernin [21] recommend a test program of 150 tests and 50,000 seconds of accumulated hot-fire time but leave out a number for the required engine hardware. Wood [118] reports that 183 tests with 18,945 seconds spread over eight plus four rebuild engine hardware sets were sufficient for the qualification of the RS-68 liquid rocket engine. Greene [119] assumes a similar test program for the J-2X. Therefore, how many tests are enough? The question remains unanswered, but Emdee [4] makes a point:

> *"The lack of guidance can be frustrating to vehicle manufactures since engine development can be one of the largest expenses ... Unfortunately, despite the significant expense allocated to engine testing, the historical record shows that propulsion system still account for over 50 percent of the launch vehicle failures."*

One must review the hot-fire test strategies that were used in the previous liquid rocket engine programs in order to understand the frustration. Initially, a formal

reliability demonstration was required for the F-1 and J-2 liquid rocket engines. This was followed by what is called a Design Verification Specifications (DVS) approach for the SSME, and the latest evolution is an objective based variable test/time philosophy for the RS-68 liquid rocket engine [10, 118]. In short, one may provocatively state that any mathematical justification for the scope definition of a test program was sacrificed in favor of program cost/price savings and development schedule reductions.

In that context, Meisl [10] argues that a formal reliability demonstration for the SSME would have required 20 more hardware sets with the associated increase in terms of development cost/price and the needed time to perform the tests. But is this the truth? Emdee [5] reports 2,805 tests and 252,958 seconds of accumulated hot-fire time for a formal reliability demonstration of 0.99 at 50% confidence for the F-1 whereas Biggs [120] reports 726 tests and 110,253 seconds of accumulated hot-fire time for the DVS approach for the SSME with a reported reliability of 0.984 [117]. The information just provided is, however, biased because the actual number of tests that was needed to attain the flight readiness was only 1081 for the F-1 [121]. Considering the thrust size of the F-1, it is still the highest thrust engine with a single combustion chamber, and the additional 350 hot-fire tests may become relative if the reliability numbers for the F-1 and SSME are compared.

The derivation of the reliability test scope for the SSME may be rather illustrative when discussing test plans and demonstrated reliability levels because the hot-fire test program for the SSME was defined by a highly respected manager who set a criterion of 65,000 seconds of accumulated hot-fire time that would qualify the

41

engine for flight without any mathematical justification. The number is, however, derived as follows: Take 40 (derived from military aircraft business) and multiply it with the nominal SSME mission time of 520 seconds to result in 20,800 seconds. Add an extra conservatism and multiply that by three (for the three SSME on the Space Transportation System) to arrive at 62,400 seconds. Take this number and round it up to 65,000 seconds [114].

The lack of guidance was recognized and expressed in an Air Force guideline (RM2000) and the DoD "Total Quality Management Initiatives," which dictate that contractors shall elevate reliability to the equal status with performance and cost [122]. O'Hara also reports that the Advanced Launch System programs have specified quantitative reliability levels at engine level, i.e., an R-by-C level of R99C90. In response to that requirement, Pugh [123] describes a reliability demonstration technique that should have been applied to the Space Transportation Main Engine (STME). The technique is based on the binomial law for zero-failure test plans and is complemented with the Crow/AMSAA model in case of failures during the development testing. In that model, Pugh addresses also equivalent full duration hot-fire tests using a simplified version, the conditional probability of a shutdown given a failure had not occurred prior to shutdown was ignored, of the original work of Lloyd and Lipow [124]. The adequacy of the Lloyd and Lipow model was demonstrated using the H-1, F-1, and J-2 liquid rocket engines as well as the SSME in Worlund et al. [125]. Note that the outgrowth of the STME is the RS-68 liquid rocket engine that should have been tested according to the defined reliability demonstration technique. However, the actual test program of the RS-68 was much

smaller in scope and had 183 hot-fire tests with 18 failures (versus 230 hot-fire tests that should have been performed without a failure) due to budgetary constraints. An official reliability figure has not been published, but it should be certainly below the R99C90 requirement.

By now, it should be evident that the reliability confidence-building game is test, test, test, and then do more testing, but no guidance exists and frustration prevails. Gut feeling and educated guess are the two resorts that determine the scope of test plans. The required test schedule and the mission reliability are only results and not input variables to size the scope of a liquid rocket engine hot-fire test plan.

2.4.3   Affordability and the Denial of the Facts

The affordability assessment for liquid rocket engines is a blend of art and science according to Hammond [1], but the space industry has been constantly trying to improve the accuracy of the cost estimates since the 1990s. However, the situation has not changed and was addressed by NASA during the 7th Annual NASA Project Management Challenge held in 2010. According to Butts [126], the following statement describes best the cost situation for NASA projects:

> *"WASHINGTON - NASA can land a spacecraft on a peanut-shaped asteroid 150 million miles away, but it doesn't come close to hitting the budget target for building its spacecraft, according to congressional auditors. NASA's top officials know it and even joke about it."*

A cost overshoot was also experienced recently for the RS-68 liquid rocket engine development for which roughly 180 million dollar [economic condition (e.c.)

2010] was spent in addition due to fail-fix efforts [118]. This amount may seem low but corresponds to a cost overrun of 45% and is based on the official price figure!

NASA [127] defines affordability as an engineering process or management discipline which assures that the final system, program, project, product, or service can be delivered within the budget constraints previously established while still meeting all approved requirements (Note that the word "system" is used in the remaining discussion on affordability but may refer to program, project, product, or service). Therefore, affordability expresses the amount of money (budget) that the purchaser is able to pay. The affordability assessment is part of decision-making that takes place usually in the system planning stage, i.e., the pre-Phase A [127] or Phase 0 [112]. Note that the financial measures Life Cycle Cost (LCC) and Total Cost of Ownership (TCO) are usually used in affordability assessments.

The LCC includes all of the costs that are accrued during a defined system life cycle spanning from requirement development through design, verification, production, operation and maintenance until recycling or disposal. The TCO refers to cost that covers the acquisition, the operation, and the maintenance of a particular system [128]. Therefore, the TCO is a subset of the LCC. This distinction is only important for systems that are publicly funded for the development and commercially operated during the utilization phase of the system life cycle, which is usually the case for space transportation systems.

A considerable body of literature related to LCC and TCO exists, but many of the materials were written by practitioners and may lack academic rigor. The exceptions are the textbooks by Dhillon [129], Fabrycky and Blanchard [130], and

Blanchard and Fabrycky [131] that treat the LCC and TCO tools and techniques in more depth.

Among the numerous publications, Gupta and Chow [132], Asiedu and Gu [133], and Christensen et al. [134] are noteworthy because the authors summarize 40 years of LCC literature, describe the mechanisms of life cycle engineering and costing relevant for complex system development, and analyze the techniques used in life cycle costing. The conclusions of the latter authors consider the 12 steps in the LCC analysis process as state-of-the-art that were defined by Blanchard and Fabrycky [131]. Among the 12 steps, the most important ones are to specify a system life cycle, to develop a Cost Breakdown Structure, to select a cost model for analysis and evaluation, to develop a cost profile and summary, to conduct a sensitivity analysis, and to evaluate feasible alternatives and select a preferred approach.

The development of a Cost Breakdown Structure is the most important task because it provides a top-down and bottom-up view of the cost structure over the complete system life cycle. Blanchard and Fabrycky [131] provide an example of such a Cost Breakdown Structure that includes research and development cost, production cost, operation and support cost, and retirement cost at the first breakdown level. A Work Breakdown Structure is usually converted into a Cost Breakdown Structure in practice. Although there is no general rule on how to generate a Work Breakdown Structure, the MIL-HDBK-881 or preferably a process-product oriented approach might be used as guideline [135].

The next crucial step in the LCC analysis process is the establishment of costs for each of the categories that are defined in the Cost Breakdown Structure. This step

is especially critical for LCC analysis that are performed in early system life cycle phases when available input data is limited and uncertainty is the highest due to the lack of detailed component design definitions. Note that about 60 percent of the LCC are committed at the end of the system planning and conceptual design stage (corresponding to the pre-Phase A or Phase 0), roughly 80 percent are committed by the end of the system definition, and 95 percent are committed after the full-scale system development [136]. Figure 2-2 depicts this fact graphically [131].
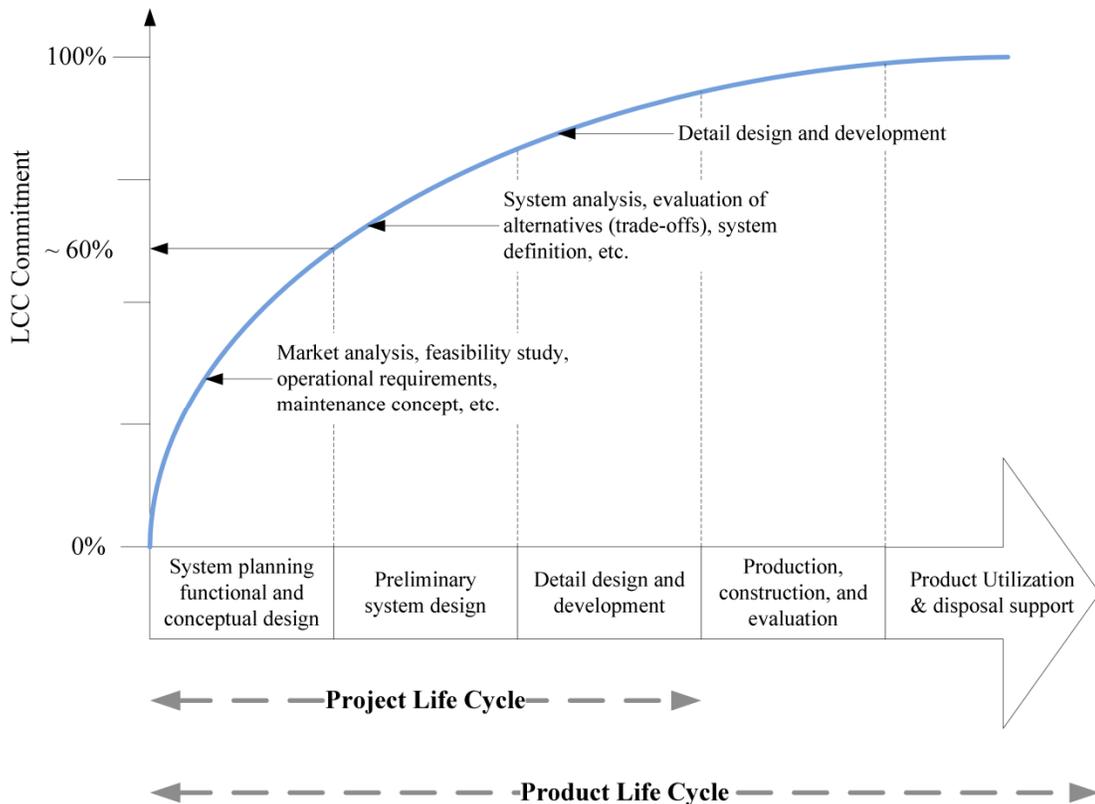
Figure 2-2:   LCC Commitment versus System Life Cycle

The most widely used cost/price estimating techniques in the space industry are the grass-root, the analogy, and the parametric approach. The application of these techniques is, however, related to the design level of the system. NASA [127] and Blanchard and Fabrycky [131] suggest the parametric approach in early project

phases, the analogy in intermediate project phases, and the grass-root costing for the production phase (see Figure 2-3 [131]). The parametric cost estimation technique is the more advantageous approach because several design and programmatic parameters can be used. The analogy models are usually limited to a single design parameter such as the thrust, capacity, or weight for liquid rocket engines [129].
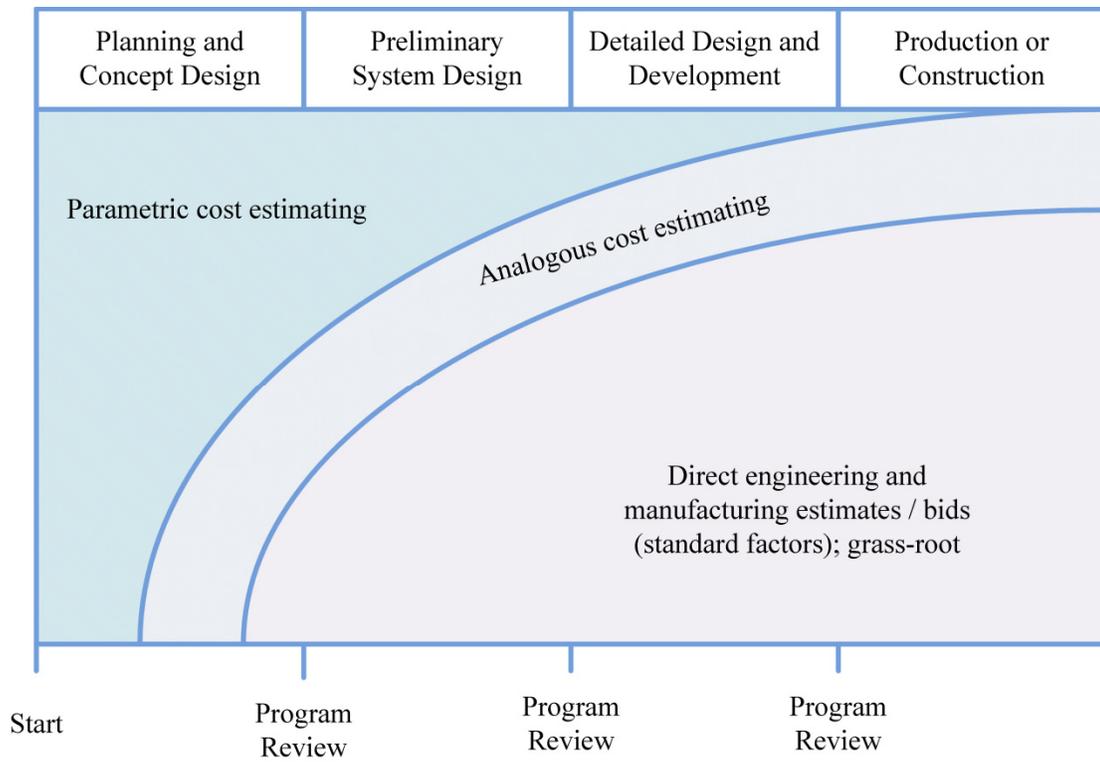
Figure 2-3:   Cost Estimation Techniques linked to System Life Cycle Phases

The grass-root costing is based on Cost Estimation Relationships (CER) using detailed, accurate capital and operational cost data. Certainly, this cost estimation technique may seem to be the most preferable, but a high degree of accuracy remains elusive in the aerospace business since the data suffer from incompleteness and small sample sizes. Unfortunately, many program managers trust grass root costing more than parametric costing because the latter results usually in higher cost estimates which they think are not competitive. Therefore, managers use grass-root costing as

justification for a lower cost but assume at the same time that the company's organization operates like their grass-root cost model dictates it. This is certainly not true, but managers deny this fact. The result is usually a major cost overrun [137].

Not shown in Figure 2-3 but also used are modern cost management systems such as Activity-Based Costing, Just-in-Time Costing, Target Costing, and Strategic Cost Management. However, Activity-Based Costing and Just-in-Time Costing have limited use during early product life cycle phases since they require the bill of activity and bill of material as input. Both inputs cannot easily be generated for conceptual designs. In later system life cycle phases, these methods are superior and should be used. Target Costing is used for the cost allocation process and might be used during the conceptual design phase. Finally, the Strategic Cost Management is focused around the value chain and can be used in conjunction with Activity-Based Costing but not during the life cycle cost assessment in early system life cycle phases [138].

Inflation and escalation are also important variables in affordability assessments. Inflation is in general the rise in the level of prices of goods and services in an economy over a period of time. It is mainly influenced by the money supply of governments by setting the interest rates. The scarcity of a certain material due to political disruption impacts also the inflation rate which is of particular interest for the aerospace business. Rising energy costs are influencing also the level of inflation which will become even more dominant in the next decades to come. Escalation is mainly linked to salary creep and grade creep due to salary upgrades and career advancements which is in particular true for aerospace companies with typically low personnel turnover [136].

The purchasing power of money is another important aspect in comparative affordability assessments among different economic markets. The appropriate metrics is the Purchasing Power Parity (PPP) which is the long term view of the value of money. This metric is important if prices from different markets are compared, e.g. the prices from U.S. liquid rocket engines versus the one from European liquid rocket engines. Although it is common practice in the liquid rocket engine business, it is absolutely wrong to use the currency exchange rate if prices of liquid rocket engines are compared. Currency exchange rates reflect the short term view of the value of money which is not the case for system life cycle times of up to 30 years [139].

Historically, the need for better cost/price estimates is first addressed by Meisl [10] who assesses the main LCC contributor of expendable and reusable space transportation systems by comparing the liquid rocket engines F-1, J-2, and the SSME. He also points out that both the data scarcity and the infrequency of development programs constitute one of the major difficulties of the LCC assessment of liquid rocket engines. The key elements that Meisl discusses are drivers for the development and production cost, the tendency of a platykurtic slightly left-skewed bell shaped development budget versus development time profile, and three testing periods that are linked to the costly elimination of failure modes (fundamental modes, repeat modes, and quality control modes). In that context, Meisl argues qualitatively about a possible development cost/price reduction if the fundamental failure modes can be avoided during the full-scale development program due to prior testing of Integrated Subsystem Test Bed (ISTB) or breadboard engine demonstrator. In addition, lessons learned indicate also the strong recommendation to test components

early in the development program, as it is the case for the DVS approach, in order to minimize the development cost. The main point of Meisl's paper is the strong relation between the schedule, reliability, and affordability that was already identified above but which cannot be stressed enough.

Meisl [10] also provides similar elements for the production cost and identifies the influences of technical parameters on the production cost that were based on a Rocketdyne parametric production cost model. Several years later, Meisl [12] includes facility cost and reliability and risk cost into his discussion on LCC.

One of Meisl's [140] last publications about the affordability subject describes the future of design integrated cost modeling with focus on process-oriented parametric cost models and quantifiable uncertainties for technical, programmatic, and cost/price parameters. One of the process-oriented parametric liquid rocket engine cost models is described in Lee [141]. The model requires specific inputs in terms of labor effort, material cost, and support cost. In order to provide credibly these elements, the bill of material is needed, but this is usually not available during early project phases. Therefore, the process-oriented parametric cost modeling approach is not adequate for concept tradeoffs and is not further discussed here. The second focus on the quantification of the uncertainties for technical, programmatic, and cost/price parameters is of much higher interest because these elements should be part of an integrated and balanced evaluation of performance or equivalently schedule and reliability. The influence of the uncertainties about the programmatic and cost/price parameters is derived from the hot-fire test plan, the failure mode description, the number of available hot-fire test facilities, and the cost model input

parameters. However, Meisl does not provide a quantitative link between these influences and in particular for reliability, program duration, and development cost/price. This link does still not exist!

Hamaker [8] supports this strong statement by suggesting a research direction that should address the project success as a function of the amount of testing. In addition, Hunt [7] points out his experience on the development costing work which he performed for the RS-84 and J2-X liquid rocket engines. In particular, he mentions the difficulties he had about the TAAF cycle assumptions and the impact on the development cost/price. Therefore, Hamaker and Hunt confirm the strong statement about the lack of existence of a link between reliability, program duration, and development cost.

Joyner et al. [6] reaffirm the strong dependency of the development cost on the TAAF assumptions and provide the following figures: only two percent is spent on the initial conceptual design effort, 15 percent is spent on the engineering design and analyses, and ten percent is spent on the qualification, reliability demonstration, and certification. The majority – more than 70 percent – is spent on the elimination of failure modes. They conclude that the key development cost/price drivers are the number of hot-fire tests and number of hardware sets required to complete the test program.

Joyner et al. [6] also review the main cost models used in the liquid rocket engine industry: PRICE-H® (Parametric Review of Information for Cost and Evaluation – Hardware) [142], SEER-H® (System Evaluation & Estimation of Resources – Hardware) [143], TRANSCOST® (Handbook of Cost Engineering for

Space Transportation Systems) [117], and the Liquid Rocket Engine Cost Model (LRECM) that is implemented in NAFCOM® (NASA/Air Force Cost Model) [13, 144]. A similar analysis is also given in Harwick [145].

Except for the LRECM, the main model parameter of these cost tools is the engine weight. Multipliers such as complexity, engineering experience, technical factors, and design maturity are then used to increase the fidelity of the models. The general tendency of the weight based tools is that a greater weight results in more development costs. The development cost of liquid rocket engines behave, however, opposite for a fixed design, i.e., increasing the weight usually reduces the development cost and vice versa. Since the LRECM is not using the weight as cost input parameter, it can be seen as an original approach to remedy the classical strong dependency of cost models on a weight based CER. The details about the LRECM can be found in Joyner et al. [6].

2.5 Reliability Growth

Lloyd and Lipow [124] introduce the subject of reliability growth as the relationship between reliability prediction (a future, projected reliability number) and reliability estimation which is estimated directly from current and previous observations. The reliability estimate generally increases during the development. However, the rate of growth, its adequacy, and the level of attainment at the end of the test program is a concern. The true reliability increases incrementally through a series of redesigns of the failure-producing piece parts. The magnitude and frequency of the redesigns may vary and depends on the type of subassembly.

Broemm et al. [146] define reliability growth as the improvement in a reliability parameter over a period of time due to changes in the product design or the manufacturing process. The changes in the product design are typically associated with an iterative TAAF cycle.

The three major areas in the field of reliability growth are planning, tracking, and projection, which can be directly derived from the definition given in Lloyd and Lipow [124], i.e., the planning is linked to the forecast of the level of attainment of the reliability metric at the end of the test program, the tracking is the reliability estimation of current and previous observations, and the projection is the prediction of the final reliability metric following the implementation of corrective actions to the observed failure modes.

The models that are most widely used are based on the empirical Duane method [147] and the US Crow/AMSAA analytical model [146]. However, both methods are based on the underlying assumption that the failure intensity function [or rate of occurrence of failures (ROCOF)] follows a non-homogeneous Poisson process (NHPP) [71]. Other models which are not based on the NHPP assumption are extensively reviewed in Hall [148].

Liquid rocket engine developments are predestinated for the iterative TAAF cycle. Codier [147] applies successfully the Golovin and the empirical Duane models to the test data of the F-1 and J-2 liquid rocket engines whereas Williams [30] reports a failure in applying the US Crow/AMSAA model for the SSME because the model initially estimated an increase of the MTBF (indicating reliability growth) but the system reliability declined towards the end of the testing profile although overall

testing experience would have suggested an increase in the system reliability. Why does the reliability growth fail for the SSME but succeeds for the F-1 and J-2?

The reason is linked to the hot-fire test philosophy that has evolved over time as already pointed out in Section 2.4.2 on the liquid rocket engine test plans. Historically, liquid rocket engine hot-fire testing profiles followed well the mission profile, i.e., the operational loads during ground tests were similar to the loads seen during the flight acceptance and actual flight. The DVS and the objective based variable test/time philosophy include extreme load points to demonstrate robustness and design margin which introduce a significant difference between the testing profile and the mission profile. Crow [149] and Krasich [31] also observe this concern in other industries and propose either the grouping of the failure times in intervals from which the classical US Crow/AMSAA model parameters are estimated or the physics-of-failure and cumulative damage models to normalize the data from which the parameters for the Duane or US Crow/AMSAA model can be estimated. Note that Safie and Fuller [150] applied successfully the Crow/AMSAA model to track the reliability growth of the Space Shuttle Main Engine reliability using data that was strongly adjusted.

Chapter 3:   Mathematical Formulation of the Risk-informed satisficed Decision-Making Methodology

The RISDM methodology is based on a constrained multiobjective satisficing problem formulation using the weighted sum method, i.e., the fitness function is the sum of normalized objectives, in which the objectives are defined as the areas of concern affordability, reliability, and IOC that are influenced by hot-fire tests, the decision variables, which are allocated to various system integration levels. The areas of concern create a conflict because they are contradicting; therefore, tradeoffs must be made to reach a satisficed solution. A genetic algorithm is used to generate vectors of decision variables that define the sets of possible solutions for a given liquid rocket engine system alternative which are influenced by stakeholder targets, weights, and uncertainties about their areas of concern. The proposed vectors of decision variables are actually used to determine the levels of attainment for the measures of effectiveness for the areas of concern. These are based on interrelated models that include non-technical and technical factors such as failure mechanisms, differences between mission profile and testing profiles, TRL, MRL, TAAF cycle assumptions based on the newness of the system that needs to be developed, system performance, product life cycle, design scope and environment, and team experience. The subsequent sections describe these interrelated models. Simple problems provide sensitivities and model validations.

3.1   Definition of Testing Profiles as Multiples of the Mission Profile

The *mission profile* or main life cycle is defined by several hot-fire events that take place during the service life of a liquid rocket engine. It may include acceptance hot-fire test(s), a possible engine ground start hold-down with launch commit criteria abort hot firing, and a single flight mission hot firing duration (or several flight hot firings in case of a reusable main stage engine) or multiple re-ignitions in case of in-flight operation. The mission profile is applicable during the production phase.

The *testing profiles* are composed of a potpourri of hot firings that may be multiples of the mission profile, completely different in terms of hot-fire duration and operational load points in order to demonstrate margin and design robustness or the combination of both. The hot firings are also executed at various system integration levels, i.e., component, subsystem, and system level. The complete set of all testing profiles defines the development hot-fire test plan. Consequently, the testing profiles are applicable during the development phase.

The *testing profiles* testing approach is not limited to liquid rocket engine systems. Gas turbine engine developments use the concept of Accelerated Mission Testing (AMT), which is an extension to the DVS that was developed by NASA for safety critical and high reliability systems [151]. In that context, the focus of AMT is to concentrate the testing on the failure-inducing agents in proportion to the mission profile. Similar testing profiles are also applied to main battle tanks, light armored vehicles, and mine-resistant ambush-protected vehicles [149].

3.2 Constrained Multiobjective Satisficing

The mathematical basis for the constrained multiobjective satisficing decision-making is described in Wierzbicki [60], who introduces the achievement scalarizing function that preserves the order of preferences among the sets of attainable measure of effectiveness within an area of concern. The normative target-based decision-making and the related truncated exponential utility function, which is also order preserving, are used as an alternative to the achievement scalarizing function in this research.

In case of several areas of concern or objectives, the multiobjective problem may be formulated either as a single-objective problem, in which the objectives are collected into a single fitness function [152], or a multiobjective problem [153]. In this research, the weighted sum method, a single-objective problem formulation, is used to define the fitness function. This function is maximized using a genetic algorithm because the multiobjective problem is convex, i.e., the generated solutions are Pareto-optimal [153-155]. A comparison between the SOGA using Palisade's Evolver® software [156] and the well-known and frequently used NSGA-II using the SolveXL® software [157] is given in Section 4.4.3.

Figure 3-1 displays the flowchart of the implemented constrained multiobjective satisficing approach. Note that the implementation follows the basic genetic algorithm except for Steps 1 and 2 (see Sections 3.2.1 and 3.3 for details concerning the objective weights and the specific models that are used to determine the measures of effectiveness for the three objectives affordability, reliability, and IOC, respectively).
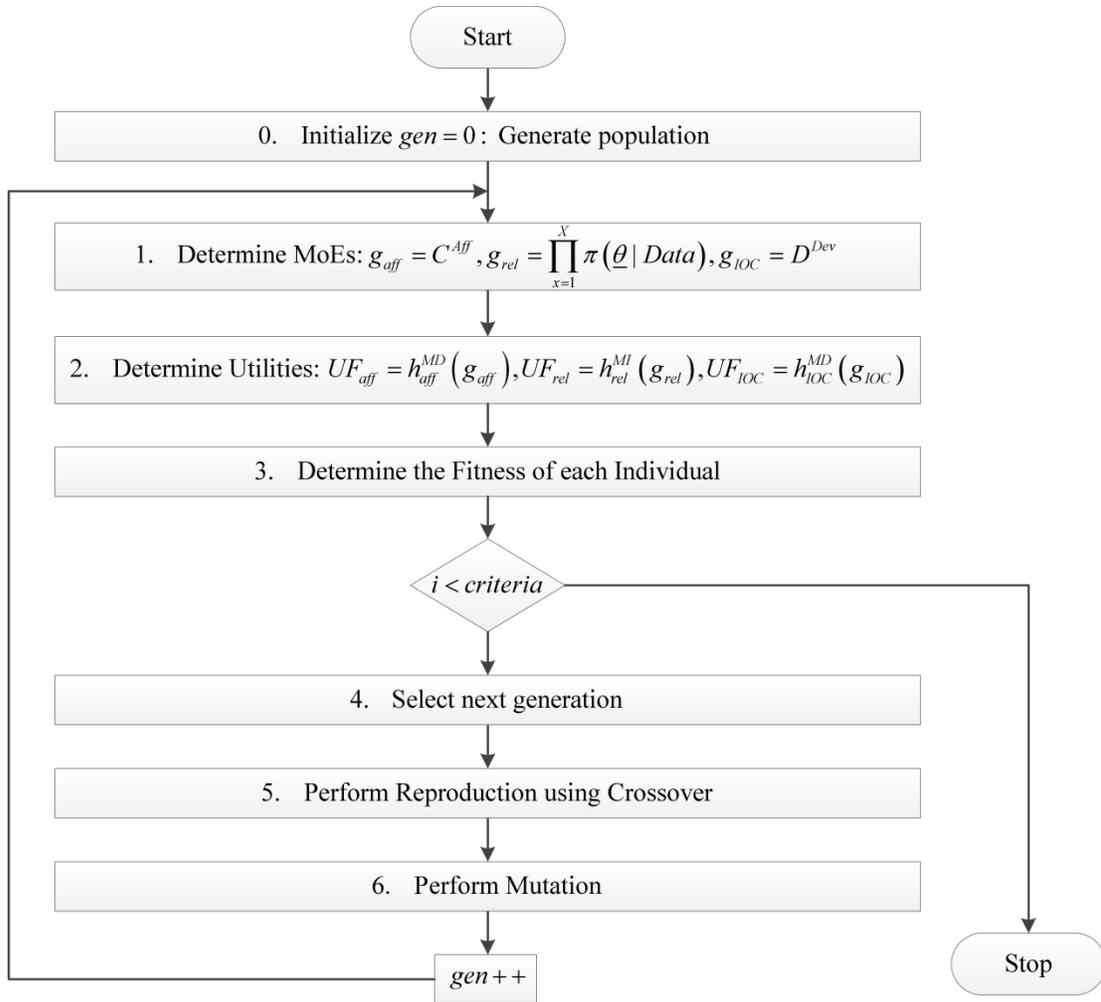
Figure 3-1:   Constrained Multiobjective Satisficing Approach

3.2.1 Mathematical Formulation

The constrained multiobjective satisficing is formulated as a constrained multiobjective optimization problem. The decision variables are the hot-fire tests, which are allocated to multilevel system integration levels.

The multilevel system integration levels are associated with component, subsystem, and system levels and are denoted as hot-fire test groups, which are indicated with subscript i. Within each hot-fire test group, there are different hot firing durations or testing profiles that are indicated with subscript j. The number of

hot-fire tests with testing profile j in hot-fire test group i is denoted as NFC_{ij}^{TP} for which the upper and lower bounds are denoted as $NFC_{ij}^{TP_{LB}}$ and $NFC_{ij}^{TP_{UB}}$, respectively. The aggregation of all hot-fire tests defines the overall hot-fire test plan, denoted as EQM^{TP}.

The weights for the normalized objectives are determined by means of the Analytic Hierarchy Process (AHP), which will be discussed in the following paragraphs. The objective functions that preserve the order preference for each area of concern are modeled using truncated exponential utility functions, i.e., if the utility score should increase as the measure of effectiveness increases, then $UF = h^{MI}$; otherwise, the utility score should decrease as the measure of effectiveness increases, and $UF = h^{MD}$. Eq. (3.1) exhibits the monotonically increasing and monotonically decreasing truncated exponential utility functions.

$$
\begin{aligned}
UF = h^{MI}\left(g, \gamma^{Eff}, LB, UB\right) &= \begin{cases} \dfrac{1 - e^{-\gamma^{Eff}(g-LB)}}{1 - e^{-\gamma^{Eff}(UB-LB)}} & \gamma^{Eff} \neq 0 \\[2ex] \dfrac{g-LB}{UB-LB} & otherwise \end{cases} \\[4ex]
UF = h^{MD}\left(g, \gamma^{Eff}, LB, UB\right) &= \begin{cases} \dfrac{1 - e^{-\gamma^{Eff}(UB-g)}}{1 - e^{-\gamma^{Eff}(UB-LB)}} & \gamma^{Eff} \neq 0 \\[2ex] \dfrac{UB-g}{UB-LB} & otherwise \end{cases}
\end{aligned}
\tag{3.1}
$$

where g is the measure of effectiveness, γ^{Eff} is the effective risk coefficient, LB is the lower bound, and UB is the upper bound. The shape of a truncated exponential utility function, which determines the utility score for a given measure of effectiveness, is influenced by the effective risk coefficient γ^{Eff}. The magnitude of

the effective risk coefficient γ^{Eff} is determined by means of the utility-probability duality. Note that the levels of attainment for the measures of effectiveness are determined by specific models that are described in Sections 3.3.1 to 0.

In general, let M be the number of areas of concern that are relevant to the problem. In this research, $M = 3$. For area of concern m, let g_m be the function that estimates the measure of effectiveness as a function of the design alternative A^n and the hot-fire test plan EQM^{TP}; let the range $\{LB, UB\}$ be the lower and upper bounds on the measure of effectiveness; let UF_m be the relevant utility function, the shape of which is defined by the relevant effective risk coefficient γ_m^{Eff}; and let w_m be the relative weight of that utility function. Let FF be the fitness function, which is the weighted sum of the utility functions. Note that all $w_m \in [0,1]$ and $\sum_{m=1}^{M} w_m = 1$. Eq. (3.2) exhibits the mathematical formulation.

$$
\begin{aligned}
\text{Maximize} \quad & FF = \sum_{m=1}^{M} w_m UF_m \\
\text{subject to} \quad & UF_m = h_m\left(g_m, \gamma_m^{Eff}, LB_m, UB_m\right) \quad m = 1,2,...,M; \\
& g_m = f_m\left(A^n, EQM^{TP}\right) \quad\quad m = 1,2,...,M; \\
& g_m \geq LB_m \quad\quad\quad\quad\quad\quad m = 1,2,...M; \\
& g_m \leq UB_m \quad\quad\quad\quad\quad\quad m = 1,2,...M; \\
& NFC_{ij}^{TP_{LB}} \leq NFC_{ij}^{TP} \leq NFC_{ij}^{TP_{UB}} \quad i = 1,2,...,I; \\
& \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad j = 1,2,...,J_i.
\end{aligned}
\tag{3.2}
$$

Analytic Hierarchy Process

Saaty [158] developed the Analytic Hierarchy Process (AHP) in the early 1970s and applied it to numerous risk-informed decision-making problems [71]. Because of its widespread use, it is assumed that the AHP application is known.

Note that the AHP is preferred over the minimum number of judgments methods, such as SWING or simple multiattribute rating technique (SMART) [159], in the frame of this research in order to remedy behavioral biases in a decision-making process that involves new technologies and risk-averse decision-makers.

Utility-Probability Duality

The normative target-based decision-making framework is used to express the uncertainty of the decision-maker's preference about an associated measure of effectiveness for each area of concern. In the context of this research, the areas of concern are assumed to be independent, which, as a consequence, requires the formulation of a specific utility function for each of the area of concern.

The uncertainty about the actual performance in each area of concern for a specific alternative can, however, be expressed, based on the knowledge of the decision-maker about the design alternative, as a range in which each measure of effectiveness should fall and a target for each area of concern. The targets correspond to the programmatic requirements. Given this limited information, the decision-maker's uncertainty for each area of concern is modeled as a subjective probability distribution. The challenge then is to find an appropriate utility function.

The utility-probability duality [55, 56] provides a framework to find appropriate utility functions because it represents the decision-maker's preference, the decision-maker's information about the uncertainty, the decision-maker's target for the specific area of concern using the aspiration-equivalent, and as a consequence the decision-maker's risk attitude, i.e., risk-averse, risk-neutral or risk-seeking, that is expressed by the effective risk coefficient γ^{Eff}.

What follows is the determination of the effective risk coefficient γ_m^{Eff} given the decision-maker's targets and the uncertainty for each area of concern using the utility-probability duality. For the sets of alternatives and areas of concern m, let $\hat{g}_m$ be the decision-maker's target (aspiration-equivalent), let $\{LB_m, UB_m\}$ be the lower and upper bound, and let $F(g_m; \alpha_m, \beta_m, LB_m, UB_m)$ be the general Beta cumulative distribution function that describes the decision-maker's uncertainty (the distribution parameters α_m and β_m can be determined based on the bounds using the method of quantiles [160]). Finally, let $uf_m = d(UF_m)/dg_m$ be the derivative of the corresponding utility function. Then, the utility-probability duality is defined by Eq. (3.3) that is solved for the effective risk coefficient γ_m^{Eff}.

$$F(\hat{g}_m) = \int_{LB_m}^{UB_m} F(g_m; \alpha_m, \beta_m, LB_m, UB_m) uf(g_m, \gamma_m^{Eff}, LB_m, UB_m) dg_m \qquad (3.3)$$

Because truncated monotonically increasing exponential utility and general Beta cumulative distribution functions are used, an expression in analytic form of the right-hand side of Eq. (3.3) can be found which is then used to solve numerically for the effective risk coefficient γ_m^{Eff} for each of the m areas of concern, i.e., in general,

$$\left. v(g)u(g) \right|_{LB}^{UB} - \int_{LB}^{UB} v'(g)u(g) dg$$

where $v(g)$ is the general Beta cumulative distribution function, $v'(g)$ is the general Beta probability density function, $u(g)$ is the truncated monotonically increasing exponential utility function, and $\{LB_m, UB_m\}$ are the bounds of the measures of

effectiveness or equivalently the bounds of the general Beta distribution that reflect the uncertainty of the decision-maker.

The utility-probability duality as described in [55, 56] applies only to monotonically increasing utility functions but the adaptation to truncated monotonically decreasing exponential utility functions is accomplished by symmetry, i.e., $uf^{MI}\left(\gamma^{Eff}\right)=1-uf^{MI}\left(-\gamma^{Eff}\right)=u^{MD}\left(\gamma^{Eff}\right)$.

3.2.2 Numerical Results

Target-based Effective Risk Coefficient

Let assume that the stakeholder uncertainties about the three areas of concern affordability (development cost), reliability, and IOC with corresponding targets were elicited as given in Table 3-1. [The data is normalized to millions of monetary units (MMU) in order to protect the proprietary nature of the data.] The targets for the affordability and IOC are based on expert opinions using historical data [4, 5] or Bayesian estimation for the reliability. The min values for the three areas of concern correspond to the lower natural bound, i.e., zero, whereas the max values are defined by an assumed maximum affordability, the natural bound of one for the reliability, and an assumed IOC, respectively. The percentiles for the affordability are based on percentage values that are subtracted and added to the target. The percentiles for the reliability are based on the two-sided credibility interval (TBCI) using the historical data that are given in [80]. Finally, the percentiles for the IOC are based on a minus three standard deviation using the data given in [4, 5] and an upper bound that includes an assumed positive slack, i.e., one year with regard to the IOC.

Table 3-1:   Stakeholder Uncertainties and Targets

Areas of concern	Min	Max	0.05 percentile	0.95 percentile	target
Affordability, MMU	0	2000	930	1350	1035
Reliability, -	0	1	0.9663	0.9974	0.956
IOC, y	0	13	7.50	12.00	10.9

Equation (3.3) is then used to assess the influence of the target $\hat{g}_m$ on the effective risk coefficient γ_m^{Eff} for the areas of concern affordability, reliability, and IOC, respectively. The calculated risk coefficients γ_m^{Eff} are depicted in Figure 3-2. Note that the abscissae for affordability and IOC have been normalized. By looking at Figure 3-2 and considering the elicited ranges, the stakeholder's or decision-maker's risk attitudes for the area of concern affordability and IOC are always risk-averse whereas the risk attitude for the area of concern reliability changes from risk-averse to risk-seeing, i.e., positive effective risk coefficients correspond risk-averse, zero to risk-neutral, and negative to risk-seeking risk attitudes, respectively. E.g., setting the target to the upper bound and expect to attain this high level is a risky (risk-seeking) endeavor considering the actual levels of reliability for liquid rocket engines.
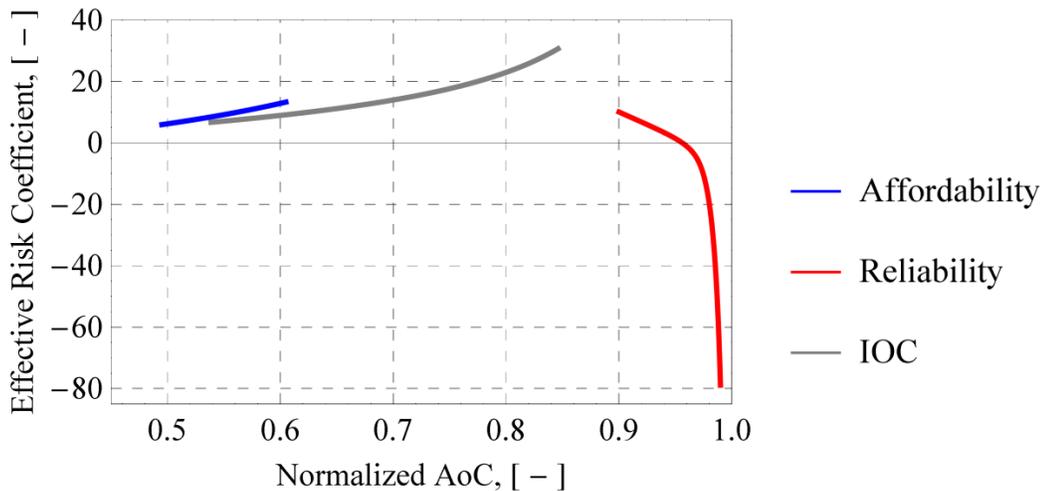
Figure 3-2:   Targets versus Effective Risk Coefficient

Satisficing results with different objective weights, decision-maker's uncertainty, and Penalty Functions

Impact of Objective Weights

The impact of the objective weights that are used to define the fitness function is studied using a mixture design with 11 runs. Table 3-2 lists the design matrix including the measures of effectiveness for the objectives affordability, reliability, and IOC. Each solution was found by running the SOGA with the parameters as listed in Table 3-3. Figure 3-3 depicts the resulting Pareto-optimal satisficed solutions and the genetic evolution progress for the various weight settings.

Table 3-2:   Mixture Design Matrix

Weights	Affordability	Reliability	IOC	Satisficed Solution		
				Aff	Rel	IOC
Case-I	1	0	0	714	0.9297	6
Case-II	1/2	1/2	0	943	0.9545	7.825
Case-III	0	1	0	2693	0.9888	28.75
Case-IV	2/3	1/6	1/6	760	0.9348	6
Case-V	1/3	1/3	1/3	896	0.9507	6.6
Case-VI	0	2/3	1/3	1344	0.9725	9.925
Case-VII	1/2	0	1/2	739	0.9118	5
Case-VIII	0	1/2	1/2	1342	0.9720	9.85
Case-IX	1/3	0	2/3	714	0.9273	6
Case-X	0	1/3	2/3	1338	0.9720	9.8
Case-XI	0	0	1	744	0.9177	5.1

Table 3-3:   Parameters of the SOGA used in Palisade's Evolver®

Population size	50
No. of generations	Progress based
Cross-over probability	0.5
Cross-over type	Arithmetic
Selector	Weighted average
Mutation probability	0.1
Mutator	Cauchy mutation
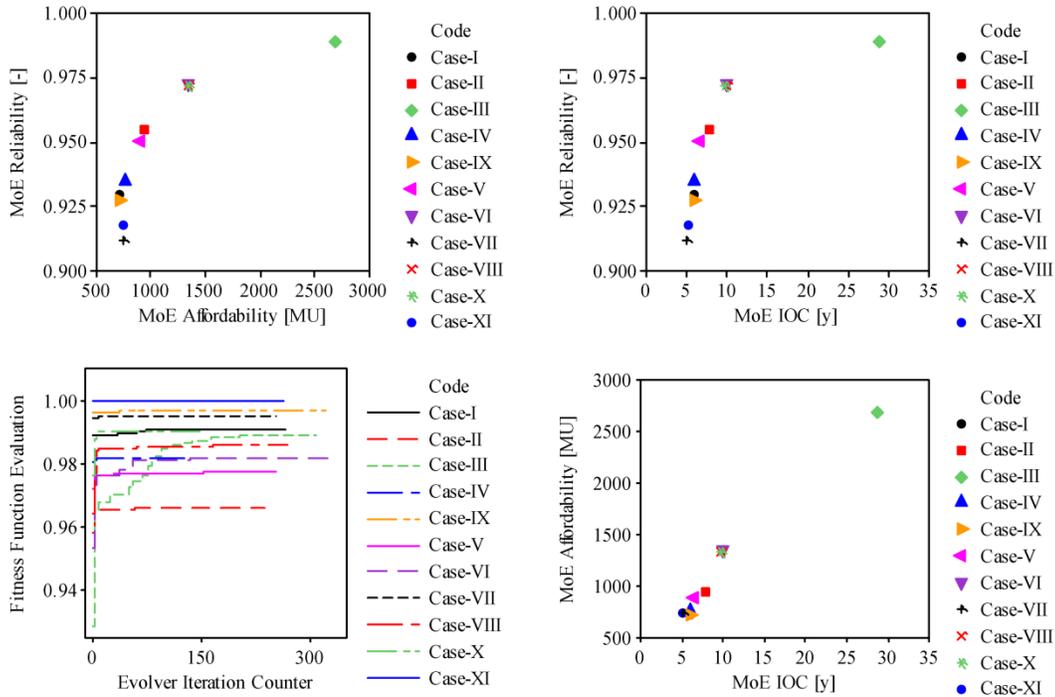
Figure 3-3:   Pareto-optimal Satisficed Solutions for different Weight Settings

Impact of Decision-maker's Uncertainty

The decision-maker's uncertainty about the objectives influences the shapes of the utility functions; therefore, the fitness function evaluation is impacted. Based on Figure 3-2, the pertinent objective that changes the risk attitude is the reliability, which is further used to study the impact of the decision-maker's uncertainty on the fitness function evaluation with equal weights for the three objectives affordability, reliability, and IOC. Figure 3-4 depicts the resulting Pareto-optimal satisficed solutions and the genetic evolution progress for three cases: risk-averse, risk-neutral, and risk-seeking, respectively.

66
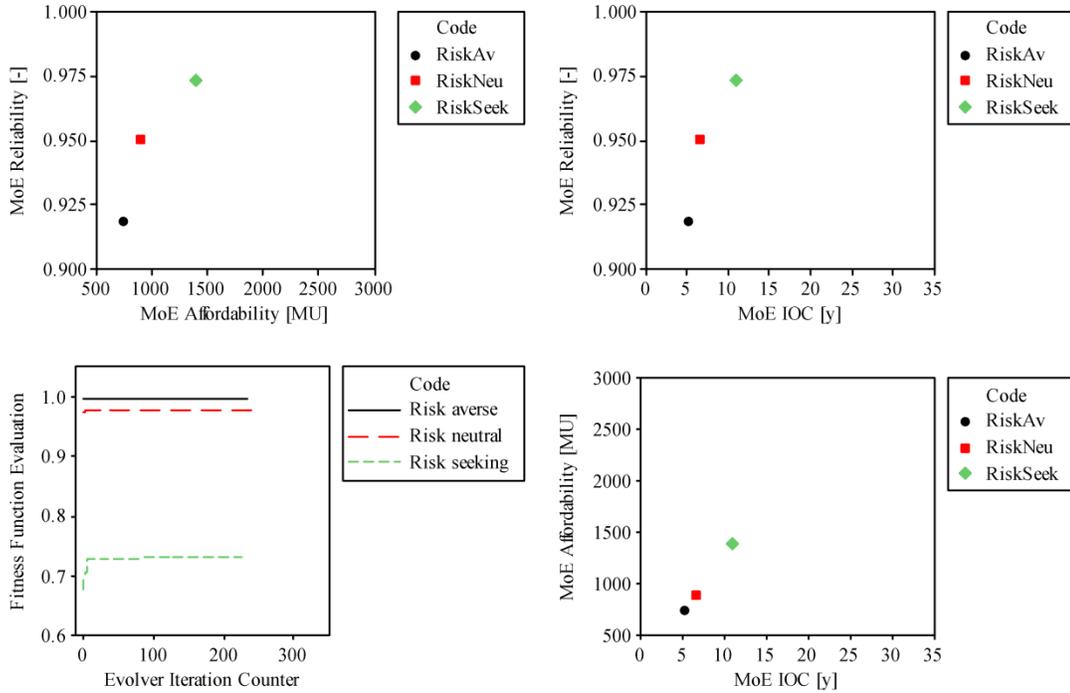
Figure 3-4: **Pareto-optimal Satisficed Solutions for the Objective Reliability with Risk Averse, Risk Neutral, and Risk Seeking Risk Attitudes**

3.3 Areas of Concern: Modeling Affordability, Reliability, and Initial Operational Capability

3.3.1 Modeling Affordability

Parametric Cost/Price Model

NATO [161] defines affordability as the degree to which the LCC of an acquisition program is in consonance with the long-range investment and force structure plans of a specific administration. In the context of this research, "in consonance" means to deliver a liquid rocket engine system that meets the customer's needs at available budget (annual funding availability) with sustainable opportunities throughout the

system life cycle. The specific administration is either NASA or the ESA member states in case of the liquid rocket engines.

The total cost, or equivalently the LCC, that is accrued throughout a typical liquid rocket engine system life cycle may be split into the classical portions: development, production, and operations and support. The development costs are associated with the technology maturation, the design and development, and the design verification by means of a test plan that include the TAAF cycles.

The Liquid Rocket Engine Cost Model (LRECM), originally developed by Rocketdyne, estimates the development and production cost. It is implemented in the NASA/Air Force Cost Model (NAFCOM®) Contractor Version [13]. Details about the LRECM evolution may be found in Meisl [10], Meisl [12], and Joyner et al. [6]. Note that the fundamental model equations are proprietary and access to NAFCOM® is given upon the acceptance of a nondisclosure agreement.

Not specific to this research but generally important is the consideration of international economy theory if two different economic markets, i.e., the prices of liquid rocket engines, are compared [139, 162]. The Purchasing Power Parity (PPP) of the price level at a specific economic condition (e.c.) is used as conversion factor. The application of economic theory to the LRECM results is validated using proprietary European cost data of existing liquid rocket engines.

Joyner et al. [6] provide the most recent description of the LRECM. The LRECM was specifically created as an alternative approach to the classical weight-based Cost Estimating Relationship (CER) models typically used in early project phases to support the decision-making concerning liquid rocket engine conceptual

design choices such as thrust, chamber pressure, engine thermodynamic cycle, technology readiness level, engineering and production processes attributes, design for producibility, etc. The input parameters are listed in Table 3-4. However, important limitations of the LRECM are the applicability of the model to NASA project phases C and D and the lack of a specific cost model that estimates the execution of a predefined hot-fire test plan, which will be developed using the RAIV strategy (see Section 3.3.2).

Table 3-4:   LRECM Input Parameters

Parameter	Development cost	Production cost	Test hardware cost
Development environment	X		
Manufacturing environment	X	X	X
Manufacturing readiness level	X	X	X
Design scope	X		
Team experience	X		
Engine cycle		X	X
Producibility		X	X
Vacuum thrust, kN		X	X
Chamber pressure, bar		X	X

Schankman [163], OSD [164], ECSS [112], and Macret [165] provide the links between project phases, design scope expressed as Technology Readiness Level (TRL) [166], team experience, Integration Readiness Level (IRL) [165, 167, 168], and Manufacturing Readiness Level (MRL) [164] as displayed in Figure 3-5. By that means, the limitation of the LRECM to the project phases C and D is no longer applicable. Note that the verbal descriptions of the design scope and team experience correspond to the factor levels used in the LRECM.

Project Phases	0+A	B	C/D	E
	MDR	PRR SRR	PDR	CDR	QR AR ORR FRR

New Technology / Unfamiliar Exp.						
	New Product / Mixed Exp.					
		New Design / Normal Exp.				
			Extensive Modif. / Extensive Exp.			
				Simple Modif. / Extensive Exp.		

IRL		1	2	3	4	5	
TRL	3	4	5	6	7	8	9	
MRL	1 - 3	4	5	6	7	8	9	10

Concepts	Technology Dev.	System Dev. & Demo.	Lean
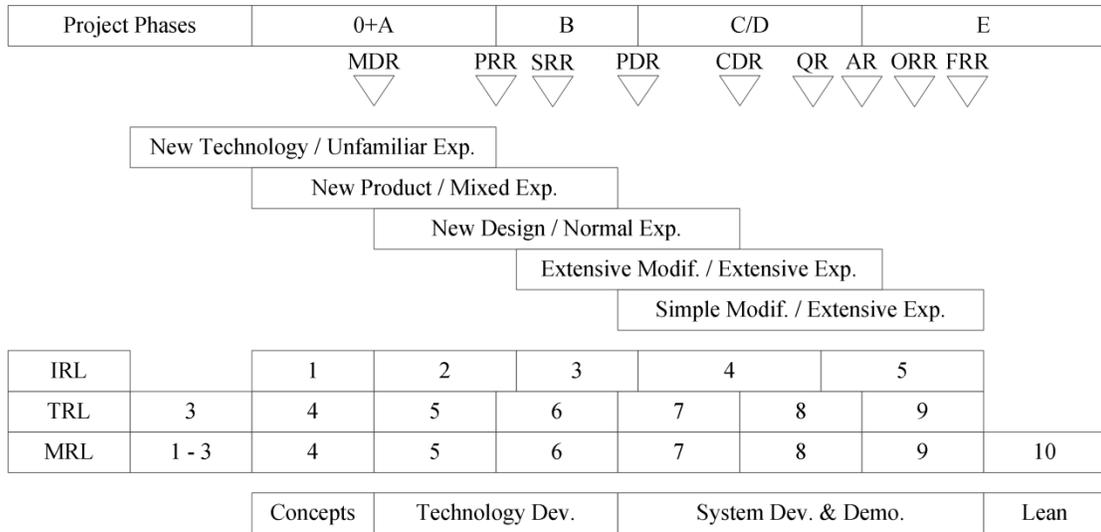
Figure 3-5:   Project Phases, Design Scope, Team Experience, IRL, TRL, and MRL

The Cost Estimating Relation (CER) for the development dost C^{Dev} is expressed as the sum of the design and development cost C^{DD}, the test hardware cost C^{THW}, and the cost to execute the testing profiles (test plan) C^{TP}. Eq. (3.4) exhibits the fundamental development CER.

$$C^{Dev} = C^{DD} + C^{THW} + C^{TP} \tag{3.4}$$

C^{DD} and C^{THW} are determined with the LRECM, and specific models are used to estimate the number of hardware sets needed to complete the hot-fire test plan (see Sections 3.3.2), whereas the test execution cost C^{TP} depends upon the construction and maintenance cost of the test facilities and the costs of the hot-fire tests performed, i.e., $C^{TP} = C^{TP}_{fix} + C^{TP}_{var}$.

Let N^{TF}_r be the number of test facilities of type r, let C^I_r be the construction cost of a test facility of type r, and let C^M_r be the annual maintenance cost of a test

facility of type r. The fixed test execution cost C_{fix}^{TP} for the test facilities is determined by Eq. (3.5).

$$C_{fix}^{TP} = \sum_r^R N_r^{TF} \left(C_r^I + D_r^{TP} C_r^{TFM} \right) \tag{3.5}$$

where D_r^{TP} is the test facility occupation duration of test facility of type r (see Section 0). Note that C_r^I is a one-time cost that may or may not be associated with the development cost and may include a complete construction of a new test facility, a major upgrade of an existing test facility, or simple modifications. Initial installment or upgrade cost may be required if existing test facility capabilities are no longer adequate to support the required testing profiles boundary conditions. For example, the thrust level of a liquid rocket engine may exceed existing facility capabilities.

The variable test execution cost C_{var}^{TP} in a test facility of type r includes the cost of operating the facility as well as the costs of the fuel, oxidizer, and consumables used during the test. Let C_{irs}^{DP} be the direct personnel cost of operating a test facility of type r for test campaign s for hot-fire test group i. This cost depends upon the time and personnel required to install the test hardware in the test facility, conduct a number of hot-fire tests with different testing profiles, and dismount the hardware from the test facility. Let C_{irs}^{Fu} be the cost of the fuel used for test campaign s for hot-fire test group i in a test facility of type r. Let C_{irs}^{Ox} be the cost of the oxidizer used for test campaign s for hot-fire test group i in a test facility of type r. These propellant costs depend upon the hot firing duration, the propellants used (determined through thrust, vacuum specific impulse, and propellant mixture ratio),

71

and the per-unit cost of fuel and oxidizer. Let C_{irs}^{Co} be the cost of the consumables used for test campaign s for hot-fire test group i in a test facility of type r. This cost is a constant and includes the cost of gases such as nitrogen or helium that are used for purging or venting operations during or in between the hot-fire tests. Given these quantities, the variable test execution cost C_{var}^{TP} for test campaign s for hot-fire group i in the test facility r is determined by Eq. (3.6).

$$C_{var}^{TP} = \sum_{i=1}^{I} \sum_{r=1}^{R} \sum_{s=1}^{S_{ir}} \left(C_{irs}^{DP} + C_{irs}^{Fu} + C_{irs}^{Ox} + C_{irs}^{Co} \right) \tag{3.6}$$

with

$$C_{irs}^{DP} = N_{irs}^{DP} \ R_{irs}^{DP} \ WY_{irs}^{DP} \ D_{irs}^{TP}$$

$$C_{irs}^{Fu} = \sum_{j=1}^{J_i} \frac{F_{ijrs}^{TP}}{Isp_{ijrs}^{TP} \ g_0 \left(MR_{ijrs}^{TP} + 1 \right)} FD_{ijrs}^{TP} \ c^{Fu}$$

$$C_{irs}^{Ox} = \sum_{j=1}^{J_i} \frac{F_{ijrs}^{TP} MR_{ijrs}^{TP}}{Isp_{ijrs}^{TP} \ g_0 \left(MR_{ijrs}^{TP} + 1 \right)} FD_{ijrs}^{TP} \ c^{Ox}$$

$$C_{irs}^{Co} = \sum_{j=1}^{J_i} C_{ijrs}^{Co}$$

where N^{DP} is the number of direct personnel operating the test facility, R^{DP} is the hourly rate, WY^{DP} are the yearly working hours of one direct personnel, D^{TP} is the test facility occupation duration, F^{TP} is the thrust level, Isp^{TP} is the vacuum specific impulse, MR^{TP} is the propellant mixture ratio, g_0 is the gravitational constant, FD^{TP} is the firing duration, c^{Ox} is the specific propellant cost for the oxidizer, and c^{Fu} is the specific propellant cost for the fuel. Note that F^{TP}, Isp^{TP}, MR^{TP}, FD^{TP}, and C^{Co}

depend on the testing profiles j associated with a specific hot-fire test group i that is performed in test campaign s in test facility of type r.

Finally, the affordability is modeled using Eq. (3.7). Note that the operations and support (O&S) costs, although part of the affordability, are not taken into account in this research because these costs are only of importance for reusable liquid rocket engines, which are not considered herein.

$$C^{Aff} = C^{Dev} + C^{Prod} \qquad (3.7)$$

where C^{Dev} is the cost associated with the development as given in Eq. (3.4) and C^{Prod} is the accumulated production cost associated with a defined product life cycle and rate of production, which is estimated with NAFCOM®.

3.3.2 Modeling Reliability as Reliability-As-an-Independent-Variable Strategy

Bayesian Multilevel Testing Profiles Aggregation

The Bayesian estimation of the multilevel testing profiles aggregation is based on a Bayesian multilevel attribute data aggregation method [17-19, 93]. The application of the Bayesian multilevel attribute data aggregation method is, however, not applicable because different hot-fire test conditions are present in the testing profiles. In order to remedy the inapplicability and apply the Bayesian multilevel attribute data aggregation method, the concept of an Equivalent Mission (EQM) is used. Note that an EQM of one simply corresponds to a single trial under the Bayesian multilevel attribute data aggregation method.

An EQM normalizes the different testing profiles with the mission profile by taking into account the *challenges* to which piece parts and subassemblies of liquid rocket engines are exposed during the operational start-up, steady, and shutdown states. The performance-requirement failure model is most applicable to liquid rocket engines in which the two dominant *failure*-inducing *agents* are *stress* and *time*, which trigger *stress-increased* and *strength-reduced failure mechanisms* [71]. The two failure mechanisms may be interrelated but certainly do not contribute equally to the well-known failure modes of liquid rocket engines; consequently, a weighting must be regarded. Mathematically, the EQM is defined in Eq. (3.8). The first term reflects the stress-increased failure mechanism, and the second term reflects the strength-reduced failure mechanism.

$$EQM^{TP} = \zeta \frac{NFC^{TP}}{NFC^{MP}} + (1-\zeta) \frac{CFD^{TP}}{CFD^{MP}} \qquad (3.8)$$

where ζ is the weighting factor of the challenges that trigger the two failure mechanisms, NFC^{TP} is the number of hot firing cycles associated with the testing profiles with the corresponding cumulative hot firing duration CFD^{TP}, and NFC^{MP} is the number of hot firing cycles associated with the mission profile with the corresponding cumulative hot firing durations CFD^{MP}. Note that the weighting factor depends on the thermodynamic engine cycle as well as the pressure and thrust level, i.e. high pressure high thrust level liquid rocket engines are more vulnerable for strength-reduced failure mechanism whereas lower level systems are more vulnerable for stress-increased failure mechanisms. Therefore, the EQM covers all possible liquid rocket engine system alternatives in a single modeling approach.

The different testing profiles are usually performed at various system integration levels, i.e., component, subsystem, and system level, which also define the test configurations. Within each test configuration, different hot firing durations FD may be defined. To account for both different test configurations and hot firing durations, hot-fire test groups are denoted with subscript i, and testing profiles are denoted with subscript j. Eq. (3.9) exhibits the introduction.

$$EQM_{ij}^{TP} = \zeta_{ij} \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1 - \zeta_{ij}\right) \frac{NFC_{ij}^{TP} FD_{ij}^{TP}}{CFD^{MP}} \tag{3.9}$$

The flight mission hot firing duration is the reference hot-fire test time. Any hot firing testing profile that is less than the full flight mission hot firing duration must be weighted with respect to this reference; otherwise, any system reliability estimate would be seriously biased [123, 125]. Lloyd and Lipow [124] derived a probabilistic model to estimate an appropriate weighting factor w_{ij}; this is given in Eq. (3.10). Note that the weighting factor w_{ij} is associated with the system level test configuration, denoted with subscript sys, because the *real* mission operational loads and a full flight mission duration FMD can be exerted only on the system level due to the limitations of the component and subsystem test facilities, i.e., limited pressure levels and firing durations.

$$w_{sys,j}^{TP} = \frac{p_{sys,j} + \varepsilon\left(1 - p_{sys,j}\right)}{p_{sys,FMD} + \varepsilon\left(1 - p_{sys,FMD}\right)} \tag{3.10}$$

where $p_{sys,j}$ is the probability of failure occurrence during the start-up and steady state up to firing duration $FD_{sys,j}^{TP}$, $p_{sys,FMD}$ is the probability of failure occurrence for the full flight mission duration, and ε is the conditional probability of failure

occurrence during the shutdown state given that no start-up and steady state failure had occurred prior to the shutdown.

Eq. (3.11) exhibits the introduction of the weighting factor w_{ij} into the EQM definition. The parameters ε and $p_{sys,j}$ are estimated using Bayesian estimation with the likelihood function that is given in Eq. (3.12) and uniform prior distributions.

$$EQM_{ij} = \zeta_{ij} \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1 - \zeta_{ij}\right) \frac{NFC_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP}}{CFD^{MP}} \tag{3.11}$$

$$L\left(Data \mid \underline{\theta}\right)_{Sys} = \prod_{j=1}^{J_{Sys}} L\left(Data \mid \underline{\theta}\right)_{Sys,j} \tag{3.12}$$

with

$$L\left(Data \mid \underline{\theta}\right)_{sys,j} = \frac{NFC_{sys,j}^{TP}!}{NFC_{sys,j}^{TP\langle S\rangle}! \, NFC_{sys,j}^{TP\langle F\rangle}! \prod_{k=1}^{j} NFC_{sys,j,k}^{TP\langle F\rangle}!} \cdot$$

$$\left[\left(1 - p_{sys,j}\right)\left(1 - \varepsilon\right)\right]^{NFC_{sys,j}^{TP\langle S\rangle}} \left[\left(1 - p_{sys,j}\right)\varepsilon\right]^{NFC_{sys,j}^{TP\langle F\rangle}} \cdot$$

$$\prod_{k=1}^{j} \left(p_{sys,k} - p_{sys,k-1}\right)^{NFC_{sys,j,k}^{TP\langle F\rangle}}$$

where $\left[\left(1 - p_{sys,j}\right)\left(1 - \varepsilon\right)\right]^{NFC_{sys,j}^{TP\langle S\rangle}}$ is the corresponding ordinary failure probability,

$\left[\left(1 - p_{sys,j}\right)\varepsilon\right]^{NFC_{sys,j}^{TP\langle F\rangle}}$ is the shutdown state failure probability, and

$\prod_{k=1}^{j} \left(p_{sys,k} - p_{sys,k-1}\right)^{NFC_{sys,j,k}^{TP\langle F\rangle}}$ is the failure probability for failures that could have

occurred in the hot firing interval with durations $\left(FD_{sys,k}^{TP} - FD_{sys,k-1}^{TP}\right)$, $NFC_{sys,j}^{TP}$ is the

total number of hot-fire tests, $NFC_{sys,j}^{TP\langle S\rangle}$ is the number of successful hot-fire tests,

$NFC_{sys,j}^{TP\langle F\rangle}$ is the number of shutdown hot-fire test failures, and $NFC_{sys,j,k}^{TP\langle F\rangle}$ is the number of hot-fire test failures that can occur in the hot firing interval.

The strength-reduced failure mechanism is influenced by the operational loads during the steady state operation; therefore, different levels of failure acceleration effects must be regarded by means of an acceleration factor AF_{ij}^{TP} [169]. Note that more research is, however, required in the field of advanced physics-of-failure models for liquid rocket engine piece parts and subassemblies and the aggregation of these individual AF into a single AF that reflects the specific hot-fire test group i. A study of combustion chambers is described by Schwarz et al. [116]. Eq. (3.13) exhibits the introduction of the acceleration factor AF_{ij}^{TP} into the EQM definition.

$$EQM_{ij}^{TP} = \zeta_{ij}\frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1-\zeta_{ij}\right)\frac{NFC_{ij}^{TP}AF_{ij}^{TP}w_{ij}^{TP}FD_{ij}^{TP}}{CFD^{MP}} \qquad (3.13)$$

Analogous to the binomial model, EQM successes, denoted by superscript S, and EQM failures, denoted by superscript F, are defined for each of the system integration levels i. Equations (3.14), (3.15), and (3.16) exhibit the relevant mathematical expressions for the EQM trials, EQM failures, and EQM successes, respectively. Note that the overall hot-fire test plan is then defined as the sum of all EQM_i, i.e., $\sum_{i=1}^{I}EQM_i^{TP}$ and denoted as EQM^{TP}.

$$EQM_i^{TP} = \sum_{j=1}^{J_i}EQM_{ij}^{TP} \qquad (3.14)$$

$$EQM_i^{TP\langle F\rangle} = \sum_{j=1}^{J_i}\left(\zeta_{ij}\frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1-\zeta_{ij}\right)\frac{NFC_{ij}^{TP}AF_{ij}^{TP}w_{ij}^{TP}FD_{ij}^{TP\langle F\rangle}}{CFD^{MP}}\right) \qquad (3.15)$$

$$EQM_i^{TP\langle S\rangle} = EQM_i^{TP} - EQM_i^{TP\langle F\rangle} \tag{3.16}$$

The final step in the Bayesian multilevel testing profiles aggregation is the construction of the underlying likelihood function using a functional node network that is similar to Reliability Block Diagrams (RBD) [170] and the definition of the prior distributions for the component reliabilities in order to define the unscaled posterior distribution $\pi(\underline{\theta}\,|\,Data)$ as given in Eq. (3.17).

$$\pi(\underline{\theta}\,|\,Data) \propto \prod_{i=1}^{I} \pi_i^{EQM_i^{TP\langle S\rangle}} \left(1-\pi_i\right)^{EQM_i^{TP}-EQM_i^{TP\langle S\rangle}} \prod_{i=1}^{I} \pi_i^0(\underline{\theta}) \tag{3.17}$$

where π_i is the functional node reliability at the various system integration levels and π_i^0 is the corresponding prior distribution. The functional node reliabilities are functions of the physical component (or any other lowest system decomposition level) or Common-Cause Component Group (CCCG) reliabilities of the physical system architecture that are subject to k risks or causes of failures, i.e., $\pi_i = f(\pi_{C_x})$. The functional node network defines also the fundamental hot-fire test plan, i.e., it specifies the hot-fire test configurations.

Figure 3-6 depicts a five component functional node network to demonstrate the construction of the likelihood function. Note that the approach is not limited to simple serial networks; complex serial-parallel networks are also possible [93].
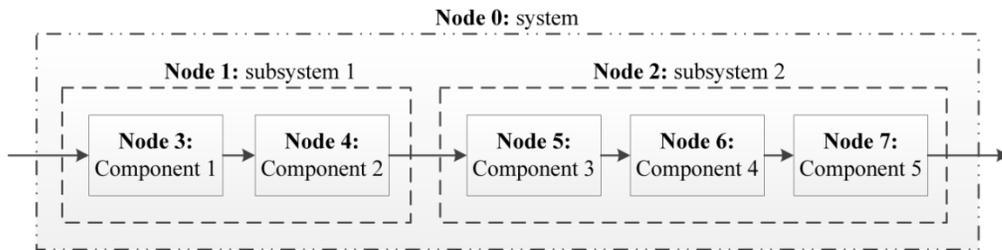
Figure 3-6:   Serial Functional Node Network

In this example, the system-level functional node reliability, denoted by Node 0, is

$\pi_0 = \pi_1 \pi_2$ with subsystem functional node reliabilities, denoted by Node 1 and Node

2, $\pi_1 = \pi_3 \pi_4$ and $\pi_2 = \pi_5 \pi_6 \pi_7$. Note that the functional node reliabilities π_3, π_4, π_5,

π_6, and π_7, correspond to the component reliabilities π_{C_1}, π_{C_2}, π_{C_3}, π_{C_4}, and π_{C_5},

respectively. The likelihood function is then found by inserting the functional

component, subsystem, and system level nodes into Eq. (3.17), i.e.,

$$
\begin{aligned}
L\left(Data \mid \underline{\theta}\right) = {}& \pi_{C_1}^{EQM_{C_1}^{TP\langle S\rangle}} \left(1 - \pi_{C_1}\right)^{EQM_{C_1}^{TP} - EQM_{C_1}^{TP\langle S\rangle}} \cdot \ldots \cdot \\
& \pi_{C_5}^{EQM_{C_5}^{TP\langle S\rangle}} \left(1 - \pi_{C_5}\right)^{EQM_{C_5}^{TP} - EQM_{C_5}^{TP\langle S\rangle}} \cdot \\
& \left(\pi_{C_1} \pi_{C_2}\right)_{subsys_1}^{EQM_{subsys_1}^{TP\langle S\rangle}} \left(1 - \pi_{C_1} \pi_{C_2}\right)^{EQM_{subsys_1}^{TP} - EQM_{subsys_1}^{TP\langle S\rangle}} \cdot \\
& \left(\pi_{C_3} \pi_{C_4} \pi_{C_5}\right)_{subsys_2}^{EQM_{subsys_2}^{TP\langle S\rangle}} \left(1 - \pi_{C_3} \pi_{C_4} \pi_{C_5}\right)^{EQM_{subsys_2}^{TP} - EQM_{subsys_2}^{TP\langle S\rangle}} \cdot \\
& \left(\pi_{C_1} \pi_{C_2} \pi_{C_3} \pi_{C_4} \pi_{C_5}\right)_{sys}^{EQM_{sys}^{TP\langle S\rangle}} \left(1 - \pi_{C_1} \pi_{C_2} \pi_{C_3} \pi_{C_4} \pi_{C_5}\right)^{EQM_{sys}^{TP} - EQM_{sys}^{TP\langle S\rangle}}
\end{aligned}
$$

Prior Distribution Choices

Section 2.3.2 discussed criticisms of the Bayesian approach related to the subjectivity

involved in the generation of the prior distributions because of the negligible,

moderate, or enormous influence on the parameter estimation. In general, two classes

of prior distributions exist: (1) minimally informative or equivalently diffuse,

noninformative, or vague and (2) informative [18].

The most common approach to define a minimally informative prior is to

apply Jeffreys' rule that may result in improper or proper distribution functions. In

case of the binomial experiment, a proper Jeffreys' prior distribution function is

given, i.e. $Be(0.5, 0.5)$ [101].

The Beta distribution function is also used to define informative prior distributions for the component level as follows. The system level Beta distribution shape parameters α_{sys} and β_{sys} are determined using the method of quantile estimates [160] that minimize

$$\left[F\left(p_{\gamma/2}\right)-\frac{\gamma}{2}\right]^2 + \left[F\left(p_{1-\gamma/2}\right)-\left(1-\frac{\gamma}{2}\right)\right]^2$$

in which the two quantiles $p_{\gamma/2}$ and $p_{1-\gamma/2}$ correspond to the predicted two-sided Bayes probability interval (TBPI) [171] or mathematically more appropriately to the TBCI of a posterior distribution [101]. Empirical data is used to calculate the required pth quantiles [4, 5, 80, 172]. Eq. (3.18) exhibits the first level Bayesian estimate of a mean predicted reliability, and Eq. (3.19) and Eq. (3.20) exhibit the lower and upper bounds of the credibility interval, respectively. The $100\left(1-\gamma/2\right)\%$ level of credibility is set to 90% that defines the 5th and 95th quantiles.

$$E\left[R_{sys}^{pred}\right] = \frac{N^{MP\langle S\rangle}+1}{N^{MP}+2} \tag{3.18}$$

$$R_{sys}^{TBCI_L} = \frac{s_M+1}{N^{MP\langle S\rangle}+1+\left(N^{MP}+2-N^{MP\langle S\rangle}-1\right)F_{1-\gamma/2}\left(\cdot\right)}$$
$$F_{1-\gamma/2}\left(\cdot\right) = F_{1-\gamma/2}\left(2N^{MP}-2N^{MP\langle S\rangle}+2,2N^{MP\langle S\rangle}+2\right) \tag{3.19}$$

$$R_{sys}^{TBCI_U} = \frac{\left(N^{MP\langle S\rangle}+1\right)F_{1-\gamma/2}\left(2N^{MP\langle S\rangle}+4,2N^{MP}-2N^{MP\langle S\rangle}+2\right)}{N^{MP}-N^{MP\langle S\rangle}+1+\left(N^{MP\langle S\rangle}+1\right)F_{1-\gamma/2}\left(\cdot\right)}$$
$$F_{1-\gamma/2}\left(\cdot\right) = F_{1-\gamma/2}\left(2N^{MP\langle S\rangle}+4,2N^{MP}-2N^{MP\langle S\rangle}+2\right) \tag{3.20}$$

where $N^{MP\langle S\rangle}$ is the number of predicted mission profile successes, N^{MP} is the number of predicted mission profile trials, and $F_{1-\gamma/2}(v_1, v_2)$ is the $(1-\gamma/2)$ quantile of the F-distribution with degree of freedoms v_1 and v_2.

The combination of both minimally informative priors and informative prior information is expressed in the form of finite mixture distributions as given in Eq. (3.21) [173].

$$f(\theta) = \sum_{l=1}^{L} \omega_l f_l(\theta \mid \underline{\eta}_l) \tag{3.21}$$

where $f_l(\theta \mid \underline{\eta}_l)$ are the population distribution functions, $\underline{\eta}_l$ is a vector of the distribution parameters for the distribution function of population l, and ω_l are the *mix parameters* with

$$\sum_{l=1}^{L} \omega_l = 1 \quad \text{and} \quad \omega_l \geq 0, \text{ for } l = 1, 2, ..., L.$$

Kleyner [82] proposes a two-component mixture distribution with the component distributions Uniform and Beta. The mix parameters are interpreted as knowledge factor expressing the similarity of a new product to the existing one and innovation factor expressing the novelty content in the new product. Eq. (3.22) exhibits Kleyner's two-component mixture distribution.

$$f(\theta; \alpha, \beta, \phi) = \begin{cases} (1-\phi) & \theta = 0 \vee \theta = 1 \\ \phi \dfrac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)} + (1-\phi) & 0 < \theta < 1 \end{cases} \tag{3.22}$$

where θ is the variable, i.e., the component reliability, ϕ and $(1-\phi)$ are the mix parameters, α and β are the shape parameters of the Beta distribution, and $\mathrm{B}(\cdot)$ is the Beta function.

Relating to the knowledge and innovation factors, Krolo [81] proposes an alternative formulation that is based on an informative Beta distribution. However, the introduction of the knowledge and innovation factors requires an adjustment of the normalization constant of a standard Beta distribution function to ensure that the total probability integrates to unity, i.e., the Eulerian integral of the first kind becomes $\int_0^1 t^{(\alpha_{C_x} \cdot \phi_{C_x})-1} \cdot (1-t)^{(\beta_{C_x}-1) \cdot \phi_{C_x}} \, dt$, which has the solution $\mathrm{B}\left[\alpha_{C_x} \phi_{C_x}, (\beta_{C_x}-1)\phi_{C_x}+1\right]$. Note that this alternative formulation was used in [28].

This research used Jeffreys' prior instead of a Uniform distribution in a finite mixture distribution because the selection of prior distributions is based on formal rules [73] and the interpretation of the mix parameters as knowledge transfer factor ϕ is similar to the knowledge and innovation factors [82]. Eq. (3.23) exhibits the two-component finite mixture distribution using the Jeffreys' prior and Beta distribution.

$$f(\theta; \alpha, \beta, \phi) = \phi \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\mathrm{B}(\alpha,\beta)} + (1-\phi)\frac{(1-\theta)^{-0.5}\theta^{-0.5}}{\mathrm{B}(\alpha,\beta)} \quad 0 < \theta < 1 \quad (3.23)$$

where θ is the variable, i.e., the component reliability, ϕ is the knowledge transfer factor, α and β are the shape parameters of the Beta distribution, and $\mathrm{B}(\cdot)$ is the Beta function.

Note that the selection of a prior distribution is used only at the lowest system decomposition level, i.e., component level, in the frame of this research. The prior

distributions of the subsystem and system level are assumed to be Uniform probability density functions, i.e., $\pi_{sys}^0 = 1$ and $\pi_{subsys_i}^0 = 1$. The implemented prior distributions on component levels are given in Eq. (3.24). The justification for the choice is deferred to Section 0.

$$\pi_{C_x}^0\left(\theta_i;\alpha_{sys},\beta_{sys},\phi_{sys}\right) = \phi_{sys}\frac{\theta_i^{\alpha_{sys}-1}\left(1-\theta_i\right)^{\beta_{sys}-1}}{B\left(\alpha_{sys},\beta_{sys}\right)} +$$
$$+\left(1-\phi_{sys}\right)\frac{\left(1-\theta_i\right)^{-0.5}\theta_i^{-0.5}}{B\left(\alpha_{sys},\beta_{sys}\right)}$$

(3.24)

where θ_i is the component level reliability, ϕ_{sys} is the system level knowledge transfer factor, $\alpha_{C_x} = \alpha_{sys}$ and $\beta_{C_x} = \beta_{sys}$ are the shape parameters of the Beta distribution, and $B(\cdot)$ is the Beta function.

The assumptions that $\alpha_{C_x} = \alpha_{sys}$ and $\beta_{C_x} = \beta_{sys}$ for the shape parameters of the Beta distribution are due to the competing risks of the system components. If a system is studied that is not following the competing risks model assumptions, the component level informative prior distribution parameters α_{C_x} and β_{C_x} are found by simulation [18]. In case of a simple series system the Beta distribution parameters are $\pi_i^0 \sim Be\left(a_{sys},\beta\right)$ with $\sum_{x=1}^{X}\beta_{C_x} = \beta$.

Predicted Test-Analyze-And-Fix Cycle Failures

The knowledge transfer factor ϕ is also used to predict the TAAF cycle failures. The level of knowledge transfer is defined by physical considerations, i.e., the power-to-weight ratio, and the expertise of the used propellants in contrast to the application of an FMEA [81], Fuzzy Model [83], pilot tests [81] or field data [81, 82].

Power is the rate at which energy is transferred, used, or transformed. In the context of liquid rocket engines, the energy equals the mass flow rate of propellants that are used to transfer chemical into kinetic energy to generate thrust. The chemical energy transfer takes place at high-temperature, high-pressure conditions that are also associated with the failure-inducing agents.

Therefore, the knowledge transfer factor on system level ϕ_{sys} is defined through the thrust and system pressure conditions that determine the adverse operational conditions in liquid rocket engines. In addition, the used propellant combination is added because new propellants add new unknown unknowns. Eq. (3.25) exhibits the mathematical formula.

$$\phi_{sys} = \left(\frac{F_{vac}^{known}}{F_{vac}^{new}} \right)^a \left(\frac{P_{cc}^{known}}{P_{cc}^{new}} \right)^b \frac{1}{I^{propellant}} \tag{3.25}$$

where F_{vac} is the vacuum thrust, P_{cc} is the main combustion chamber pressure, a and b are constants, and $I^{propellant}$ is an indicator variable. Knowledge from existing similar systems is denoted by superscript $known$, the new system of interest is denoted by superscript new, and $I^{propellant} = 1$ if the propellant is new but equals 2 if the propellant is well-known.

The method introduced by Waterman et al. [79] is used to estimate the required distribution parameters $\alpha_{sys} = \kappa_{sys}$ and $\beta_{sys} = \upsilon_{sys} - \kappa_{sys}$ but modified with a proper two-component mixture distribution, i.e., Eq. (3.22). In addition, the method also requires a (κ, υ) parameterized Beta distribution instead of a $Be(\alpha, \beta)$. Hence,

Eq. (3.26) and Eq. (3.27) can be defined to solve numerically for the parameter υ_{sys} followed by the calculation of the parameter κ_{sys}.

$$R_{sys}^{TBCI_U} - R_{sys}^{TBCI_U} \phi_{sys} + \left\{ \phi_{sys} \frac{B_{R_{sys}^{TBCI_U}}\left[\upsilon R_{sys}^{pred}, \upsilon\left(1-R_{sys}^{pred}\right)\right]}{B\left[\upsilon R_{sys}^{pred}, \upsilon\left(1-R_{sys}^{pred}\right)\right]} \right\} = \left(1 - \frac{\gamma}{2}\right) \qquad (3.26)$$

$$\kappa = \nu R_{sys}^{pred} \qquad (3.27)$$

Next, the predicted system level reliability R_{sys}^{pred} is corrected with the system level knowledge factor ϕ_{sys} to obtain the corrected predicted system level reliability $R_{sys}^{\phi_{corr}}$ using Eq. (3.28).

$$R_{sys}^{\phi_{corr}} = \int R_{sys}^{pred} \left[\left(1-\phi_{sys}\right) \frac{\left(1-R_{sys}^{pred}\right)^{-0.5} R_{sys}^{pred-0.5}}{B\left(\alpha_{sys}, \beta_{sys}\right)} + \right.$$
$$\left. +\phi_{sys} \frac{\left(R_{sys}^{pred}\right)^{\alpha_{sys}-1} \left(1-R_{sys}^{pred}\right)^{\beta_{sys}-1}}{B\left(\alpha_{sys}, \beta_{sys}\right)} \right] dR_{sys}^{pred} \qquad (3.28)$$

Finally, Eq. (3.29) approximates the number of TAAF cycle failures τ_{TAAF} by assuming that the delta in the number of successes corresponds to the number of failures, i.e., $\Delta s = \Delta n \Delta R$.

$$\tau_{TAAF} = \left\lceil \left(N^{MP} - \upsilon_{sys}\right)\left(R_{sys}^{pred} - R_{sys}^{\phi_{corr}}\right)\right\rceil \qquad (3.29)$$

Figure 3-7 depicts the system level knowledge factor ϕ_{sys} versus the number of TAAF failures τ_{TAAF}.
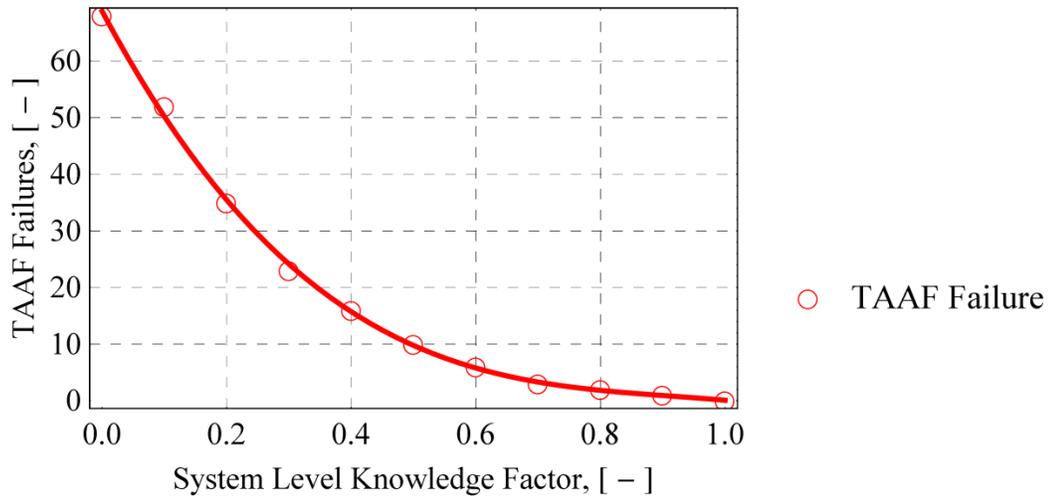
Figure 3-7:   System Knowledge Factor versus Number of System Failures

The number of predicted TAAF cycle failures is then allocated to the relevant system components [174], as listed in Table 3-5, using the SSME experience in accordance to the failure occurrence experience [125], i.e., 60% of the failures occur during the start-up, 20% within the first one-third of the full flight mission duration, and the remaining 20% up to flight mission completion. Other failure information will not be made available in this research to protect the proprietary nature of the data.

Table 3-5:   Failure Allocation to System Components

Component	Failure fraction
High pressure fuel turbopump	0.150
High pressure ox turbopump	0.076
Low pressure turbopumps	0.023
Nozzle extension	0.091
Combustion devices	0.170
Valves, sensors, and controls	0.184
Ducts	0.106
Other	0.184

Bayesian Estimation using the Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm is required to estimate the parameters of the Bayesian multilevel testing profiles aggregation because of nontrivial unscaled

posteriors. A blockwise Metropolis-Hastings algorithm with an independent candidate density is selected because of the computational efficiency, i.e., typically 10,000 iterations are needed to meet the convergence criteria even for high dimensional problems.

In particular, the blockwise Metropolis-Hastings algorithm loops through all unknown parameters θ_i conditional on all the other parameters θ_{-i} that are not in that block. At each iteration step, a new candidate value for the unknown parameters θ_i is proposed from an independent candidate density. The candidate value $q(\theta_i *)$ is either accepted or rejected according to the detailed balance condition that satisfies the kernel of the Metropolis-Hastings algorithm which is drawn on the logit-scale according to Eq. (3.30) [19].

$$\text{logit}\,\theta_i * \sim N\left(\text{logit}\,\theta_i^{(m)}, \sigma_i\right) \tag{3.30}$$

where $\text{logit}\,\theta_i *$ is defined as $\ln(\theta_i *) - \ln(1 - \theta_i *)$ [175], σ_i is the standard deviation that is used as tuning constant of the Markov chain with acceptance probability

$$\alpha\left(\theta_i^{(m)}, \theta_i * \mid \underline{\theta}_{-i}\right) = \min\left\{1, \frac{\pi(\theta_i * \mid Data)}{\pi\left(\theta_i^{(m)} \mid Data\right)} \frac{\theta_i * (1 - \theta_i *)}{\theta_i^{(m)}\left(1 - \theta_i^{(m)}\right)}\right\} \tag{3.31}$$

where $\pi(\theta_i * \mid Data)$ is the unscaled target density (posterior) that is evaluated with the new candidate value $\theta_i *$, $\pi\left(\theta_i^{(m)} \mid Data\right)$ is the unscaled target density (posterior) that is evaluated at the previously accepted value $\theta_i^{(m)}$.

The computational implementation of Eq. (3.30) is given in Eq. (3.32). In addition, Eq. (3.31) was modified, as given in Eq. (3.33), to solve numerical

instabilities that are caused by small or large numbers [176]. The overall computational implementation is depicted in Figure 3-8.

$$\theta_i^* = \frac{1}{e^{-F_X^{-1}(u)} + 1}$$

(3.32)

where $F_X^{-1}(u)$ is the equated inverse cumulative density function of $X \sim N\left(\text{logit}\,\theta_i^{(m)}, \sigma_i\right)$ at the random number u generated by $U \sim U(0,1)$.

$$\alpha\left(\theta_i^{(m)}, \theta_i^* \mid \underline{\theta}_{-i}\right) = \min\left\{1, e^{\left\{\ln\left[\pi(\theta_i^*|Data)\right] - \ln\left[\pi\left(\theta_i^{(m)}|Data\right)\right]\right\}} \frac{\theta_i^*\left(1-\theta_i^*\right)}{\theta_i^{(m)}\left(1-\theta_i^{(m)}\right)}\right\}$$

(3.33)

Markov chain Monte Carlo samples are not independent random samples; therefore, the following convergence criteria must be met: convergence to the stationary distribution, convergence of averages, and convergence to iid sampling [98]. The burn-in period, the acceptance rate, and the autocorrelation of the samples are a concern but can also be used to influence the Markov chain behavior in order to meet the convergence criteria. Unfortunately, no mathematical treatment is given that determines the length of the required burn-in period [97]. In this research, it turned out that 1000 iterations are sufficient for the burn-in period reflecting the two considerations: convergence to the stationary distribution using mainly trace plots as well as minimum scatter of the standard deviation σ_i of the independent candidate density that influences the acceptance rate and consequently the autocorrelation as measure for the convergence to iid sampling. Gregory [107], Liu [108], and Graves and Hamada [19] suggest acceptance rates close to 0.35 for problems that are similar to the ones treated in this research. However, an empirical study using the posteriors in the frame of this research suggests acceptance rates of 0.40 (see Figure 3-15). The

final step is to remove the iterations of the burn-in period and thin the remaining iterations of the Markov chain using the lag at which the autocorrelation is below the 0.95 confidence level.
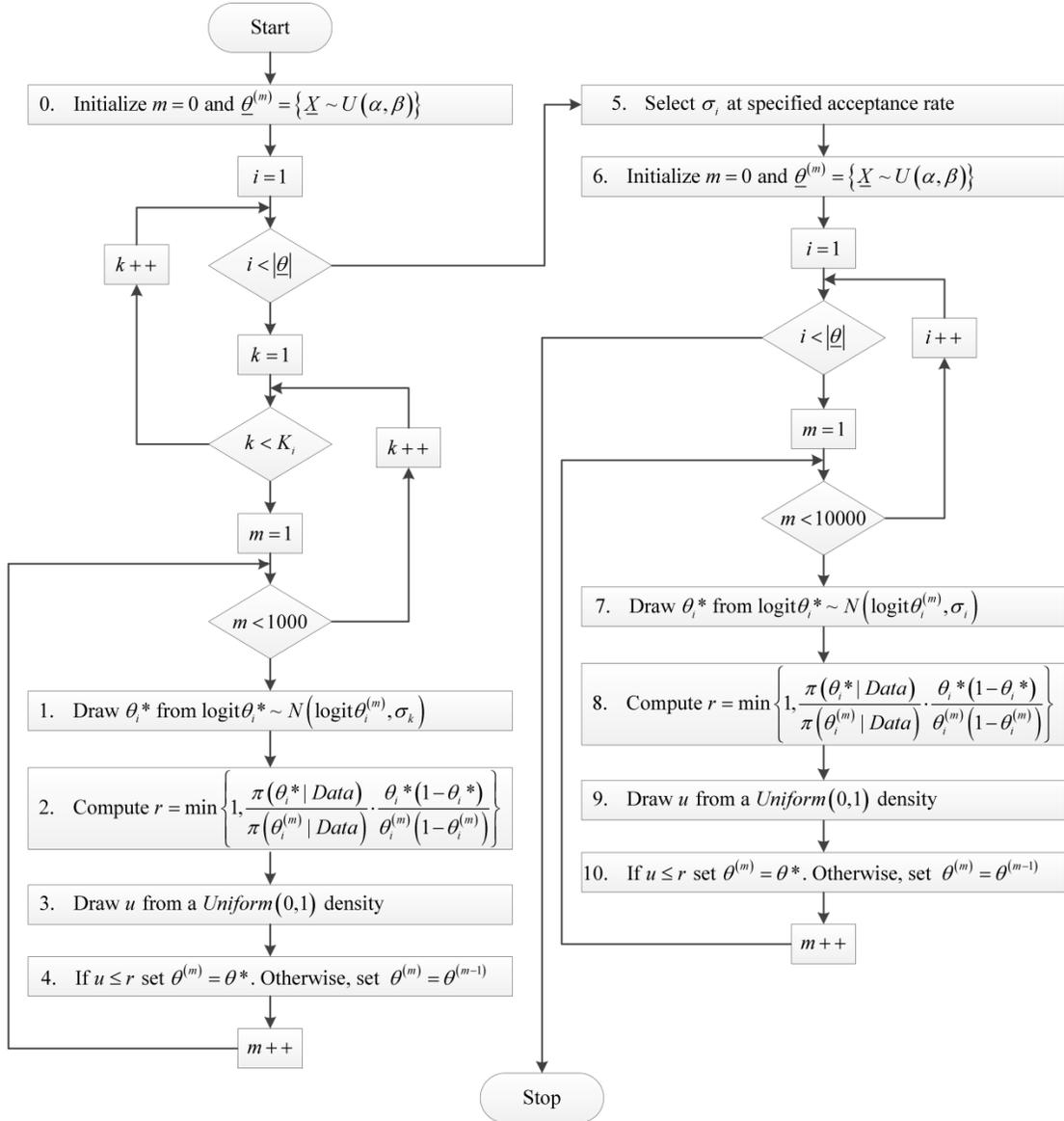
Figure 3-8:   Computational Algorithm of the Bayesian multilevel Testing Profiles Aggregation Method

Number of Development Hardware based on Bayesian Success Mission Profile

Testing

The number of hardware sets that are required to verify the inherent mission profile reliability-by-credibility (R-by-C) requirement is based on the Bayesian success testing under an exponential distribution assumption [115]. An expression in analytic form is found by the Bayesian estimation of a failure fraction.

The likelihood function for the failure fraction is a binomial distribution in which the number of trials n is replaced by the Equivalent Mission notion as given in Eq. (3.34).

$$L\left(Data \mid q\right) = \begin{pmatrix} EQM_{RbyC}^{TP} \\ r \end{pmatrix} q^{r} \left(1-q\right)^{EQM^{TP}-r} \tag{3.34}$$

where q is the failure fraction, EQM_{RbyC}^{MP} is the number of mission profile EQMs associated with the R-by-C requirement, and r is the number of observed failures during the hot-fire test plan. Note that the number of failures r is set to zero in the Bayesian success testing under an exponential distribution assumption.

The prior distribution for the failure fraction is a two-component mixture distribution in which the mixture components are a Uniform and a Beta distribution. Eq. (3.35) exhibits the two-component mixture distribution [82].

$$f\left(q;\alpha_q,\beta_q,\phi_{sys}\right) = \left(1-\phi_{sys}\right) + \phi_{sys} \frac{q^{\alpha_q-1}\left(1-q\right)^{\beta_q-1}}{\mathrm{B}\left(\alpha_q,\beta_q\right)} \tag{3.35}$$

where q is the failure fraction, ϕ_{sys} is the knowledge transfer factor, α_q and β_q are the Beta distribution shape parameters, and $\mathrm{B}(\cdot)$ is the Beta function.

The posterior of the failure fraction q is found using the Bayesian estimation. Eq. (3.36) exhibits the resulting posterior.

$$\pi\left(q;\alpha_q,\beta_q,\phi\mid Data\right)=\frac{\left(1-q\right)^{EQM_{RbyC}^{MP}}\left[\dfrac{\phi_{sys}q^{\alpha_q-1}\left(1-q\right)^{\beta_q-1}}{B\left(\alpha_q,\beta_q\right)}-\phi_{sys}+1\right]}{\dfrac{\phi_{sys}\Gamma\left(\alpha_q\right)\Gamma\left(\beta_q+EQM_{RbyC}^{MP}\right)}{B\left(\alpha_q,\beta_q\right)\Gamma\left(\cdot\right)}+\dfrac{1-\phi_{sys}}{EQM_{RbyC}^{MP}+1}} \tag{3.36}$$

with

$$\Gamma\left(\cdot\right)=\Gamma\left(\alpha_q+\beta_q+EQM_{RbyC}^{MP}\right)$$

where q is the failure fraction, EQM_{RbyC}^{MP} is the number of mission profile EQMs associated with the R-by-C requirement, ϕ_{sys} is the knowledge transfer factor, α_q and β_q are the Beta distribution shape parameters, $B(\cdot)$ is the Beta function, and $\Gamma\left(z\right)=\int_0^\infty t^{z-1}e^{-t}dt$ is the Gamma function.

The percentiles on the posterior distribution of the failure fraction q are given by Eq. (3.37).

$$\Pr\left(q\le q_u\right)=\int_0^{q_u}\pi\left(q;\alpha_q,\beta_q,\phi_{sys}\mid Data\right)dq=C \tag{3.37}$$

where q_U is the upper percentile of the posterior distribution of the failure fraction q and C is the level of credibility.

The expression in analytic form of the upper percentile failure fraction q_U is given in Eq. (3.38).

$$\frac{\left\{\left[1-\left(1-q_U\right)^{EQM_{RbyC}^{MP}+1}\right]\left(\phi_{sys}-1\right)B\left(\cdot\right)-a\phi_{sys}B_{q_U}\left(\cdot\right)\right\}\Gamma^{\langle1\rangle}\left(\cdot\right)}{\left(\phi_{sys}-1\right)B\left(\cdot\right)\Gamma^{\langle1\rangle}\left(\cdot\right)-a\phi_{sys}\Gamma\left(\alpha_q\right)\Gamma^{\langle2\rangle}\left(\cdot\right)}=C \tag{3.38}$$

with

$$a = \left(EQM_{RbyC}^{MP} + 1 \right),$$

$$\mathrm{B}(\cdot) = \mathrm{B}\left(\alpha_q, \beta_q \right),$$

$$\Gamma^{\langle 1 \rangle}(\cdot) = \Gamma\left(\alpha_q + \beta_q + EQM_{RbyC}^{MP} \right), \text{ and}$$

$$\Gamma^{\langle 2 \rangle}(\cdot) = \Gamma\left(\beta_q + EQM_{RbyC}^{MP} \right)$$

The final step is to transfer the upper percentile failure fraction q_U to the lower bound mission profile reliability R_{LB}^{MP}, i.e., $R_{LB}^{MP} = 1 - q_U$ and $q_U = 1 - R_{LB}^{MP}$. Then, Eq. (3.38) exhibits the expression in analytic form of the mission profile reliability-by-credibility (R-by-C) requirement that is required in the Bayesian success testing under an exponential distribution assumption. Note that a minimally informative prior distribution is assumed for the failure fraction q in this research, i.e., Jeffreys' prior with distribution parameters $\alpha_q = \beta_q = 0.5$ [101].

The number of hardware sets depends on the *capacity* of the piece parts and subassembly designs to withstand the thermofluid-mechanical *challenges* that are caused by *stress* and *time*, the two different *failure*-inducing *agents* [71]. Like the notion of Equivalent Mission (EQM), which accumulates the challenges, the notion of Equivalent Life (EQL) is used for the capacity in this research. Note that an EQL without an associated R-by-C requirement is, however, useless. In the automotive industry, the R-by-C requirement is also referred to as a *test bogey* [115]. Eq. (3.39) exhibits the definition for the reliable EQL EQL_{RbyC}^{MP} that is multiplied with a safety factor.

$$EQL_{RbyC}^{MP} = SF\left[\xi\frac{c_{RbyC}^{MP}}{NFC^{MP}} + (1-\xi)\frac{t_{RbyC}^{MP}}{CFD^{MP}}\right] \qquad (3.39)$$

where SF is the safety factor, ξ is the weighting factor of the capacity to withstand the challenges that trigger the two failure mechanisms, c_{RbyC}^{MP} is the number of reliable cycles, t_{RbyC}^{MP} is the reliable time, and NFC^{MP} is the number of hot firing cycles associated to the mission profile with the corresponding cumulative hot firing durations CFD^{MP}. Note that the values for the reliable cycles c_{RbyC}^{MP} and the reliable life t_{RbyC}^{MP} are based primarily on engineering judgment and simplified engineering models. Advanced physics-of-failure models for liquid rocket engine piece parts or subassemblies are still an area of active research [116].

Using the results of Eq. (3.38) and Eq. (3.39), the number of hardware sets HW_{RbyC}^{MP} that are needed to verify the inherent mission profile R-by-C requirement is finally given in Eq. (3.40). Note that Eq. (3.40) applies only to the system level test configuration.

$$HW_{RbyC}^{MP} = \frac{EQM_{RbyC}^{MP}}{EQL_{RbyC}^{MP}} \qquad (3.40)$$

Remaining Number of Development Hardware based on the Median Equivalent Life

Depending on the liquid rocket engine system's maturity, expressed as knowledge transfer factor ϕ_{sys}, the EQMs of the complete hot-fire test plan EQM^{TP} may exceed

the EQM that is associated with the R-by-C requirement, i.e., EQM_{RbyC}^{MP}. Therefore, Eq. (3.41) exhibits the remaining EQMs EQM_{rem}^{TP} that are in excess to the EQM_{RbyC}^{MP}.

$$EQM_{rem}^{TP} = EQM^{TP} - EQM_{RbyC}^{MP} \tag{3.41}$$

As a strategy for reducing the number of hardware sets, the remaining EQM testing profiles EQM_{rem}^{TP} are to be performed in excess to the reliable EQL EQL_{RbyC}^{MP} up to the median EQL $\widetilde{EQL}^{TP}$. The determination of the median EQL $\widetilde{EQL}^{TP}$ requires defining the underlying distributions that describe the two different failure-inducing agents (stress and time), which are the Poisson and the Weibull, respectively. The Poisson distribution is a proper choice for cyclic loads since it describes a random discrete variable with no upper bound. The Weibull distribution governs the time to occurrence of the weakest link of many competing failure processes. The median is chosen in preference over the average statistics in cases of small Weibull shape parameter [177]. Eq. (3.42) exhibits the fundamental definition. The subassemblies of liquid rocket engines that typically dominate the time to failure are the turbine(s), bearings, and combustion chamber liner.

$$\widetilde{EQL}^{TP} = SF\left[\xi \frac{\tilde{\lambda}^{TP}}{NFC^{MP}} + (1-\xi)\frac{\tilde{t}^{TP}}{CFD^{MP}} \right] \tag{3.42}$$

where SF is the safety factor, ξ is the weighting factor of the capacity to withstand the challenges that trigger the two failure mechanisms, $\tilde{\lambda}^{TP}$ is the median number of cycles to failure, $\tilde{t}^{TP}$ is the median life, and NFC^{MP} is the number of hot firing cycles associated to the mission profile with the corresponding cumulative hot firing durations CFD^{MP}.

The median number of cycles $\tilde{\lambda}$ is given in Eq. (3.43). Note that the median $\tilde{\lambda}^{TP}$ is actually calculated as mean $\bar{\lambda}^{TP}$ which does not impact the overall approach because the Poisson distribution is approximated with the Normal distribution if the mean $\bar{\lambda}^{TP}$ is above nine, and the median and the mean of a Normal distribution are equal.

$$P\left(NFC^{TP} \le c_{RbyC}^{MP}\right) = R\left(c_{RbyC}^{MP}\right) = 1 - \frac{\Gamma\left(1 + \left\lfloor c_{RbyC}^{MP} \right\rfloor, \bar{\lambda}^{TP}\right)}{\Gamma\left(1 + \left\lfloor c_{RbyC}^{MP} \right\rfloor\right)} \tag{3.43}$$

where $P\left(NFC^{TP} \le c_{RbyC}^{MP}\right)$ is the probability of failure associated with the test bogey, $R\left(c_{RbyC}^{MP}\right)$ is the reliable cycles, $\bar{\lambda}^{TP}$ is the mean of the Poisson distribution, $\lfloor \cdot \rfloor$ is the floor function, and $\Gamma(z) = \int_0^\infty t^{z-1} \cdot e^{-t} dt$ is the Gamma function.

The median time $\tilde{t}^{TP}$ is given in (3.44).

$$\tilde{t}^{TP} = t_{RbyC}^{MP} \left[\frac{\ln(2)}{-\ln\left(R_{RbyC}^{MP}\right)}\right]^{\frac{1}{\beta}} \tag{3.44}$$

where t_{RbyC}^{MP} is the reliable life, R_{RbyC}^{MP} is the reliability associated with the R-by-C requirement, and β is the shape parameter of the Weibull distribution.

Similarly to Eq. (3.40), the number of hardware sets HW_{rem}^{TP} that are in excess to the number of hardware sets HW_{RbyC}^{MP} but needed to complete the overall hot-fire test plan EQM_i^{TP} is defined in Eq. (3.45). Note that Eq. (3.45) is applied to all system integration levels, i.e., component, subsystem, and system level.

$$HW_{rem}^{TP} = \frac{EQM_{rem}^{TP}}{\widetilde{EQL}^{TP}} \tag{3.45}$$

Numerical Results

Testing Profiles Weighting according to the Lloyd-Lipow Model

Worlund et al. [125] provides data for the weighting factor w_{ij} that were estimated using the SSME and H-1, F-1, and J-2 liquid rocket engines. For ease of reference, the plot is reproduced in Figure 3-9.

To study further the general behavior of the testing profiles weighting factor w_{ij}, consider an arbitrary liquid rocket engine test plan that consists of five testing profiles with hot firing interval durations $\left(FD^{TP}_{sys,k} - FD^{TP}_{sys,k-1} \right)$, i.e., $[0,10)$, $[10,100)$, $[100,180)$, $[180,240)$, and $[240,300]$. For this example, assume that the total number of hot-fire tests remains constant in the study, i.e., 200, and that the numbers of failures depend on the system level knowledge factor ϕ_{sys} knowledge factor. In addition, the failure occurrence assumptions follow the empirically observed ones, i.e., 60% occur within the first couple of seconds, an additional 20% occur within one-third of the flight mission hot firing duration, and the remaining failures occur up to flight mission hot firing completion [125]. The results are depicted in Figure 3-10.

By looking at Figure 3-10, the general behavior of the testing profiles weighting factor w_{ij} is consistent with a rocket scientist's belief, i.e., if a certain hot-fire time is past the likelihood of a failure is lower and that the additional gain in demonstrating system reliability is minor. Prominent examples are the SSME and the F-1 liquid rocket engine. The critical time of the SSME is 1.5 seconds during start-up due to a thermodynamic instability [120] whereas the F-1 featured a 110 second

turbopump phenomenon problem that remains a mystery [178]. Current flight liquid rocket engines observe similar phenomena but cannot be disclosed in this research.
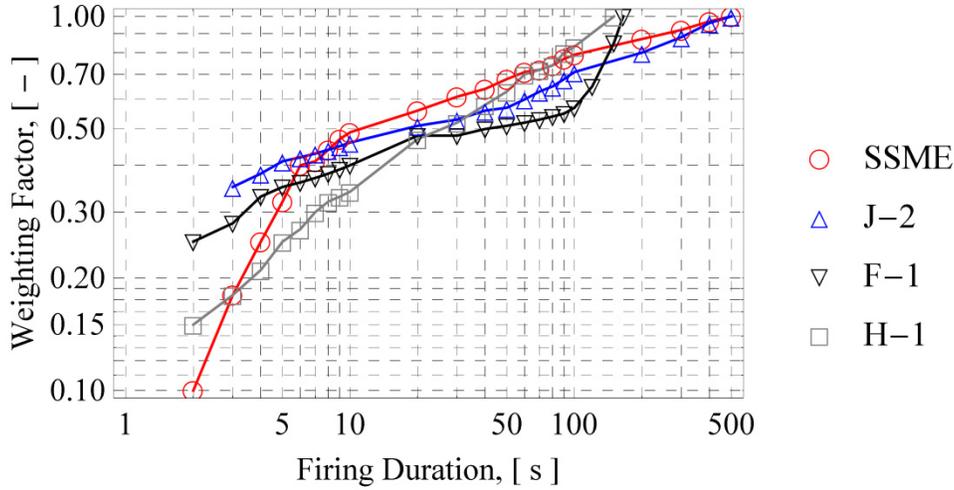
Figure 3-9: **Test Firings versus Weighting Factor**
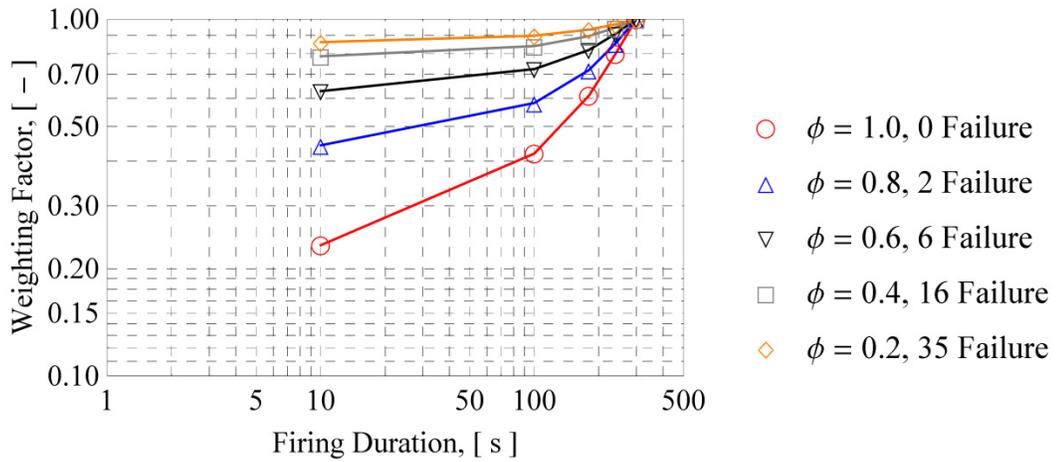
Figure 3-10: **Influence of Knowledge Factor Level on Testing Profiles Weighting Factor**

To conclude the discussion on the testing profiles weighting according to the Lloyd-Lipow model and to demonstrate the coherence between theory and actual data, both the empirical and the model results are overlaid in Figure 3-11.
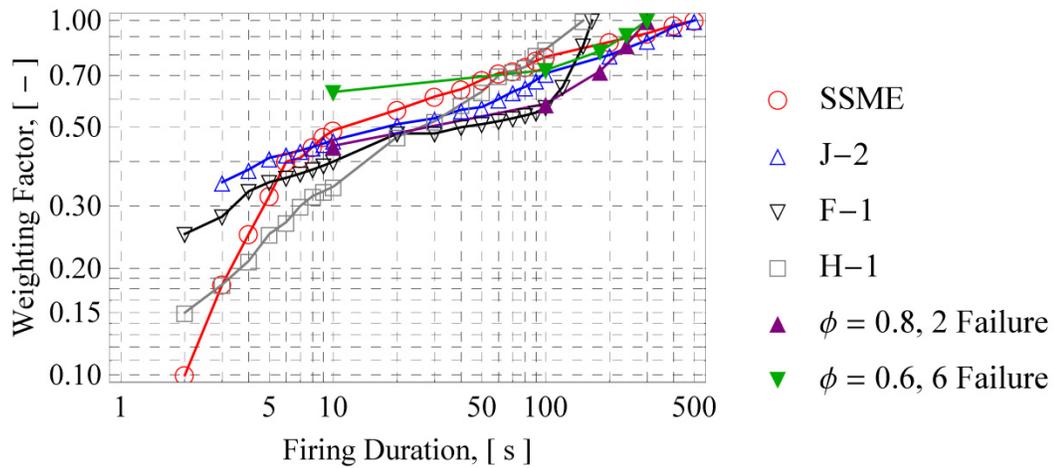
Figure 3-11: Historical Weighting Factors compared to Lloyd-Lipow Model based Testing Profiles Weighting Factor

Multilevel Bayesian Attribute Test Data Aggregation

Hamada [18] provides data for a three-component series system that is described in Table 3-6.

Table 3-6: Bayesian Aggregation: Three Component Series Test Data

Integration Level	Success	Failures	Units tested
Component 1	5	1	6
Component 2	6	0	6
Component 3	9	1	10
System	10	2	12

Each component is modeled as $p_i \sim Bi(n_i, \pi_i)$ where n_i is the number of units tested and π_i is the success probability for each of the $i = 1, 2, 3$ components. If common cause failures are excluded, i.e., the component failures are independent, the system reliability is $\pi_{sys} = \pi_1 \pi_2 \pi_3$. The prior distributions, Uniform density functions, are assumed to be independent for each π_i. The unscaled posterior distribution is, therefore, given as

$$\pi\left(\underline{\theta} \mid Data\right) = \pi\left(\pi_1, \pi_2, \pi_3 \mid \underline{x}\right) = \pi_1^8 \left(1 - \pi_1\right)^2 \pi_2^7 \left(1 - \pi_2\right)^2 \pi_3^3 \left(1 - \pi_3\right)$$

where the vector $\underline{x}$ corresponds to the data given in Table 3-6.

The Metropolis-Hastings algorithm starts with the tuning of the standard deviation σ_i for each of the i component probabilities. Figure 3-12 depicts the results of the sweep. A classical regression is used to select the proper tuning constant that meets an acceptance rate of 0.35. Noteworthy is the dependency of the acceptance rate and tuning constant on the number of units tested, i.e., the slight shift of the tuning constant sweep for the component 3. This effect is also applicable to the RISDM methodology.

The Metropolis-Hastings algorithm is applied again, with the tuning constants selected to result in acceptance rates of 0.35, in order to estimate the component reliabilities. Note that the actual acceptance rates of the Markov chains were 0.3557, 0.3503, and 0.3471, respectively.

Before accepting the results of the Metropolis-Hastings algorithm, the convergence criteria are checked by means of trace plots and the autocorrelation function [97, 98, 102] as depicted in Figure 3-13 for the current example. All three Markov chains provide adequate levels of convergence to the stationary distribution, convergence of averages, and convergence to iid sampling, respectively.

The remaining steps are the dropping of the burn-in iterations and the thinning of the Markov chains according to the lags, which do no longer feature a strong autocorrelation based on the results of the autocorrelation function. A lag of 10 is adequate for the given three-component series system example. Finally, the results can be used to estimate the system reliability as given in Table 3-7 and Figure 3-14.
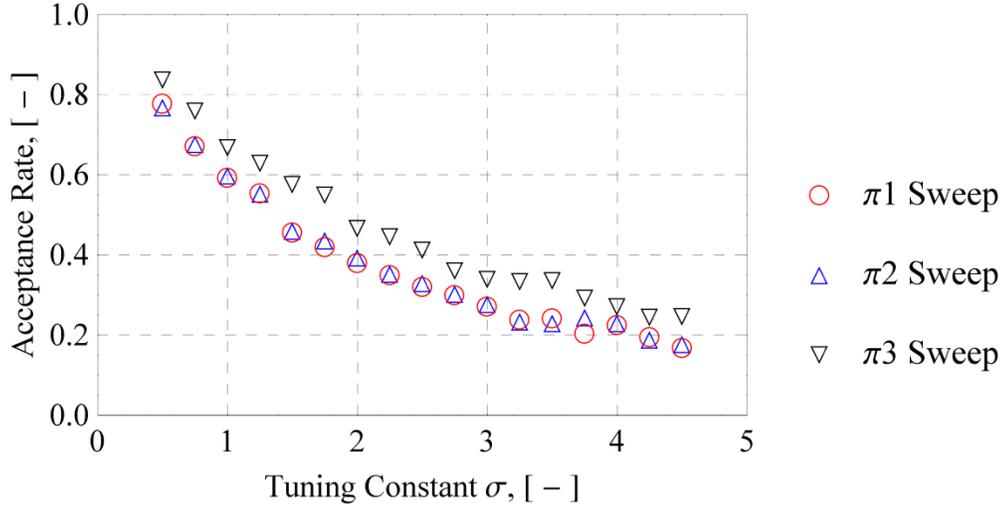
Figure 3-12: **Tuning Constant versus Acceptance Rate of Markov Chain**

Table 3-7: **Bayesian Aggregation: Three Components Series Results without Inclusion of System Level Data**

Parameter	Results by Hamada [18] Mean	StDev	Results by blockwise MH Mean	StDev
π_1	0.75	0.12	0.7511	0.1198
π_2	0.73	0.13	0.7316	0.1291
π_3	0.67	0.18	0.6715	0.1766
π_{sys}	0.36	0.13	0.3693	0.1339

The Metropolis-Hastings algorithm was also applied to the same three-component series system but evaluated with the system level data, and results similar to those reported in Hamada [18] were obtained.

As already mentioned above, the convergence criteria of a Markov chain are a concern, and the tuning constant σ is used to influence the behavior. To assess the level of influence, a parametric study was performed, and the results, shown in Figure 3-15, suggest setting the acceptance rate to a value near 0.40 in order to minimize the thinning of the Markov chain and, as a consequence, the number of iterations.
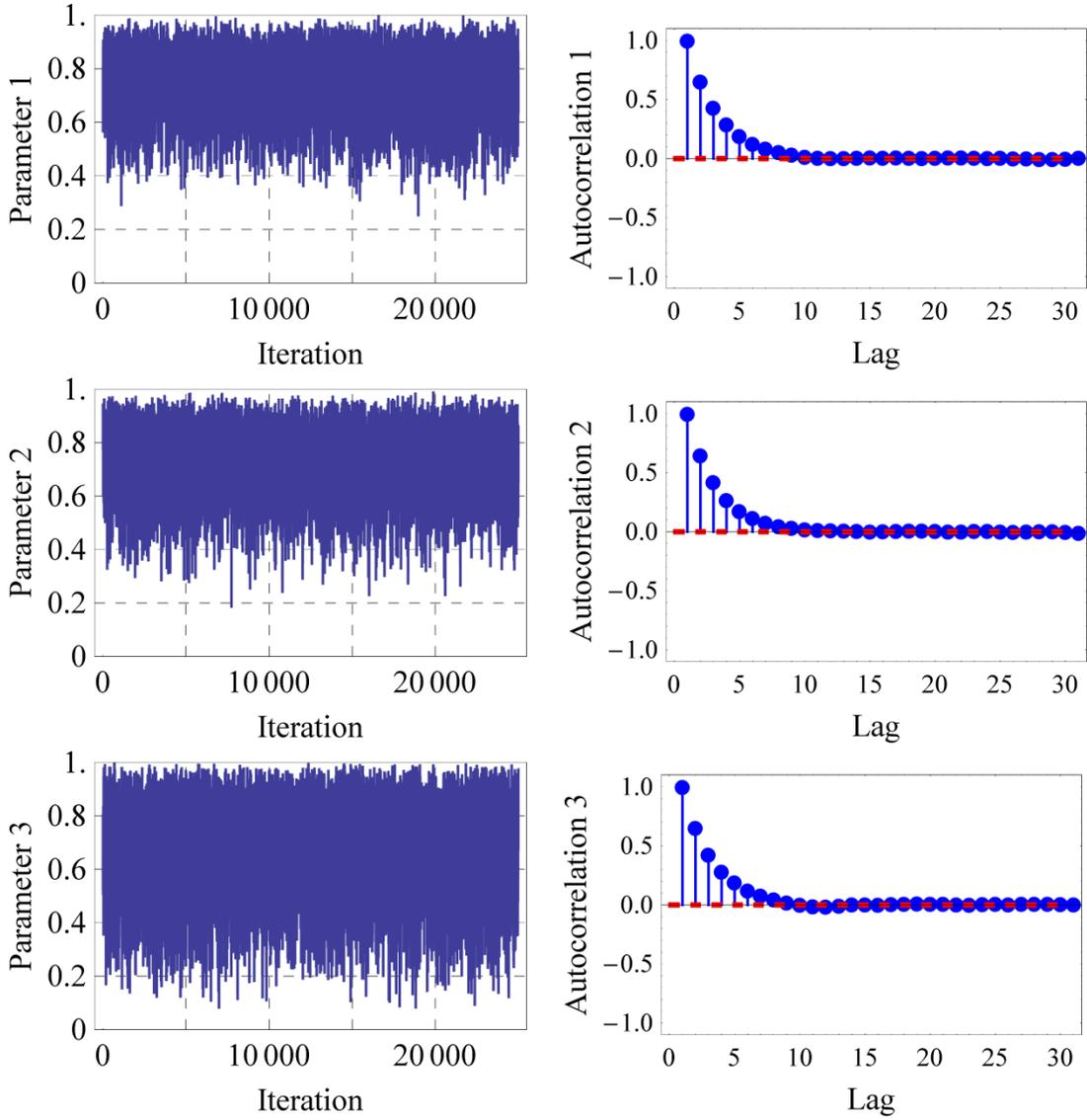
Figure 3-13: **Metropolis-Hastings Algorithm Convergence Criteria Check with 0.35 Acceptance Rate: Convergence to the Stationary Distribution, Convergence of Averages, and Convergence to iid Sampling**
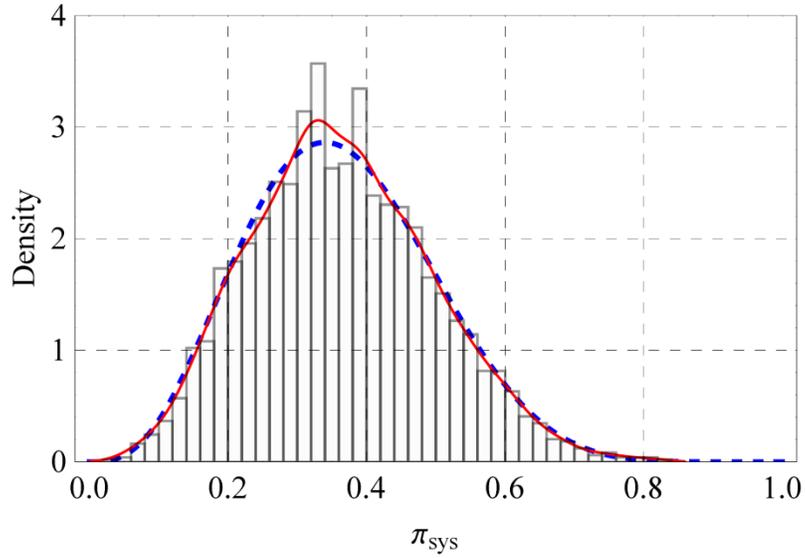
Figure 3-14: Thinned (Lag = 10) Metropolis-Hastings Algorithm Result overlaid on Analytic Solution of Three-component Series System
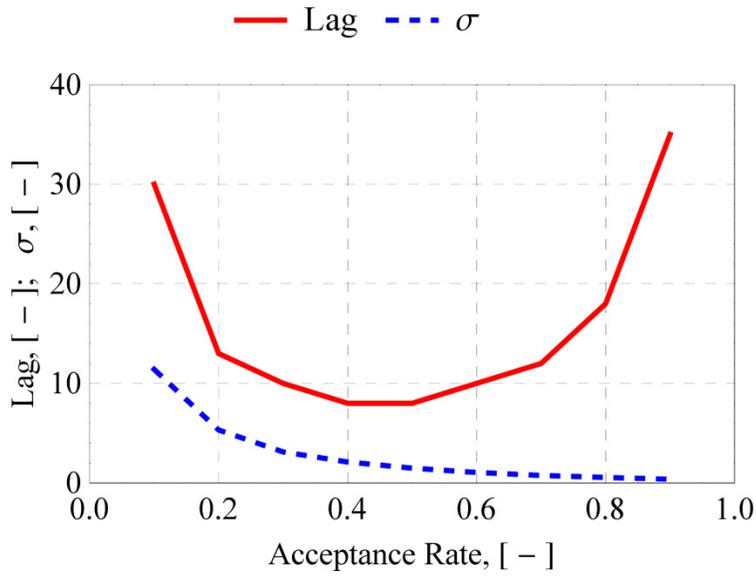
Figure 3-15: Influence of Acceptance Rate on ACF Lag and Tuning Constant

System Reliability Metamodel

The Bayesian estimation using the Metropolis-Hastings algorithm is the most time consuming model of all the models that are implemented in the RISDM methodology. A single estimation loop takes about 20 seconds using four cores of a Quad Core CPU 2.40GHz, e.g., 1000 iterations take about 5.5 hours. If a genetic algorithm satisficing run needed to be superimposed with 500 Monte Carlo simulations within each of the 1000 genetic algorithm iterations, the total simulation would then take approximately six days. In early project/program phases, high fidelity models are prerequisite to explore all possible design alternatives; therefore, metamodels, if accurate enough, should be used.

The Response Surface Methodology and regression-kriging technique are used in general to generate metamodels [179, 180]. However, the particularity of weighting the testing profiles is given in the Bayesian estimation of the system reliability, as depicted in Figure 3-16, which limits the applicability due to the lack of model accuracy. In particular, the two parameters ε and $p_{sys,j}$ depend on the total number of testing profiles $NFC_{sys,j}^{TP}$, the number of successful testing profiles $NFC_{sys,j}^{TP\langle S\rangle}$, and the number of failed testing profiles $NFC_{sys,j}^{TP\langle F\rangle}$ which influence likewise the EQMs that are used to estimate the system level reliability. Therefore, if the testing profiles weighting approximation is inaccurate, then the EQMs are erroneous and likewise the system level reliability.
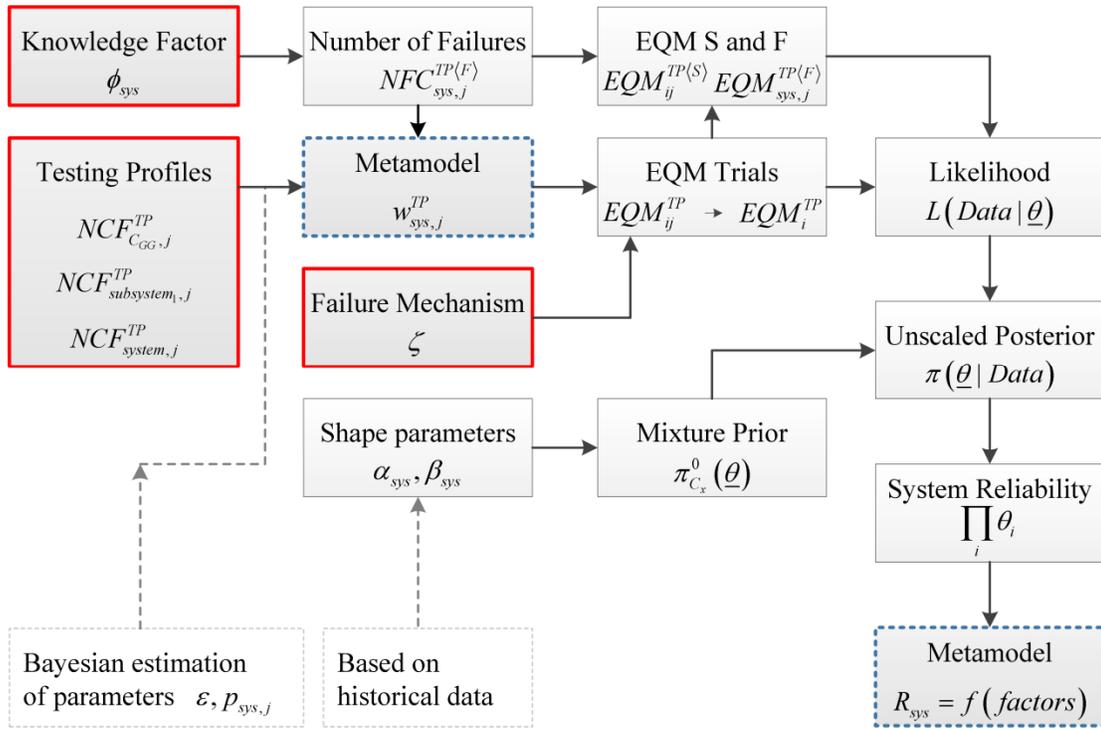
Figure 3-16: System Reliability Metamodel Factors

To study the metamodel accuracies, a D-optimal baseline design with 66 runs was selected to obtain minimum variance metamodel parameter estimates for the ten factors that are given in Table 3-8. Then, the baseline design was augmented by adding 300 design points using the strategy "minimum Euclidean distance" to obtain an overall design matrix that features a fraction of design space that is flat with a low standard error [179].

The analyzed design of experiment results are given in Table 3-9 in terms of R^2_{adj} and R^2_{pred} as measures of adequacy and predictive capability of the regression model, respectively [181]. Although the differences are small, they significantly adversely affect the system reliability approximations; therefore, the metamodel approach cannot be used in the frame of this research.

Table 3-8:   System Reliability Metamodel Design Details

Design Summary			
Study Type	Response Surface	Runs	366
Design Type	Distance	Coordinate Exchange	
Design Model	Quadratic		

Factor Name		Min	Max
Knowledge transfer factor	ϕ_{sys}	0.6124	0.7484
Component	$EQM^{TP}_{C_{GG},j}$	30	90
Subsystem	$EQM^{TP}_{subsystem_1,j}$	5	50
System Testing Profile 1	$EQM^{TP}_{sys,1}$	20	100
System Testing Profile 2	$EQM^{TP}_{sys,2}$	20	100
System Testing Profile 3	$EQM^{TP}_{sys,3}$	10	200
System Testing Profile 4	$EQM^{TP}_{sys,4}$	10	300
System Testing Profile 5	$EQM^{TP}_{sys,5}$	10	300
System Testing Profile 6	$EQM^{TP}_{sys,6}$	10	300
Failure mechanisms weighting	ζ	0.3	0.7

Table 3-9:   Design of Experiment Results

	$w_{sys,1}$	$w_{sys,2}$	$w_{sys,3}$	$w_{sys,4}$	$w_{sys,5}$	System reliability
R^2_{adj}	0.9866	0.9796	0.9813	0.9852	0.9843	0.9938
R^2_{pred}	0.9793	0.9681	0.9719	0.9752	0.9745	0.9888
Difference R^2	0.0073	0.0115	0.0094	0.01	0.0098	0.005
Difference in %	0.7	1.2	1.0	1.0	1.0	0.5

Knowledge Transfer Factor and Predicted Number of Test-Analyze-And-Fix Failures

The knowledge transfer factor ϕ is estimated using the SSME and the RS-68 liquid rocket engine data by assessing the prior information with respect to thrust and combustion chamber pressure levels. Table 3-10 and Table 3-11 list these examples and include the resulting knowledge factor levels and the number of TAAF cycle failures. The predicted numbers of TAAF cycle failures are generally coherent with the experienced number of failures during the hot-fire test plan execution.

Table 3-10: Knowledge Transfer Factor for the SSME

Case: F-1 to SSME	Thrust, kN	Pressure, bar	Propellants
F-1 (old)	6672	70	LOx/RP1
SSME (new)	2279	206.4	LOx/LH2
Factors of Eq. (3.25)	1 [1)]	0.418	1 [2)]
Knowledge transfer factor		0.421	
Predicted TAAF failures		14 [3)]	

[1)] Higher thrust level not taken into account as additional experience
[2)] Propellants are different; however, propellant experience from J-2
[3)] Number of TAAF failures are in accordance with the data given in [120]

Table 3-11: Knowledge Transfer Factor for the RS-68

Case: SSME to RS-68	Thrust, kN	Pressure, bar	Propellants
SSME (old)	2279	206.4	LOx/LH2
RS-68 (new)	3370	97	LOx/LH2
Factors of Eq. (3.25)	0.67	1 [1)]	1
Knowledge transfer factor		0.676	
Predicted TAAF failures		3 [2)]	

[1)] Higher pressure level not taken into account as additional experience
[2)] Number of TAAF failures may not seem in accordance with the data given in [118], i.e., 18 on engine level; however, if one analyzes the publication in detail, there are only 3 main failure modes addresses: shortfall of turbopump power, fatigue life of turbine blisks, and damping of turbine blisks. Thus, the number of failures seems to follow actual experience.

Bayesian Success Equivalent Mission

The equivalency of the expression given in Eq. (3.38) with the well-known frequentist binomial model $(1-C)=R^n$ may not be obvious, but if Eq. (3.38) is rewritten using a vague prior (parameters α_q, β_q, and ϕ_{sys} are set to one), the Bayesian-like binomial model can be stated as

$$Bi(0,q,n+1)=(1-q)^{n+1}=(1-C)\Leftrightarrow R^{n+1}=(1-C)$$

where $Bi(0,q,n+1)$ is the binomial probability density function including the Bayesian adjustment of the vague prior by the quantity $n+1$ instead of only n in the

frequentist framework and C as the confidence level. E.g., let be $R = 0.99$ and $C = 0.9$ then $n = 229.105$ in the frequentist estimation.

Similarly, let $R = 0.99$, $C = 0.9$, $\alpha_q = 1$, $\beta_q = 1$, and $\phi_{sys} = 1$ then $n = 228.105$ in the Bayesian estimation. The difference of one is due to the Bayesian adjustment, i.e., $n + 1$ instead of only n if a uniform prior distribution is assumed on the failure fraction q [115].

Next, the influence of the knowledge factor ϕ_{sys}, the lower bound mission profile reliability R_{LB}^{MP}, and the credibility level C on the Bayesian success EQM is studied. Figure 3-17 depicts the influence of the lower bound mission profile reliability R_{LB}^{MP} and Figure 3-18 the credibility level C on the Bayesian success EQM, respectively. The influence of the lower bound mission profile reliability R_{LB}^{MP} is slightly higher than the influence of the credibility level C on the Bayesian success EQM.
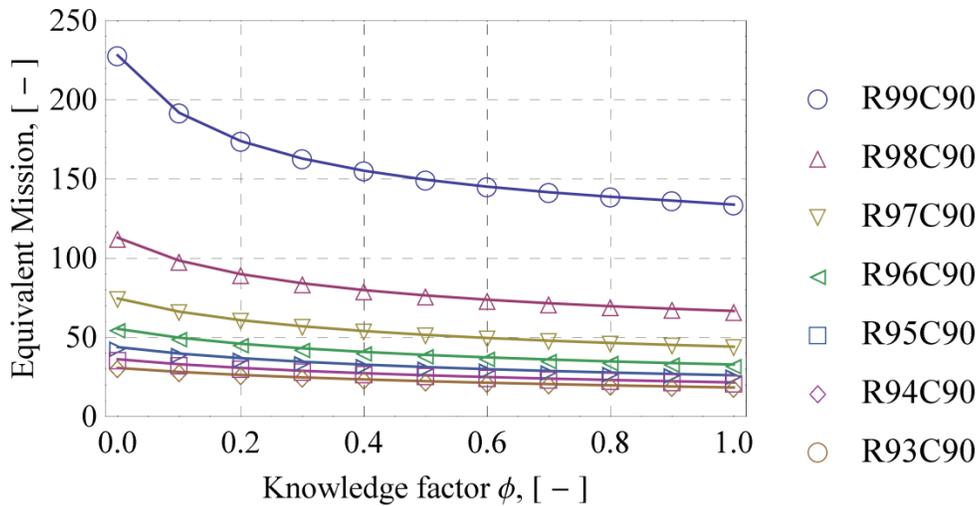
Figure 3-17: Knowledge Factor versus Bayesian Success Mission Profile for R9xC90
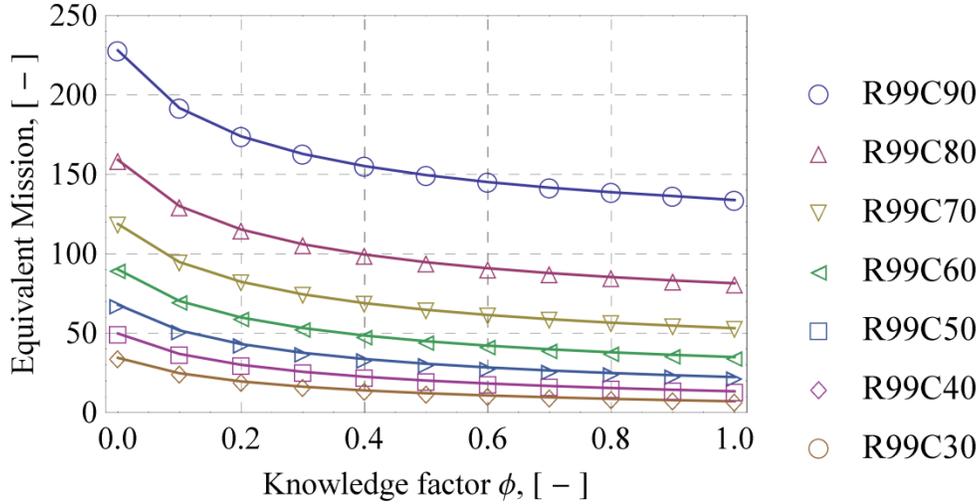
Figure 3-18: **Knowledge Factor versus Bayesian Success Mission Profile for R99Cx0**

Mission Profile and Median Equivalent Life

Richards [9] provides quantitative values for design starts and the design life for various liquid rocket engines. Eq. (3.39) is used to calculate the reliable EQL EQL_{RbyC}^{MP} assuming a weighting factor level ξ of 0.5. Table 3-12 lists the results.

Table 3-12: **Bogey EQLs for various Liquid Rocket Engines**

Engine Designation	Design starts	Design life, s	Mission profile cycles	Mission profile life, s	EQM_{RbyC}^{MP}
SSME	55	22700	4	821.5	20.7
F-1	20	2250	3	365	6.4
J-2	30	3750	3	680	7.8
RL10	20	4500	3	1000	5.6
LR87	12	1980	3	365	4.7
LR91	12	2700	3	425	5.2

RAIV Strategy Validation

The RAIV strategy is validated against the empirical mission reliabilities that are given in [117] using the published SSME and F-1 liquid rocket engine hot-fire test plans [120, 121, 178]. Table 3-13 lists and Figure 3-19 depicts the RAIV strategy

108

results. Note that Table 3-13 includes the error between the nonlinear fit and the RAIV strategy estimated median system reliabilities.

Table 3-13:   RAIV Strategy Validation

No. of hot-fire tests	SSME System Reliability	Error	No. of hot-fire tests	F-1 Liquid Rocket Engine System Reliability	Error
726	0.9833	0.0040	1081	0.9894	0.0020
2476	0.9936	0.0027	1437	0.9930	0.0004
2930	0.9948	0.0022	2740	0.9952	0.0015

The SSME system reliability figure of 0.9948 may also be compared to the estimated engine reliability for a nominal mission firing duration of 520 seconds of 0.9924 using the Crow/AMSAA reliability growth model with the hot-fire test data after the Challenger accident [150].
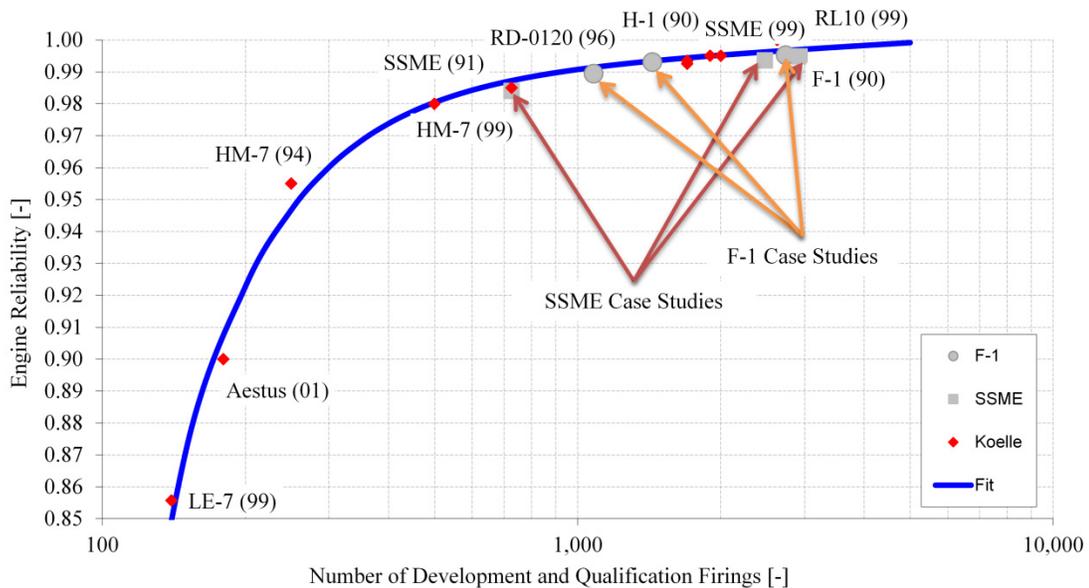
Figure 3-19: RAIV Strategy Validation

3.3.3   Modeling Initial Operational Capability

The IOC depends on the design maturity (TRL), the design process maturity (experience of the team), the R-by-C requirement that determines the hot-fire test

plan, the hot-fire test cadence, the number of test facilities, and the yearly funding level. Therefore, the Schedule Estimating Relation (SER) for the IOC, also known as the development duration D^{Dev}, can be expressed as the sum of the design and development duration D^{DD} and the test facility occupation duration D^{TP}. Eq. (3.46) exhibits the fundamental SER.

$$D^{Dev} = D^{DD} + D^{TP} \tag{3.46}$$

D^{DD}, the development duration in years, is based on the associated design and development cost estimation using the LRECM divided by MAF^{DD}, the mean annual funding level, and DDF^{DD}, a design and development factor that expresses the technology maturation effort. Eq. (3.47) exhibits the mathematical expression.

$$D^{DD} = \frac{C^{DD}}{DDF^{DD} MAF^{DD}} \tag{3.47}$$

The values for the design and development factor DDF^{DD}, as listed in Table 3-14, are derived from previous development programs. Note that the numerical values are linked to the LRECM input parameter design scope.

Table 3-14:   Numerical Values for the Design and Development Factor Levels

Design scope	Factor level
Simple modification	0.9
Extensive modification	0.95
New design	1
New product	1.25
New technology	1.5
Advanced state-of-the-art	2

D^{TP}, the test facility occupation duration in years, is driven by the hot-fire test plan, which is specified by the number of hot firing cycles associated with the testing profiles NFC_{ij}^{TP} as a result of the RAIV strategy, the number of available test

facilities TF that are suitable to provide the relevant testing conditions for the different system integration levels, and the test cadence TC of the utilized test facility to perform the hot firings (TC is in the range of four to six days [182]). D^{TP} is also influenced by the limited bogey EQL EQL_{RbyC}^{MP} of the piece parts and subassemblies of the system components. The RAIV strategy allocates more testing profiles (challenges) on a specific hardware than the hardware is capable to withstand (capacity) due to the failure-inducing agents; therefore, the testing profiles are spread over several hardware sets that requires hot-fire test free mounting and dismounting activities.

Let S_{ir} be the number of test campaigns for hot-fire test group i that is to be performed in a test facility of type r. Let NFC_{ijrs}^{TP} be number of hot firing cycles associated with the testing profiles j for hot-fire test group i that is to be performed in test campaign s in a test facility of type r. Let TC_{irs} be the test cadence (number of firing cycles per year per facility) that a test facility of type r can perform for test campaign s in hot-fire test group i. Let N_{irs}^{TF} be the number of test facilities of type r that can perform test campaign s in hot-fire test group i. Note that all of the test campaigns within a hot-fire test group at any type of test facility must be done sequentially, but other types of facilities can do other campaigns in parallel, and other hot-fire test groups can be done in parallel. Therefore, let I be the number of distinct hot-fire test groups and let R be the number of types of distinct test facilities. Then, Eq. (3.48) defines the maximum hot-fire D^{TP}.

$$D^{TP} = \max \left\{ \sum_{s=1}^{S_{ir}} \frac{\sum_{j=1}^{J_i} NFC_{ijrs}^{TP}}{TC_{irs} \cdot N_{irs}^{TF}} : i = 1,...,I; r = 1,...,R \right\} \tag{3.48}$$

3.4 Sensitivity Assessment

The RISDM methodology uses models, i.e. the multilevel EQM attribute data sampling for reliability as well as CERs and SERs for affordability and IOC, with imprecisely known parameters. Therefore, epistemic uncertainty is associated with the results of the specific affordability, reliability, and IOC models, respectively. In that context, the prior distribution conveys the epistemic uncertainty about possible model parameter values [183].

The sensitivity assessment is principally based on the objective based variable test/time philosophy that was applied to the RS-68 qualification/certification. The node representation that is used in the Bayesian multilevel testing profiles aggregation is depicted in Figure 3-20. Sections 3.4.1 to 3.4.4 discuss not only the dependencies of the parameters within the models but also the interdependencies between the areas of concern affordability, reliability, and IOC. Section 3.4.5 summarizes the results of a Monte Carlo simulation to assess the epistemic uncertainties of the RISDM methodology. Finally, Section 3.4.6 assesses the most pertinent epistemic uncertainty that is the component level node prior distributions that were discussed in Section 3.3.2. A justified selection for the range of the shape parameter β_{sys} is presented.
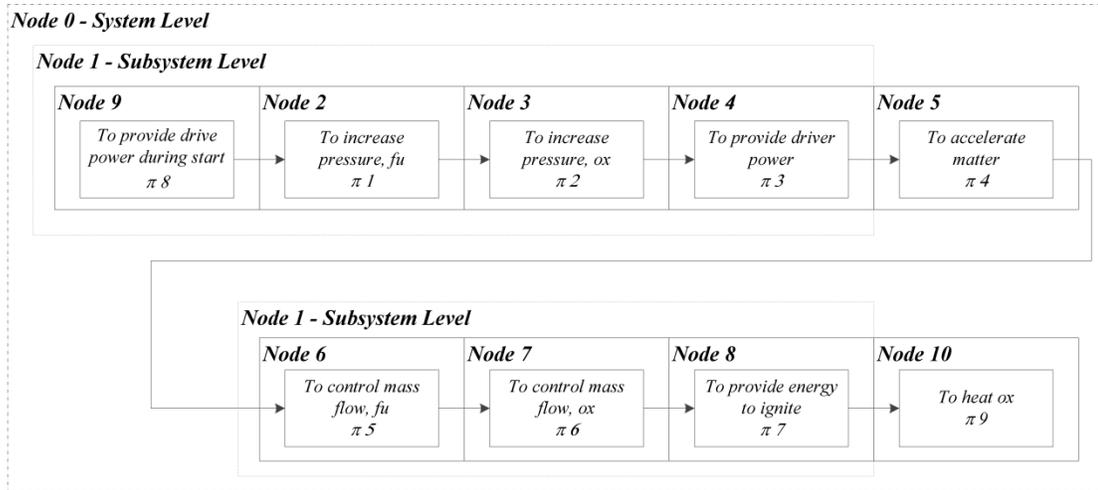
Figure 3-20:   Node Representation of Sensitivity Study Test Plan

3.4.1   Modeling Affordability

Figure 3-21 depicts the modeling strategy for the area of concern affordability, i.e., the epistemic parameters, the decision variables, and the models are indicated. Figure 3-22 is an extension to Figure 3-21 with focus on the hot-fire test cost model. Note that the production cost C^{Prod} is not further explained in Figure 3-21 because the cost drivers are well-known from manufacturing progress models (learning curve), i.e., learning and production rate assumptions as well as the level of producibility [142, 184]. However, the Monte Carlo simulation includes the production cost drivers for completeness.

3.4.2   Modeling Reliability

Figure 3-23 depicts the modeling strategy for the area of concern reliability, i.e., the epistemic parameters, the decision variables, and the models are indicated.

3.4.3 Modeling Initial Operational Capability

Figure 3-24 depicts the modeling strategy for the area of concern IOC, i.e., the epistemic parameters, the decision variables, and the models are indicated.

3.4.4 Composite Fitness Function

The composite fitness function involves two sources of epistemic uncertainties, i.e. the weighting of the areas of concern and the shapes of the utility functions that reflect the risk attitude of the decision maker. Figure 3-25 depicts the relations of these epistemic uncertainties and links them to the models for the area of concerns.

3.4.5 Monte Carlo Simulations

A simple Monte Carlo simulation was performed to assess not only the ranges for the epistemic uncertainties but also the ranges for the decision variables on the results of the models for the areas of concern affordability, reliability, and IOC. In addition, the sensitivity on the fitness function that is used in the satisficing using a genetic algorithm is given. The inputs terms of name, distributions, and ranges as well as the results are presented. The practical importance of input variables on the model outputs is also discussed.

Input variables

The Monte Carlo simulation input variable values are depicted in Figure 3-26 to Figure 3-30. They include the minimum and maximum values that are based on physical considerations and natural limits. Epistemic uncertainty parameters are prefixed with EpiUn/.
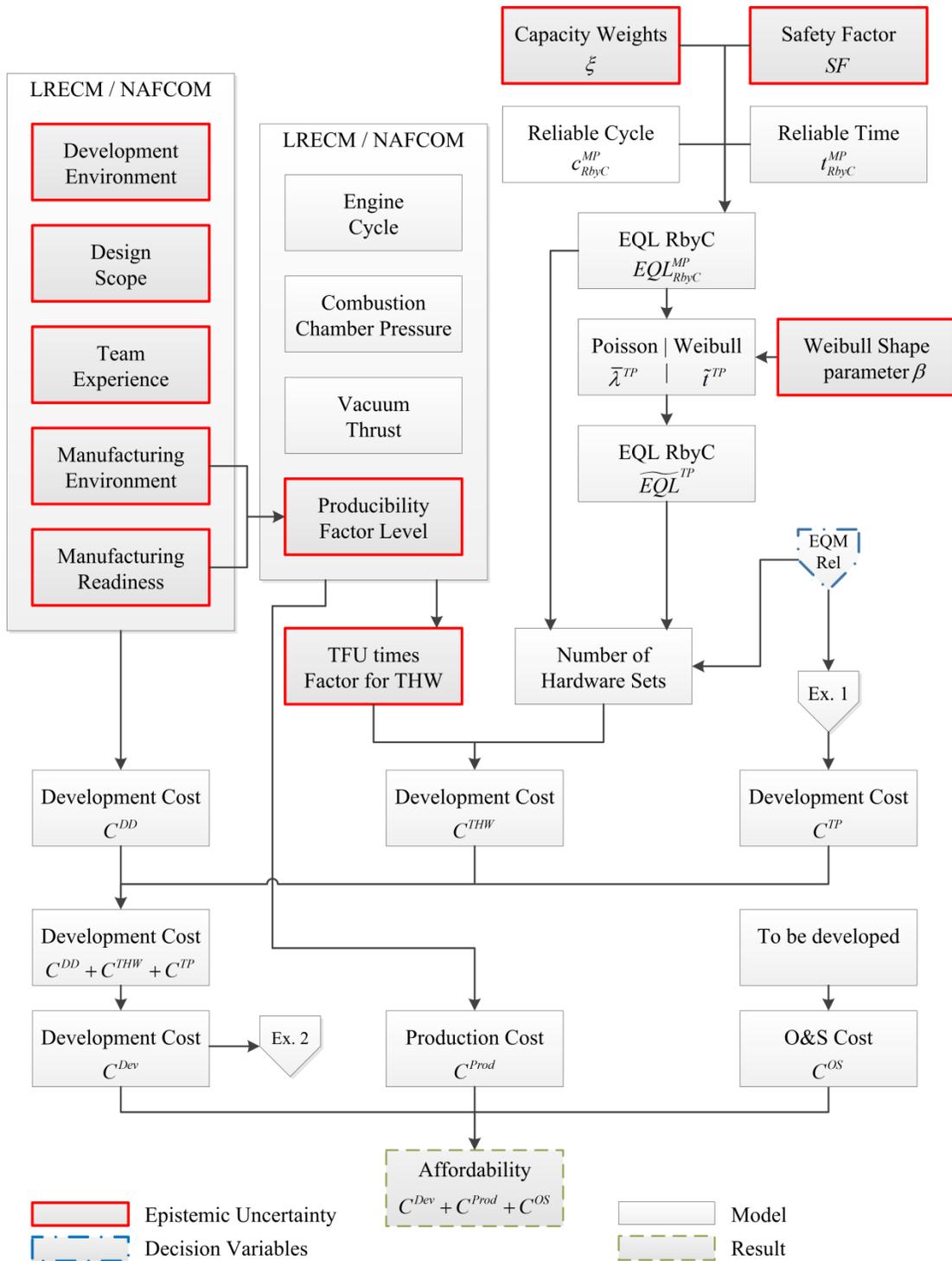
Figure 3-21: **Epistemic Uncertainty, Decision Variables, and Models of the Area of Concern Affordability – Main Process**
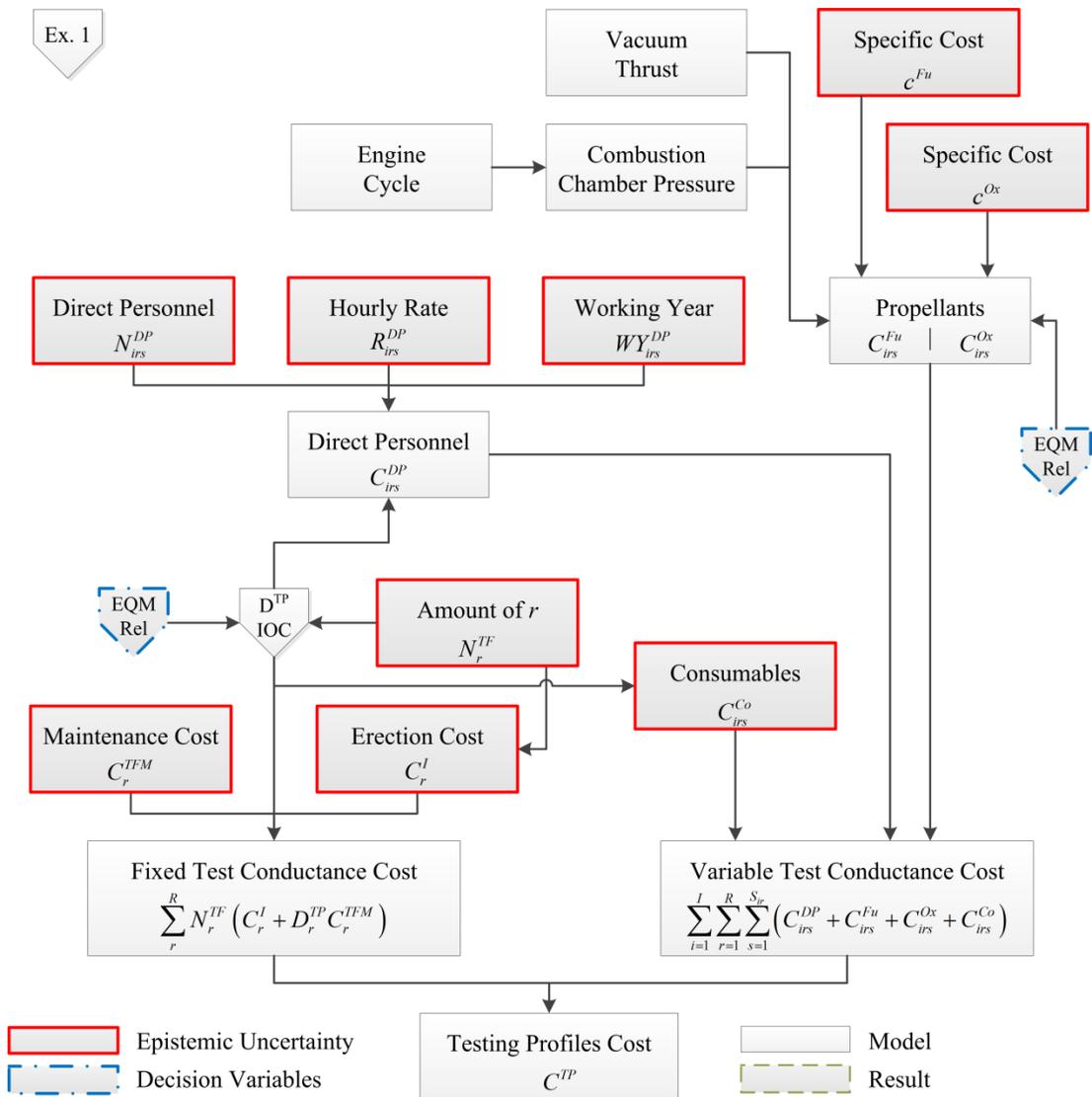
Figure 3-22: **Epistemic Uncertainty, Decision Variables, and Models of the Area of Concern Affordability – Testing Profiles Cost**
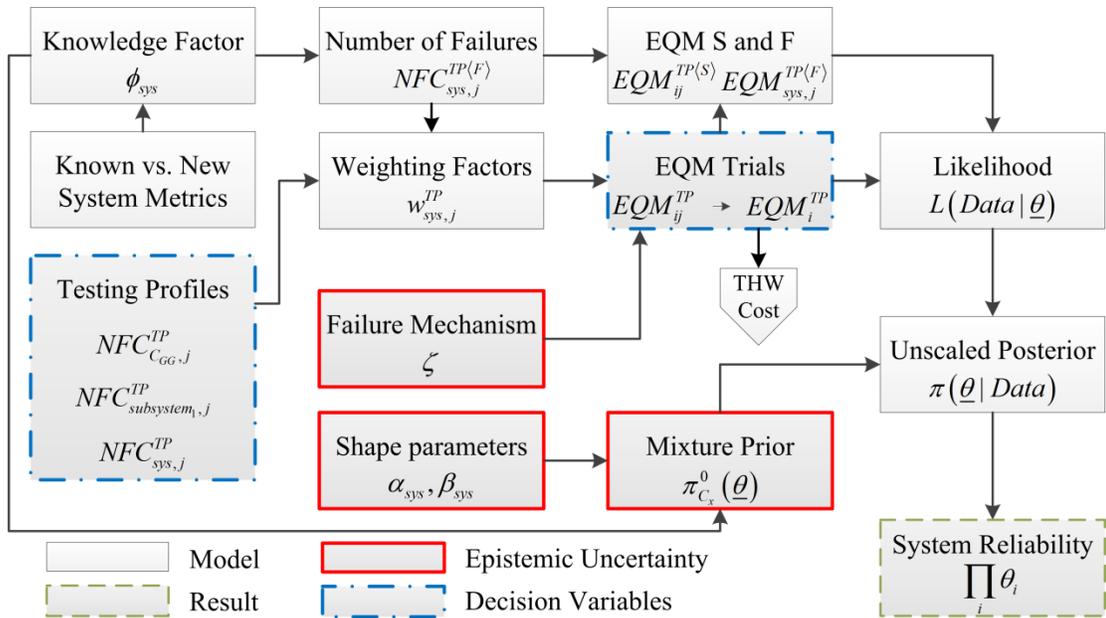
Figure 3-23: **Epistemic Uncertainty, Decision Variables, and Models of the Area of Concern Reliability**
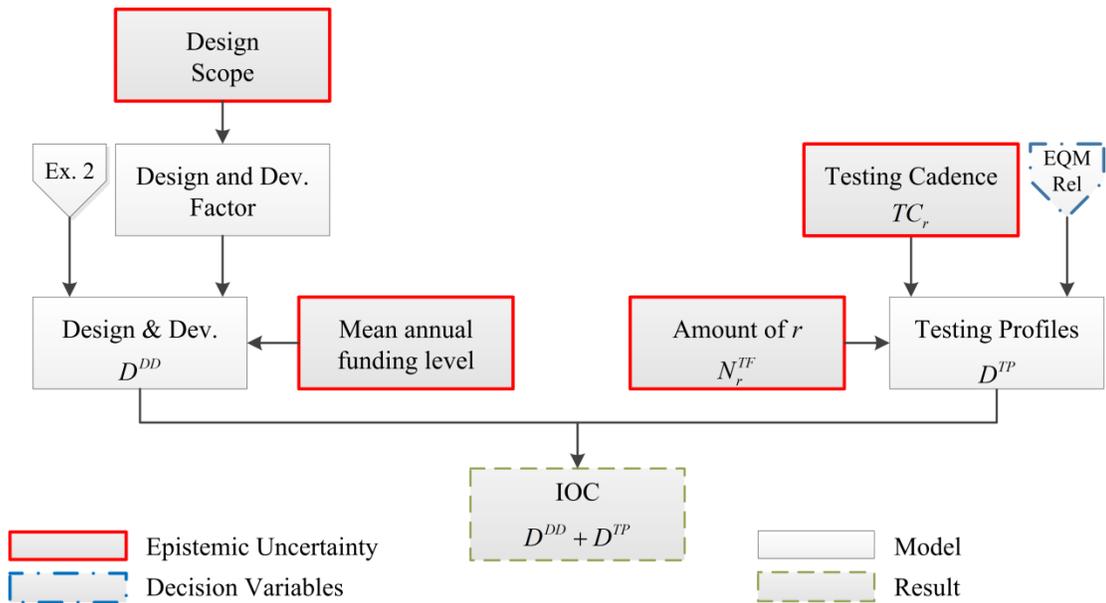
Figure 3-24: **Epistemic Uncertainty, Decision Variables, and Models of the Area of Concern IOC**
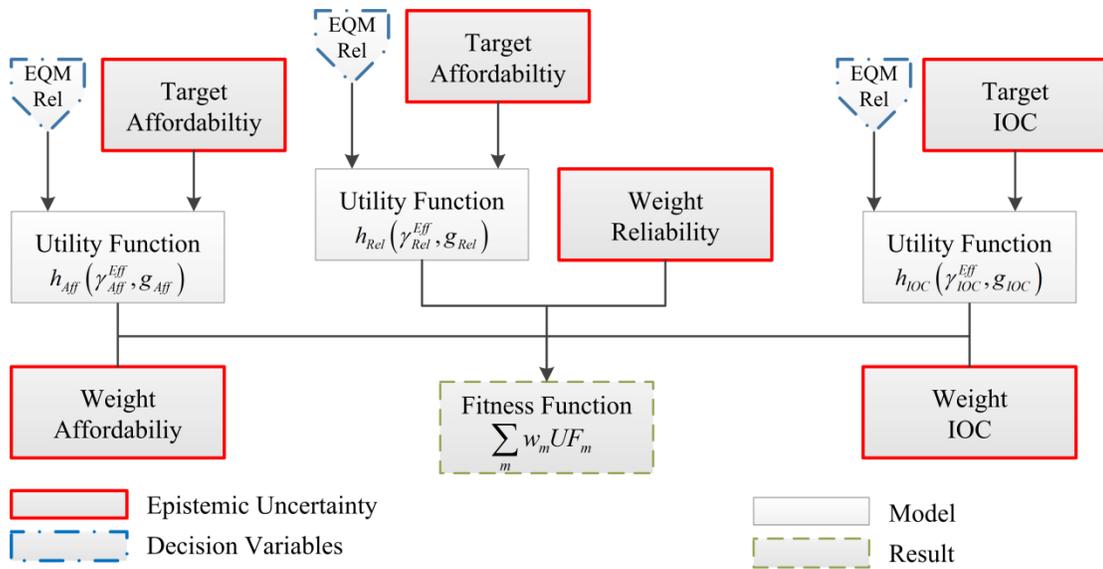
Figure 3-25: Epistemic Uncertainty of Composite Weighted Fitness Function and Relations to the Areas of Concern

Name	Cell	Graph	Min	Max
Category: Aff				
Aff/ Service Life [y]	J12		20	30
Aff/ Flights per Year [1/y]	J13		3	15
Category: CompTF				
CompTF/ No of Test Facility	Q207		1	2
CompTF/ Initial or Modification Facility Cost [MMU]	R207		0.01	19.98
CompTF/ No of People operating Facility	S207		10	25
CompTF/ Hourly Rate [MU/h]	T207		90	150
CompTF/ Yearly Working Hours [h]	U207		1500	2200
CompTF/ Test Cadence [1/w]	V207		20	50
CompTF/ Yearly Maintenance Cost [MMU]	W207		0.10	1.00
Category: DT				
DT/ Average Fund Availability [MMU]	F520		80	200
Category: EpiUn				
EpiUn/ SF (Safety Factor)	G450		1.3	10.0
EpiUn/ Knowledge Transfer Factor	K4		0.00	1.00
EpiUn/ Weibull Shape Parameter p1	L450		5.0	7.0
EpiUn/ Weibull Shape Parameter p2	L451		5.0	7.0
EpiUn/ Weibull Shape Parameter p3	L452		2.0	4.0
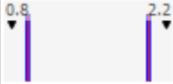
Figure 3-26:   Monte Carlo Simulation Inputs – Part I

Name	Cell	Graph	Min	Max
EpiUn/ Weibull Shape Parameter p4	L453		2.0	4.0
EpiUn/ Weibull Shape Parameter p5	L454		2.0	4.0
EpiUn/ Weibull Shape Parameter p6	L455		2.0	4.0
EpiUn/ Weibull Shape Parameter p7	L456		2.0	4.0
EpiUn/ Weibull Shape Parameter p8	L457		2.0	4.0
EpiUn/ Weibull Shape Parameter p9	L458		2.0	4.0
EpiUn/ Failure Mechanisms Weight	M207		0.3	0.7
EpiUn/ Capacity Weight	M450		0.3	0.7
EpiUn/ Prior Beta (alpha)	N205		38.0	67.0
EpiUn/ Prior Beta (beta)	O205		0.30	0.70
Category: FF				
FF/ Target Reliability	C105		0.900	0.990
FF/ Target IOC [y]	C110		8.00	12.00
FF/ Target Affordability [MMU]	C115		950 MU	1,300 MU
FF/ Reliability Weight	R20		0.00	1.00
FF/ Affordability Weight	R21		0.00	1.00
FF/ IOC Weight	R22		0.00	1.00
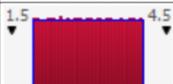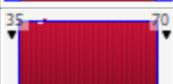
Figure 3-27:   Monte Carlo Simulation Inputs – Part II

120

Name	Cell	Graph	Min	Max
Category: LRECM				
LRECM/ Dev Environment	F500		1	2
LRECM/ Design Scope	F503		0	5
LRECM/ Mfg Environment	F501		0	3
LRECM/ Mfg MRL	F502		0	2
LRECM/ Team Experience	F504		0	3
LRECM/ Producibility	F506		0.30	0.80
LRECM/ Unit LC Factor	F512		0.85	0.95
LRECM/ Overhead on TFU for Dev System Price	F514		1.25	1.50
LRECM/ Oxidizer Cost per kilogramm [MU/kg]	F551		0.11	0.20
LRECM/ Fuel Cost per kilogramm [MU/kg]	F552		8.6	20.0
Category: NFC TP				
NFC TP/Component	J207		30	90
NFC TP/ Subsystem	J225		5	100
NFC TP/ System1	J235		50	200
NFC TP/ System2	J236		50	200
NFC TP/ System3	J237		50	200
NFC TP/ System4	J238		50	200

Figure 3-28: Monte Carlo Simulation Inputs – Part III

Name	Cell	Graph	Min	Max
NFC TP/ System5	J239		50	200
NFC TP/ System6	J240		50	200
Category: SubsystemTF				
SubsystemTF/ No of Test Facility	Q225		1	2
SubsystemTF/ Initial or Modification Facility Cost [MMU]	R225		0	60
SubsystemTF/ No of People operating Facility	S225		10	25
SubsystemTF/ Hourly Rate [MU/h]	T225		90	150
SubsystemTF/ Yearly Working Hours [h]	U225		1500	2200
SubsystemTF/ Test Cadence [1/w]	V225		20	50
SubsystemTF/ Yearly Maintenance Cost [MMU]	W225		0.10	1.00
Category: SystemTF				
SystemTF/ No of Test Facility	Q235		1	4
SystemTF/ Initial or Modification Facility Cost	R235		0	200
SystemTF/ No of People operating Facility	S235		30	50
SystemTF/ Hourly Rate [MU/h]	T235		90	150
SystemTF/ Yearly Working Hours	U235		1500	2200
SystemTF/ Test Cadence [1/w]	V235		20	50
SystemTF/ Yearly Maintenance Cost [MMU]	W235		0.50	2.00

Figure 3-29:   Monte Carlo Simulation Inputs – Part IV

122

Name	Cell	Graph	Min	Max
Category: TF				
TF/ Consumables [MMU per test]	F554		0.02	0.05

Figure 3-30: Monte Carlo Simulation Inputs – Part V

Output analysis

The Monte Carlo simulation results are analyzed with regard to their importance using the Spearman's rank correlation coefficient and the associated hypothesis test (or confidence interval) at a significance level of 0.001. The significance level differs significantly from the common 0.05 level because only convincing correlations are of practical importance in the frame of this research (a level of 0.05 is considered suggestive but inconclusive [185]). Figure 3-31 to Figure 3-33 depict the Monte Carlo simulation output results followed by Table 3-15 to Table 3-34 that list the Spearman's rank correlation coefficients and the associated p-value of the corresponding hypothesis test.

Based on the p-values, the only important epistemic uncertainty is linked to the shape parameters α and β of the prior distribution for the node reliabilities. Therefore, the influence is further studied in Section 3.4.6. All other epistemic uncertainties are of minor importance. The Monte Carlo simulation revealed also the strong influence of the decision variables, the hot-fire tests, and the non-technical TRL, MRL, product life cycle, design scope and environment, and team experience. Therefore, the RISDM methodology results are mainly determined by aleatory model parameters that reflect non-technical and technical stakeholder inputs as well as the decision variables of the problem formulation.

Name	Cell	Graph	Min	Mean	Max
No. of Failures Sys TP 1	B250		0	11	43
No. of Failures Sys TP 2	B251		0	4	14
No. of Failures Sys TP 3	B252		0	4	14
Engine Level THW	D30		9	30	109
BogeyEQL based	D31		1	1	13
p1 SubSys THW	D34		1	2	9
p2 SubSys THW	D35		1	2	9
p3 SubSys THW	D36		1	2	9
p4 SubSys THW	D37		0	0	0
p5 SubSys THW	D38		1	2	9
p6 SubSys THW	D39		1	2	9
p7 SubSys THW	D40		1	2	9
p8 SubSys THW	D41		1	2	9
p3 Comp THW	D55		1	1	5
0.5-percentile System Level	F24		0.9703	0.9852	0.9931
p-th Percentile Rel Sys	F26		0.9671	0.9833	0.9921
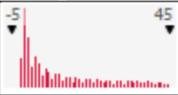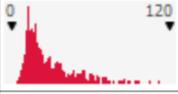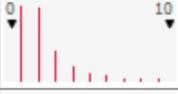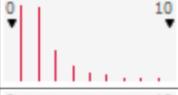
Figure 3-31:   Monte Carlo Simulation Outputs – Part I

124

Name	Cell	Graph	Min	Mean	Max
Risk Coeff Rel	H105		-79.6	-2.9	10.0
Risk Coeff Dev Dur	H110		9.6	23.8	63.3
Risk Coeff Dev Cost	H115		4.9	10.4	17.5
Service Life HW Tot	J16		80	250.4	480
Development Cost (D&D), [MMU]	J20		54	431	1228
Development Cost (HW) [MMU]	J21		376	1863	9235
Development Cost (DD+HW) [MMU]	J22		460	2294	9528
Development Cost (Testing) [MMU]	J23		219	512	1061
DDE&T Cost (Total Dev) [MMU]	J24		825	2806	10188
TFU [MMU]	J30		24.4	45.4	71.2
AUC [MMU]	J31		17.3	34.7	60.3
Production Cost Tot [MMU]	J35		1590	8698	23443
Affordability [MMU]	J36		3460	11504	28596
MoE Reliabiltiy	P5		0.9671	1.0	0.9921
MoE Affordability [MMU]	P6		825	2806	10188
MoE IOC [y]	P7		3.3	15.0	48.1
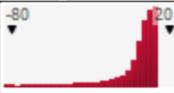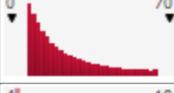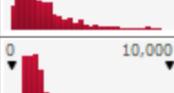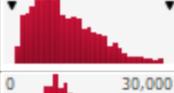
Figure 3-32: Monte Carlo Simulation Outputs – Part II

Name	Cell	Graph	Min	Mean	Max
weighted FF	P30		-1.5E+57	-1.5E+54	1.00
Comp / Test Cost [MMU]	Y207		4	22	58
Subsystem Test Cost [MMU]	Y225		3	48	99
System Test Cost [MMU]	Y235		169	443	970
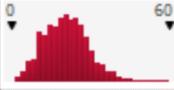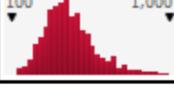
Figure 3-33:   Monte Carlo Simulation Outputs – Part III

Table 3-15: **Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix I-a**

Output	EpiUn/ Knowledge Transfer Factor rho	p-val	Aff/ Service Life [y] rho	p-val	Aff/ Flights per Year rho	p-val	FF/ Reliability Weight rho	p-val	FF/ Affordability Weight rho	p-val	FF/ IOC Weight rho	p-val	FF/ Target Reliability rho	p-val
MoE Reliability	0.06	0.06	-0.02	0.55	0.00	0.95	0.00	0.92	0.00	0.91	-0.03	0.37	0.05	0.10
MoE Affordability [MMU]	-0.11	0.00	-0.02	0.59	-0.03	0.35	-0.05	0.13	0.00	0.97	-0.01	0.73	0.01	0.76
MoE IOC [y]	0.07	0.04	-0.01	0.65	0.01	0.78	0.05	0.11	-0.03	0.38	0.01	0.69	0.02	0.44
Service Life HW Tot	0.02	0.54	0.32	0.00	0.95	0.00	0.02	0.58	-0.01	0.70	0.02	0.59	-0.01	0.79
Development Cost (D&D), [MMU]	0.02	0.50	0.00	0.92	-0.02	0.58	-0.02	0.50	-0.05	0.15	0.03	0.37	0.00	0.92
Development Cost (HW) [MMU]	-0.14	0.00	-0.01	0.74	-0.04	0.24	-0.05	0.15	0.00	0.96	-0.03	0.43	0.01	0.81
Development Cost (DD+HW) [MMU]	-0.12	0.00	0.00	0.89	-0.03	0.29	-0.05	0.09	0.00	0.89	-0.01	0.77	0.01	0.73
Development Cost (Testing) [MMU]	0.05	0.14	-0.07	0.04	0.01	0.74	0.04	0.16	0.02	0.46	-0.01	0.64	0.00	0.95
0.5-percentile (Median) System Level	0.06	0.07	-0.02	0.58	0.00	1.00	-0.01	0.81	0.00	0.88	-0.03	0.34	0.05	0.08
DDE&T Cost (Total Dev) [MMU]	-0.11	0.00	-0.02	0.59	-0.03	0.35	-0.05	0.13	0.00	0.97	-0.01	0.73	0.01	0.76
p-th Percentile Rel Sys	0.06	0.06	-0.02	0.55	0.00	0.95	0.00	0.92	0.00	0.91	-0.03	0.37	0.05	0.10
Engine Level THW	-0.13	0.00	0.00	0.94	-0.05	0.09	-0.05	0.15	0.02	0.53	-0.03	0.40	-0.02	0.63
TFU [MMU]	-0.08	0.01	-0.01	0.64	0.01	0.73	-0.02	0.58	-0.03	0.36	0.00	0.95	0.03	0.28
weighted FF	-0.01	0.71	0.02	0.52	0.00	1.00	0.02	0.55	-0.07	0.03	0.01	0.81	0.01	0.65
BogeyEQL based Reliability	-0.33	0.00	0.00	0.94	0.01	0.80	0.01	0.77	0.00	0.95	0.00	0.92	-0.04	0.27
AUC [MMU]	-0.08	0.01	-0.01	0.73	0.00	0.88	-0.03	0.42	-0.02	0.49	0.01	0.86	0.02	0.46
p1 SubSys THW	-0.10	0.00	-0.04	0.19	0.03	0.41	-0.05	0.12	0.03	0.41	-0.05	0.14	0.01	0.83
p2 SubSys THW	-0.10	0.00	-0.04	0.19	0.03	0.41	-0.05	0.12	0.03	0.41	-0.05	0.14	0.01	0.83
Production Cost Tot [MMU]	-0.02	0.46	0.26	0.00	0.78	0.00	0.00	0.94	-0.02	0.50	0.02	0.45	0.01	0.64
p3 SubSys THW	-0.10	0.00	-0.04	0.19	0.03	0.41	-0.05	0.12	0.03	0.41	-0.05	0.14	0.01	0.83
Affordability [MMU]	-0.05	0.15	0.23	0.00	0.70	0.00	-0.01	0.69	-0.03	0.42	0.02	0.50	0.02	0.52
p5 SubSys THW	-0.10	0.00	-0.04	0.19	0.03	0.41	-0.05	0.12	0.03	0.41	-0.05	0.14	0.01	0.83

Table 3-16:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix I-b

Output	EpiUn/ Knowledge Transfer Factor		Aff/ Service Life [y]		Aff/ Flights per Year		FF/ Reliability Weight		FF/ Affordability Weight		FF/ IOC Weight		FF/ Target Reliability	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
p6 SubSys THW	-0.10	0.00	-0.04	0.19	0.03	0.41	-0.05	0.12	0.03	0.41	-0.05	0.14	0.01	0.83
p7 SubSys THW	-0.10	0.00	-0.04	0.19	0.03	0.41	-0.05	0.12	0.03	0.41	-0.05	0.14	0.01	0.83
p8 SubSys THW	-0.10	0.00	-0.04	0.19	0.03	0.41	-0.05	0.12	0.03	0.41	-0.05	0.14	0.01	0.83
p3 Comp THW	-0.05	0.12	-0.01	0.84	-0.07	0.03	-0.05	0.11	0.00	0.89	-0.01	0.78	0.02	0.59
Risk Coeff Rel	-0.01	0.76	0.01	0.87	0.00	0.96	0.02	0.55	0.02	0.50	0.02	0.55	-1.00	0.00
Risk Coeff Dev Dur	0.02	0.46	-0.01	0.81	0.01	0.73	0.01	0.69	0.00	0.92	0.00	0.99	-0.02	0.63
Risk Coeff Dev Cost	-0.04	0.18	-0.03	0.35	-0.03	0.32	0.05	0.11	0.02	0.51	0.03	0.37	-0.03	0.30
Comp / Test Cost [MMU]	-0.02	0.50	0.00	0.94	0.05	0.15	0.00	0.93	0.02	0.53	0.05	0.15	-0.02	0.46
Subsystem Test Cost [MMU]	0.01	0.74	-0.07	0.02	-0.02	0.47	-0.03	0.35	0.03	0.32	0.01	0.64	0.03	0.30
System Test Cost [MMU]	0.04	0.17	-0.06	0.06	0.01	0.80	0.05	0.14	0.02	0.63	-0.02	0.58	-0.01	0.78
No. of Failures Sys TP 1	-1.00	0.00	-0.03	0.33	-0.01	0.66	-0.01	0.82	0.04	0.22	0.00	0.96	-0.02	0.64
No. of Failures Sys TP 2	-0.99	0.00	-0.03	0.28	-0.01	0.78	0.00	0.96	0.04	0.22	0.00	0.99	-0.01	0.65
No. of Failures Sys TP 3	-0.97	0.00	-0.02	0.46	-0.02	0.63	-0.02	0.62	0.05	0.13	0.02	0.61	-0.01	0.78

Table 3-17: **Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix II-a**

Output	FF/ Target IOC rho	p-val	FF/ Target Affordability rho	p-val	EpiUn/ Prior Beta (alpha) rho	p-val	EpiUn/ Prior Beta (beta) rho	p-val	NFC TP/Component rho	p-val	EpiUn/ Failure Mechanisms Weight rho	p-val	CompTF/ No of Test Facility rho	p-val
MoE Reliability	0.03	0.29	0.02	0.61	0.11	0.00	-0.81	0.00	-0.03	0.32	0.09	0.00	0.02	0.46
MoE Affordability [MMU]	-0.01	0.76	0.00	0.91	0.05	0.08	-0.01	0.80	-0.04	0.20	0.00	0.88	0.01	0.76
MoE IOC [y]	0.05	0.14	0.03	0.31	0.06	0.06	-0.01	0.72	0.03	0.31	-0.01	0.85	-0.05	0.10
Service Life HW Tot	0.01	0.77	-0.04	0.23	0.02	0.50	0.02	0.49	0.01	0.77	-0.05	0.14	0.01	0.66
Development Cost (D&D), [MMU]	0.02	0.61	-0.02	0.61	-0.02	0.55	-0.03	0.43	0.07	0.03	0.02	0.44	-0.01	0.76
Development Cost (HW) [MMU]	-0.02	0.58	0.01	0.70	0.06	0.05	0.01	0.74	-0.06	0.08	0.00	0.93	0.02	0.52
Development Cost (DD+HW) [MMU]	-0.02	0.54	0.01	0.87	0.05	0.12	-0.01	0.87	-0.04	0.21	0.01	0.79	0.01	0.71
Development Cost (Testing) [MMU]	0.05	0.14	-0.01	0.81	0.01	0.73	0.00	0.93	0.00	1.00	0.01	0.82	-0.04	0.18
0.5-percentile (Median) System Level	0.03	0.33	0.01	0.64	0.11	0.00	-0.83	0.00	-0.03	0.34	0.09	0.01	0.02	0.44
DDE&T Cost (Total Dev) [MMU]	-0.01	0.76	0.00	0.91	0.05	0.08	-0.01	0.80	-0.04	0.20	0.00	0.88	0.01	0.76
p-th Percentile Rel Sys	0.03	0.29	0.02	0.61	0.11	0.00	-0.81	0.00	-0.03	0.32	0.09	0.00	0.02	0.46
Engine Level THW	-0.02	0.46	0.00	0.88	0.07	0.03	0.04	0.17	-0.06	0.04	-0.03	0.36	-0.02	0.61
TFU [MMU]	-0.01	0.80	0.00	0.91	0.02	0.56	-0.05	0.09	0.01	0.81	0.04	0.19	0.05	0.10
weighted FF	-0.13	0.00	-0.13	0.00	-0.07	0.02	0.02	0.61	0.00	0.99	0.01	0.74	0.03	0.39
BogeyEQL based Reliability	-0.01	0.65	0.03	0.39	-0.01	0.80	0.02	0.48	-0.03	0.40	-0.05	0.10	0.02	0.56
AUC [MMU]	-0.01	0.81	0.00	0.95	0.01	0.82	-0.04	0.19	0.01	0.80	0.03	0.28	0.05	0.11
p1 SubSys THW	-0.03	0.41	-0.02	0.54	0.04	0.24	0.02	0.51	-0.01	0.80	0.09	0.01	0.07	0.03
p2 SubSys THW	-0.03	0.41	-0.02	0.54	0.04	0.24	0.02	0.51	-0.01	0.80	0.09	0.01	0.07	0.03
Production Cost Tot [MMU]	0.00	0.98	-0.03	0.33	0.02	0.47	-0.01	0.80	0.02	0.61	-0.01	0.64	0.04	0.24
p3 SubSys THW	-0.03	0.41	-0.02	0.54	0.04	0.24	0.02	0.51	-0.01	0.80	0.09	0.01	0.07	0.03
Affordability [MMU]	0.00	1.00	-0.02	0.53	0.03	0.34	-0.01	0.84	0.01	0.74	-0.02	0.55	0.04	0.19
p5 SubSys THW	-0.03	0.41	-0.02	0.54	0.04	0.24	0.02	0.51	-0.01	0.80	0.09	0.01	0.07	0.03

Table 3-18:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix II-b

Output	FF/ Target IOC rho	p-val	FF/ Target Affordability rho	p-val	EpiUn/ Prior Beta (alpha) rho	p-val	EpiUn/ Prior Beta (beta) rho	p-val	NFC TP/Component rho	p-val	EpiUn/ Failure Mechanisms Weight rho	p-val	CompTF/ No of Test Facility rho	p-val
p6 SubSys THW	-0.03	0.41	-0.02	0.54	0.04	0.24	0.02	0.51	-0.01	0.80	0.09	0.01	0.07	0.03
p7 SubSys THW	-0.03	0.41	-0.02	0.54	0.04	0.24	0.02	0.51	-0.01	0.80	0.09	0.01	0.07	0.03
p8 SubSys THW	-0.03	0.41	-0.02	0.54	0.04	0.24	0.02	0.51	-0.01	0.80	0.09	0.01	0.07	0.03
p3 Comp THW	-0.04	0.21	0.02	0.55	0.04	0.17	0.04	0.26	0.29	0.00	0.19	0.00	-0.04	0.20
Risk Coeff Rel	0.02	0.63	0.03	0.30	-0.02	0.48	0.04	0.18	-0.03	0.33	-0.02	0.59	-0.02	0.62
Risk Coeff Dev Dur	1.00	0.00	-0.01	0.75	-0.01	0.83	0.00	0.88	0.00	0.97	0.02	0.60	0.03	0.37
Risk Coeff Dev Cost	-0.01	0.75	1.00	0.00	0.02	0.45	-0.02	0.45	0.06	0.06	0.02	0.56	0.05	0.10
Comp / Test Cost [MMU]	0.00	0.92	-0.03	0.34	-0.07	0.02	0.02	0.44	0.48	0.00	0.04	0.18	-0.02	0.55
Subsystem Test Cost [MMU]	0.05	0.15	0.00	0.91	0.03	0.29	0.00	0.96	0.00	0.93	0.04	0.25	0.03	0.36
System Test Cost [MMU]	0.05	0.15	-0.01	0.82	0.01	0.77	0.00	0.93	-0.03	0.29	0.00	0.97	-0.05	0.10
No. of Failures Sys TP 1	-0.02	0.43	0.05	0.14	-0.04	0.21	0.02	0.63	0.02	0.54	-0.01	0.86	-0.01	0.85
No. of Failures Sys TP 2	-0.03	0.35	0.04	0.16	-0.03	0.27	0.02	0.54	0.02	0.62	0.00	1.00	-0.01	0.83
No. of Failures Sys TP 3	-0.02	0.60	0.03	0.36	-0.04	0.26	0.02	0.51	0.01	0.64	0.00	0.97	-0.01	0.69

Table 3-19: **Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix III-a**

Output	CompTF/ Initial or Modification Facility Cost		CompTF/ No of People operating Facility		CompTF/ Hourly Rate [MU/h]		CompTF/ Yearly Working Hours		CompTF/ Test Cadence		CompTF/ Yearly Maintenance Cost [MMU]		NFC TP/ Subsystem	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
MoE Reliability	-0.01	0.77	0.02	0.57	0.02	0.60	-0.02	0.56	-0.05	0.09	-0.01	0.69	0.06	0.06
MoE Affordability [MMU]	0.00	0.90	0.01	0.84	0.04	0.21	0.04	0.26	-0.01	0.70	0.06	0.08	-0.04	0.26
MoE IOC [y]	0.04	0.24	0.00	0.88	0.02	0.46	-0.02	0.48	0.00	0.90	-0.02	0.50	-0.05	0.08
Service Life HW Tot	0.06	0.07	-0.03	0.27	-0.04	0.24	-0.03	0.42	-0.01	0.86	0.02	0.43	0.03	0.28
Development Cost (D&D), [MMU]	0.07	0.04	-0.02	0.50	0.02	0.61	0.03	0.36	-0.02	0.62	-0.01	0.73	-0.05	0.13
Development Cost (HW) [MMU]	-0.03	0.30	0.01	0.85	0.04	0.16	0.02	0.61	-0.02	0.62	0.07	0.04	-0.04	0.24
Development Cost (DD+HW) [MMU]	-0.01	0.68	-0.01	0.86	0.04	0.26	0.04	0.25	-0.02	0.61	0.06	0.07	-0.04	0.18
Development Cost (Testing) [MMU]	0.06	0.07	0.08	0.01	0.02	0.48	-0.01	0.66	0.02	0.52	-0.01	0.74	0.04	0.23
0.5-percentile (Median) System Level	-0.01	0.75	0.02	0.57	0.02	0.53	-0.02	0.53	-0.05	0.10	-0.01	0.71	0.06	0.07
DDE&T Cost (Total Dev) [MMU]	0.00	0.90	0.01	0.84	0.04	0.21	0.04	0.26	-0.01	0.70	0.06	0.08	-0.04	0.26
p-th Percentile Rel Sys	-0.01	0.77	0.02	0.57	0.02	0.60	-0.02	0.56	-0.05	0.09	-0.01	0.69	0.06	0.06
Engine Level THW	-0.02	0.47	0.00	0.94	0.01	0.75	0.01	0.68	0.01	0.83	0.03	0.37	-0.04	0.23
TFU [MMU]	-0.03	0.40	0.01	0.70	0.05	0.15	0.01	0.85	-0.03	0.30	0.07	0.03	-0.02	0.62
weighted FF	-0.05	0.10	-0.03	0.27	-0.04	0.17	0.01	0.65	-0.01	0.81	0.02	0.57	0.04	0.16
BogeyEQL based Reliability	0.04	0.27	0.01	0.85	0.04	0.25	0.00	1.00	0.00	0.99	0.06	0.04	0.02	0.45
AUC [MMU]	-0.03	0.40	0.01	0.74	0.04	0.17	0.00	1.00	-0.03	0.39	0.06	0.08	-0.02	0.61
p1 SubSys THW	-0.02	0.58	0.00	0.92	-0.01	0.64	-0.05	0.14	0.00	0.91	0.06	0.08	0.73	0.00
p2 SubSys THW	-0.02	0.58	0.00	0.92	-0.01	0.64	-0.05	0.14	0.00	0.91	0.06	0.08	0.73	0.00
Production Cost Tot [MMU]	0.03	0.41	-0.02	0.51	-0.01	0.77	-0.02	0.45	-0.01	0.74	0.06	0.08	0.02	0.51
p3 SubSys THW	-0.02	0.58	0.00	0.92	-0.01	0.64	-0.05	0.14	0.00	0.91	0.06	0.08	0.73	0.00
Affordability [MMU]	0.02	0.52	-0.02	0.57	0.00	0.89	-0.02	0.60	-0.02	0.53	0.07	0.02	0.01	0.73
p5 SubSys THW	-0.02	0.58	0.00	0.92	-0.01	0.64	-0.05	0.14	0.00	0.91	0.06	0.08	0.73	0.00

Table 3-20:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix III-b

Output	CompTF/ Initial or Modification Facility Cost		CompTF/ No of People operating Facility		CompTF/ Hourly Rate [MU/h]		CompTF/ Yearly Working Hours		CompTF/ Test Cadence		CompTF/ Yearly Maintenance Cost [MMU]		NFC TP/ Subsystem	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
p6 SubSys THW	-0.02	0.58	0.00	0.92	-0.01	0.64	-0.05	0.14	0.00	0.91	0.06	0.08	0.73	0.00
p7 SubSys THW	-0.02	0.58	0.00	0.92	-0.01	0.64	-0.05	0.14	0.00	0.91	0.06	0.08	0.73	0.00
p8 SubSys THW	-0.02	0.58	0.00	0.92	-0.01	0.64	-0.05	0.14	0.00	0.91	0.06	0.08	0.73	0.00
p3 Comp THW	-0.01	0.69	0.03	0.29	0.05	0.13	0.00	0.91	-0.02	0.47	0.08	0.02	-0.02	0.61
Risk Coeff Rel	0.05	0.14	0.02	0.61	0.00	0.90	-0.02	0.52	0.01	0.83	-0.06	0.07	-0.02	0.44
Risk Coeff Dev Dur	-0.01	0.87	0.03	0.42	-0.01	0.76	0.01	0.74	0.02	0.45	0.00	0.97	0.03	0.36
Risk Coeff Dev Cost	-0.07	0.02	0.02	0.57	-0.04	0.26	0.04	0.16	0.00	0.91	-0.02	0.47	-0.04	0.26
Comp / Test Cost [MMU]	0.76	0.00	0.26	0.00	0.12	0.00	0.11	0.00	-0.27	0.00	0.09	0.00	0.01	0.67
Subsystem Test Cost [MMU]	-0.01	0.86	0.01	0.68	-0.01	0.74	-0.04	0.16	-0.06	0.04	-0.01	0.71	0.43	0.00
System Test Cost [MMU]	0.02	0.62	0.07	0.03	0.02	0.62	-0.02	0.54	0.05	0.10	-0.01	0.68	-0.03	0.28
No. of Failures Sys TP 1	0.01	0.73	-0.04	0.23	0.04	0.17	0.01	0.77	-0.01	0.85	0.08	0.01	0.02	0.51
No. of Failures Sys TP 2	0.02	0.48	-0.05	0.15	0.04	0.17	0.01	0.76	-0.01	0.84	0.08	0.01	0.03	0.36
No. of Failures Sys TP 3	0.00	0.92	-0.04	0.20	0.04	0.18	0.01	0.70	0.00	0.94	0.07	0.03	0.01	0.76

Table 3-21: Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix IV-a

Output	Subsystem/TF/ No of Test Facility		Subsystem/TF/ Initial or Modification Facility Cost		Subsystem/TF/ No of People operating Facility		Subsystem/TF/ Hourly Rate [MU/h]		Subsystem/TF/ Yearly Working Hours		Subsystem/TF/ Test Cadence		Subsystem/TF/ Yearly Maintenance Cost [MMU]	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
MoE Reliability	0.01	0.85	0.02	0.49	0.00	0.93	0.04	0.18	-0.01	0.77	0.018	0.572	0.023	0.477
MoE Affordability [MMU]	0.03	0.34	0.01	0.85	-0.03	0.39	0.05	0.13	-0.03	0.35	0.02	0.50	0.02	0.43
MoE IOC [y]	0.01	0.68	0.01	0.83	-0.02	0.57	0.00	0.94	0.03	0.32	-0.01	0.67	0.01	0.65
Service Life HW Tot	-0.04	0.20	-0.06	0.08	0.01	0.66	-0.05	0.12	-0.05	0.09	0.02	0.63	0.02	0.57
Development Cost (D&D), [MMU]	0.06	0.04	-0.02	0.53	-0.03	0.36	-0.01	0.82	0.00	0.87	0.03	0.42	0.00	0.93
Development Cost (HW) [MMU]	0.01	0.79	-0.01	0.84	-0.03	0.37	0.06	0.06	-0.01	0.66	0.02	0.46	0.02	0.58
Development Cost (DD+HW) [MMU]	0.03	0.38	-0.01	0.84	-0.04	0.26	0.05	0.12	-0.03	0.31	0.03	0.39	0.02	0.47
Development Cost (Testing) [MMU]	0.04	0.26	0.13	0.00	0.00	1.00	0.02	0.61	0.02	0.60	-0.04	0.21	-0.02	0.61
0.5-percentile (Median) System Level	0.00	0.91	0.02	0.51	0.00	0.94	0.04	0.20	-0.01	0.75	0.02	0.62	0.03	0.39
DDE&T Cost (Total Dev) [MMU]	0.03	0.34	0.01	0.85	-0.03	0.39	0.05	0.13	-0.03	0.35	0.02	0.50	0.02	0.43
p-th Percentile Rel Sys	0.01	0.85	0.02	0.49	0.00	0.93	0.04	0.18	-0.01	0.77	0.02	0.57	0.02	0.48
Engine Level THW	0.02	0.52	0.01	0.73	-0.03	0.29	0.04	0.16	-0.02	0.63	0.03	0.40	-0.03	0.37
TFU [MMU]	-0.02	0.63	-0.02	0.58	0.03	0.36	0.03	0.31	0.01	0.79	0.02	0.57	0.10	0.00
weighted FF	-0.02	0.63	0.00	0.93	-0.02	0.53	0.00	0.95	0.02	0.60	0.01	0.75	-0.02	0.62
BogeyEQL based Reliability	0.02	0.59	-0.05	0.14	0.00	0.94	0.04	0.19	-0.07	0.04	0.01	0.82	-0.03	0.43
AUC [MMU]	-0.01	0.65	-0.01	0.74	0.03	0.36	0.02	0.52	0.02	0.58	0.01	0.72	0.11	0.00
p1 SubSys THW	0.05	0.10	-0.04	0.26	-0.04	0.23	-0.01	0.68	0.02	0.62	0.01	0.87	-0.06	0.07
p2 SubSys THW	0.05	0.10	-0.04	0.26	-0.04	0.23	-0.01	0.68	0.02	0.62	0.01	0.87	-0.06	0.07
Production Cost Tot [MMU]	-0.05	0.14	-0.06	0.07	0.03	0.33	-0.03	0.36	-0.03	0.34	0.02	0.61	0.07	0.02
p3 SubSys THW	0.05	0.10	-0.04	0.26	-0.04	0.23	-0.01	0.68	0.02	0.62	0.01	0.87	-0.06	0.07
Affordability [MMU]	-0.03	0.35	-0.05	0.10	0.03	0.38	-0.01	0.65	-0.03	0.28	0.02	0.51	0.07	0.02
p5 SubSys THW	0.05	0.10	-0.04	0.26	-0.04	0.23	-0.01	0.68	0.02	0.62	0.01	0.87	-0.06	0.07

Table 3-22:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix IV-b

Output	Subsystem TF / No of Test Facility		Subsystem TF / Initial or Modification Facility Cost		Subsystem TF / No of People operating Facility		Subsystem TF / Hourly Rate [MU/h]		Subsystem TF / Yearly Working Hours		Subsystem TF / Test Cadence		Subsystem TF / Yearly Maintenance Cost [MMU]	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
p6 SubSys THW	0.05	0.10	-0.04	0.26	-0.04	0.23	-0.01	0.68	0.02	0.62	0.01	0.87	-0.06	0.07
p7 SubSys THW	0.05	0.10	-0.04	0.26	-0.04	0.23	-0.01	0.68	0.02	0.62	0.01	0.87	-0.06	0.07
p8 SubSys THW	0.05	0.10	-0.04	0.26	-0.04	0.23	-0.01	0.68	0.02	0.62	0.01	0.87	-0.06	0.07
p3 Comp THW	0.03	0.33	-0.02	0.52	-0.08	0.01	0.00	0.97	0.01	0.87	0.02	0.45	-0.03	0.36
Risk Coeff Rel	0.04	0.20	-0.02	0.48	0.00	0.96	-0.01	0.69	0.00	0.93	0.02	0.53	-0.04	0.24
Risk Coeff Dev Dur	-0.03	0.34	0.04	0.27	-0.02	0.56	0.01	0.69	-0.04	0.24	0.05	0.13	0.02	0.46
Risk Coeff Dev Cost	-0.05	0.12	0.02	0.48	-0.02	0.63	0.01	0.72	0.00	0.91	-0.03	0.36	-0.01	0.85
Comp / Test Cost [MMU]	0.00	0.96	0.01	0.81	0.01	0.87	0.01	0.65	0.00	0.89	0.00	0.88	0.01	0.82
Subsystem Test Cost [MMU]	0.01	0.74	0.86	0.00	0.09	0.01	0.10	0.00	0.03	0.34	-0.07	0.02	-0.01	0.87
System Test Cost [MMU]	0.03	0.27	-0.01	0.75	-0.01	0.67	0.00	0.89	0.01	0.75	-0.03	0.34	-0.01	0.64
No. of Failures Sys TP 1	0.02	0.63	-0.02	0.55	-0.02	0.59	0.00	0.93	0.00	0.89	0.01	0.74	-0.02	0.46
No. of Failures Sys TP 2	0.02	0.43	-0.02	0.59	-0.02	0.54	-0.01	0.79	0.01	0.83	0.01	0.66	-0.03	0.41
No. of Failures Sys TP 3	0.02	0.60	-0.01	0.83	-0.01	0.67	0.01	0.84	0.01	0.77	0.01	0.77	-0.02	0.51

Table 3-23: Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix V-a

Output	NFC TP / System]		System TF / No of Test Facility		System TF / Initial or Modification Facility Cost		System TF / No of People operating Facility		System TF / Hourly Rate [MU/h]		System TF / Yearly Working Hours		Subsystem TF / Test Cadence	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
MoE Reliability	0.089	0.005	0.00	0.94	0.02	0.452	0.00	0.892	0.00	0.98	0.07	0.034	-0.02	0.442
MoE Affordability [MMU]	0.05	0.09	-0.03	0.35	0.08	0.01	-0.01	0.76	0.08	0.02	0.01	0.71	-0.04	0.20
MoE IOC [y]	0.13	0.00	-0.79	0.00	-0.03	0.36	-0.01	0.84	0.04	0.20	0.02	0.43	-0.38	0.00
Service Life HW Tot	0.03	0.27	0.04	0.20	-0.03	0.42	0.00	0.96	0.04	0.21	0.05	0.15	-0.06	0.06
Development Cost (D&D), [MMU]	0.03	0.40	0.01	0.71	-0.03	0.38	0.01	0.79	0.05	0.12	-0.01	0.64	-0.02	0.44
Development Cost (HW) [MMU]	0.04	0.19	0.03	0.30	0.03	0.40	-0.02	0.45	0.03	0.30	-0.01	0.86	0.00	0.91
Development Cost (DD+HW) [MMU]	0.04	0.18	0.04	0.18	0.01	0.69	-0.02	0.55	0.05	0.12	-0.01	0.82	0.00	0.97
Development Cost (Testing) [MMU]	0.10	0.00	-0.48	0.00	0.50	0.00	0.07	0.03	0.14	0.00	0.13	0.00	-0.24	0.00
0.5-percentile (Median) System Level	0.09	0.01	0.00	0.92	0.02	0.44	0.00	0.89	0.00	0.91	0.06	0.04	-0.03	0.41
DDE&T Cost (Total Dev) [MMU]	0.05	0.09	-0.03	0.35	0.08	0.01	-0.01	0.76	0.08	0.02	0.01	0.71	-0.04	0.20
p-th Percentile Rel Sys	0.09	0.00	0.00	0.94	0.02	0.45	0.00	0.89	0.00	0.98	0.07	0.03	-0.02	0.44
Engine Level THW	0.04	0.21	0.00	0.99	0.03	0.38	-0.01	0.75	0.04	0.21	-0.01	0.84	0.04	0.23
TFU [MMU]	0.02	0.60	0.08	0.01	-0.01	0.79	-0.02	0.52	0.01	0.83	0.01	0.87	-0.03	0.31
weighted FF	-0.08	0.01	0.55	0.00	-0.02	0.50	0.01	0.81	-0.08	0.01	-0.02	0.52	0.24	0.00
BogeyEQL based Reliability	0.01	0.68	0.03	0.34	0.00	0.99	-0.01	0.78	0.07	0.04	-0.09	0.00	0.01	0.71
AUC [MMU]	0.03	0.34	0.07	0.03	-0.02	0.61	-0.03	0.41	0.01	0.69	0.00	0.94	-0.03	0.36
p1 SubSys THW	-0.04	0.19	-0.01	0.86	-0.01	0.78	0.02	0.43	0.01	0.87	-0.04	0.17	0.05	0.12
p2 SubSys THW	-0.04	0.19	-0.01	0.86	-0.01	0.78	0.02	0.43	0.01	0.87	-0.04	0.17	0.05	0.12
Production Cost Tot [MMU]	0.04	0.22	0.08	0.02	-0.04	0.27	-0.02	0.58	0.04	0.20	0.04	0.23	-0.07	0.03
p3 SubSys THW	-0.04	0.19	-0.01	0.86	-0.01	0.78	0.02	0.43	0.01	0.87	-0.04	0.17	0.05	0.12
Affordability [MMU]	0.05	0.14	0.06	0.05	-0.02	0.52	-0.02	0.49	0.06	0.07	0.03	0.27	-0.07	0.02
p5 SubSys THW	-0.04	0.19	-0.01	0.86	-0.01	0.78	0.02	0.43	0.01	0.87	-0.04	0.17	0.05	0.12

Table 3-24:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix V-b

Output	NFC TP / System1 rho	p-val	SystemTF / No of Test Facility rho	p-val	SystemTF / Initial or Modification Facility Cost rho	p-val	SystemTF / No of People operating Facility rho	p-val	SystemTF / Hourly Rate [MU/h] rho	p-val	SystemTF / Yearly Working Hours rho	p-val	SubsystemTF / Test Cadence rho	p-val
p6 SubSys THW	-0.04	0.19	-0.01	0.86	-0.01	0.78	0.02	0.43	0.01	0.87	-0.04	0.17	0.05	0.12
p7 SubSys THW	-0.04	0.19	-0.01	0.86	-0.01	0.78	0.02	0.43	0.01	0.87	-0.04	0.17	0.05	0.12
p8 SubSys THW	-0.04	0.19	-0.01	0.86	-0.01	0.78	0.02	0.43	0.01	0.87	-0.04	0.17	0.05	0.12
p3 Comp THW	0.01	0.78	-0.03	0.42	0.00	0.99	-0.04	0.25	0.03	0.39	-0.05	0.12	0.01	0.75
Risk Coeff Rel	0.03	0.33	0.00	0.99	0.05	0.14	0.00	0.88	0.01	0.82	-0.01	0.78	0.02	0.52
Risk Coeff Dev Dur	-0.09	0.00	0.00	0.91	0.02	0.59	0.03	0.39	-0.05	0.12	0.01	0.82	-0.04	0.16
Risk Coeff Dev Cost	0.05	0.10	-0.02	0.51	-0.05	0.11	0.02	0.52	-0.04	0.20	0.00	0.99	-0.03	0.33
Comp / Test Cost [MMU]	-0.04	0.27	-0.04	0.27	-0.01	0.72	-0.01	0.87	0.00	0.95	0.00	0.92	0.02	0.64
Subsystem Test Cost [MMU]	-0.02	0.46	-0.01	0.65	-0.02	0.54	-0.03	0.42	-0.03	0.34	-0.01	0.69	0.04	0.18
System Test Cost [MMU]	0.10	0.00	-0.49	0.00	0.52	0.00	0.07	0.02	0.15	0.00	0.14	0.00	-0.25	0.00
No. of Failures Sys TP 1	0.01	0.64	0.08	0.01	-0.03	0.38	-0.04	0.25	0.04	0.18	-0.03	0.32	-0.01	0.71
No. of Failures Sys TP 2	0.02	0.61	0.08	0.01	-0.03	0.39	-0.05	0.12	0.03	0.28	-0.03	0.36	-0.01	0.75
No. of Failures Sys TP 3	0.02	0.57	0.07	0.02	-0.04	0.26	-0.03	0.38	0.04	0.21	-0.04	0.25	-0.02	0.48

Table 3-25:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix VI-a

Output	System TF / Yearly Maintenance Cost [MMU] rho	p-val	NFC TP / System2 rho	p-val	NFC TP / System3 rho	p-val	NFC TP / System4 rho	p-val	NFC TP / System5 rho	p-val	NFC TP / System6 rho	p-val	EpiUn / SF rho	p-val
MoE Reliability	-0.01	0.764	0.15	0.00	0.11	0.00	0.16	0.00	0.22	0.00	0.32	0.00	0.00	0.99
MoE Affordability [MMU]	0.07	0.02	0.08	0.01	0.06	0.05	0.08	0.01	0.13	0.00	0.18	0.00	-0.01	0.75
MoE IOC [y]	0.00	0.96	0.05	0.14	0.03	0.38	0.09	0.00	0.11	0.00	0.12	0.00	0.07	0.03
Service Life HW Tot	0.07	0.03	0.00	0.96	-0.01	0.87	0.01	0.77	-0.02	0.52	0.00	0.90	-0.02	0.52
Development Cost (D&D), [MMU]	0.01	0.74	-0.03	0.28	-0.03	0.29	-0.05	0.15	0.02	0.49	0.03	0.32	0.01	0.75
Development Cost (HW) [MMU]	0.06	0.06	0.09	0.01	0.06	0.06	0.07	0.03	0.11	0.00	0.15	0.00	-0.02	0.55
Development Cost (DD+HW) [MMU]	0.07	0.04	0.07	0.02	0.05	0.12	0.05	0.10	0.11	0.00	0.15	0.00	-0.02	0.51
Development Cost (Testing) [MMU]	0.04	0.16	0.08	0.02	0.09	0.00	0.20	0.00	0.16	0.00	0.27	0.00	0.05	0.12
0.5-percentile (Median) System Level	-0.01	0.81	0.15	0.00	0.10	0.00	0.15	0.00	0.22	0.00	0.30	0.00	0.00	0.90
DDE&T Cost (Total Dev) [MMU]	0.07	0.02	0.08	0.01	0.06	0.05	0.08	0.01	0.13	0.00	0.18	0.00	-0.01	0.75
p-th Percentile Rel Sys	-0.01	0.76	0.15	0.00	0.11	0.00	0.16	0.00	0.22	0.00	0.32	0.00	0.00	0.99
Engine Level THW	0.05	0.09	0.11	0.00	0.07	0.03	0.12	0.00	0.13	0.00	0.19	0.00	-0.02	0.57
TFU [MMU]	0.01	0.67	-0.02	0.57	0.00	1.00	-0.04	0.17	0.01	0.71	-0.02	0.55	0.01	0.72
weighted FF	-0.02	0.51	-0.08	0.01	-0.03	0.28	-0.09	0.00	-0.13	0.00	-0.17	0.00	-0.05	0.14
BogeyEQL based Reliability	0.04	0.21	0.06	0.05	-0.02	0.48	-0.01	0.82	-0.02	0.63	0.01	0.77	-0.49	0.00
AUC [MMU]	0.01	0.67	-0.02	0.50	0.00	0.98	-0.03	0.30	0.01	0.75	-0.01	0.79	0.01	0.81
p1 SubSys THW	0.01	0.85	0.02	0.48	-0.03	0.36	-0.02	0.62	-0.03	0.41	0.01	0.81	-0.02	0.57
p2 SubSys THW	0.01	0.85	0.02	0.48	-0.03	0.36	-0.02	0.62	-0.03	0.41	0.01	0.81	-0.02	0.57
Production Cost Tot [MMU]	0.06	0.04	0.00	0.88	-0.01	0.71	-0.02	0.63	-0.01	0.76	-0.01	0.69	-0.01	0.75
p3 SubSys THW	0.01	0.85	0.02	0.48	-0.03	0.36	-0.02	0.62	-0.03	0.41	0.01	0.81	-0.02	0.57
Affordability [MMU]	0.07	0.02	0.02	0.54	0.00	0.98	0.00	0.93	0.02	0.49	0.03	0.31	-0.02	0.54
p5 SubSys THW	0.01	0.85	0.02	0.48	-0.03	0.36	-0.02	0.62	-0.03	0.41	0.01	0.81	-0.02	0.57

Table 3-26: Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix VI-b

Output	SystemTF / Yearly Maintenance Cost [MMU] rho	p-val	NFC TP / System2 rho	p-val	NFC TP / System3 rho	p-val	NFC TP / System4 rho	p-val	NFC TP / System5 rho	p-val	NFC TP / System6 rho	p-val	EpiUn / SF rho	p-val
p6 SubSys THW	0.01	0.85	0.02	0.48	-0.03	0.36	-0.02	0.62	-0.03	0.41	0.01	0.81	-0.02	0.57
p7 SubSys THW	0.01	0.85	0.02	0.48	-0.03	0.36	-0.02	0.62	-0.03	0.41	0.01	0.81	-0.02	0.57
p8 SubSys THW	0.01	0.85	0.02	0.48	-0.03	0.36	-0.02	0.62	-0.03	0.41	0.01	0.81	-0.02	0.57
p3 Comp THW	0.08	0.01	0.01	0.74	-0.01	0.81	0.01	0.75	0.00	0.98	0.00	0.92	0.00	0.97
Risk Coeff Rel	-0.01	0.70	-0.06	0.06	0.02	0.60	-0.01	0.72	0.00	0.89	-0.01	0.79	0.01	0.67
Risk Coeff Dev Dur	-0.03	0.29	0.05	0.11	0.02	0.52	0.04	0.20	0.06	0.04	0.02	0.46	-0.01	0.76
Risk Coeff Dev Cost	0.00	0.91	-0.01	0.72	0.08	0.02	-0.02	0.44	0.00	0.88	-0.04	0.24	-0.02	0.57
Comp / Test Cost [MMU]	0.02	0.63	0.01	0.69	-0.01	0.78	0.01	0.79	0.00	0.98	0.03	0.38	-0.02	0.56
Subsystem Test Cost [MMU]	-0.03	0.40	0.07	0.02	0.01	0.64	0.01	0.64	-0.02	0.49	0.02	0.58	0.04	0.23
System Test Cost [MMU]	0.05	0.15	0.07	0.03	0.09	0.00	0.21	0.00	0.17	0.00	0.27	0.00	0.04	0.16
No. of Failures Sys TP 1	-0.01	0.75	0.06	0.06	-0.04	0.17	0.02	0.62	-0.05	0.12	0.01	0.79	-0.02	0.53
No. of Failures Sys TP 2	-0.02	0.58	0.06	0.06	-0.05	0.14	0.02	0.59	-0.05	0.13	0.01	0.80	-0.02	0.49
No. of Failures Sys TP 3	-0.02	0.61	0.05	0.14	-0.04	0.21	0.02	0.47	-0.04	0.25	0.01	0.74	-0.02	0.60

Table 3-27: **Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix VII-a**

Output	Shape Parameter p1 EpiUn/Weibull rho	p-val	EpiUn/Capacity Weight rho	p-val	Shape Parameter p2 EpiUn/Weibull rho	p-val	Shape Parameter p3 EpiUn/Weibull rho	p-val	Shape Parameter p4 EpiUn/Weibull rho	p-val	Shape Parameter p5 EpiUn/Weibull rho	p-val	Shape Parameter p6 EpiUn/Weibull rho	p-val
MoE Reliability	-0.01	0.73	0.00	0.92	0.04	0.22	0.03	0.35	0.00	0.97	0.00	0.96	0.02	0.51
MoE Affordability [MMU]	0.03	0.32	-0.03	0.31	-0.01	0.80	0.02	0.56	0.05	0.13	-0.02	0.57	0.02	0.46
MoE IOC [y]	-0.10	0.00	-0.03	0.39	0.04	0.16	0.04	0.26	0.01	0.69	-0.08	0.01	-0.01	0.76
Service Life HW Tot	-0.01	0.79	0.04	0.25	0.01	0.77	0.02	0.48	-0.03	0.42	0.01	0.65	-0.01	0.72
Development Cost (D&D), [MMU]	0.00	1.00	-0.03	0.38	-0.06	0.04	0.01	0.75	0.00	0.88	-0.04	0.23	0.01	0.78
Development Cost (HW) [MMU]	0.03	0.31	-0.04	0.20	0.01	0.64	0.02	0.54	0.05	0.13	0.00	0.89	0.03	0.40
Development Cost (DD+HW) [MMU]	0.03	0.31	-0.03	0.28	-0.01	0.76	0.02	0.58	0.04	0.17	-0.01	0.72	0.03	0.38
Development Cost (Testing) [MMU]	-0.01	0.72	0.00	0.95	-0.01	0.87	0.01	0.86	0.03	0.36	-0.05	0.12	0.00	0.96
0.5-percentile (Median) System Level	-0.01	0.79	0.00	0.94	0.04	0.23	0.03	0.34	0.00	0.95	0.00	0.92	0.02	0.53
DDE&T Cost (Total Dev) [MMU]	0.03	0.32	-0.03	0.31	-0.01	0.80	0.02	0.56	0.05	0.13	-0.02	0.57	0.02	0.46
p-th Percentile Rel Sys	-0.01	0.73	0.00	0.92	0.04	0.22	0.03	0.35	0.00	0.97	0.00	0.96	0.02	0.51
Engine Level THW	0.04	0.25	-0.04	0.22	0.02	0.44	0.06	0.06	0.05	0.15	0.00	0.95	0.02	0.48
TFU [MMU]	0.01	0.84	0.00	0.90	-0.01	0.68	-0.05	0.11	0.01	0.72	-0.01	0.74	0.00	0.97
weighted FF	0.02	0.47	-0.01	0.85	-0.02	0.57	-0.01	0.86	-0.04	0.21	0.05	0.09	0.00	0.97
BogeyEQL based Reliability	0.01	0.74	0.00	0.88	-0.04	0.26	0.05	0.12	0.00	0.88	0.00	0.95	0.01	0.74
AUC [MMU]	0.01	0.78	-0.03	0.39	-0.01	0.68	-0.06	0.08	0.03	0.41	-0.01	0.68	0.01	0.76
p1 SubSys THW	0.06	0.07	0.02	0.55	-0.03	0.37	0.05	0.14	0.00	0.99	-0.01	0.68	-0.01	0.87
p2 SubSys THW	0.06	0.07	0.02	0.55	-0.03	0.37	0.05	0.14	0.00	0.99	-0.01	0.68	-0.01	0.87
Production Cost Tot [MMU]	-0.01	0.67	0.01	0.71	0.00	0.97	-0.02	0.60	0.00	0.93	0.00	0.96	0.00	0.99
p3 SubSys THW	0.06	0.07	0.02	0.55	-0.03	0.37	0.05	0.14	0.00	0.99	-0.01	0.68	-0.01	0.87
Affordability [MMU]	-0.01	0.83	0.00	0.97	0.00	0.92	-0.01	0.80	0.01	0.77	0.00	0.90	0.01	0.71
p5 SubSys THW	0.06	0.07	0.02	0.55	-0.03	0.37	0.05	0.14	0.00	0.99	-0.01	0.68	-0.01	0.87

Table 3-28: **Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix VII-b**

Output	EpiUn/Weibull Shape Parameter p1		EpiUn/Capacity Weight		EpiUn/Weibull Shape Parameter p2		EpiUn/Weibull Shape Parameter p3		EpiUn/Weibull Shape Parameter p4		EpiUn/Weibull Shape Parameter p5		EpiUn/Weibull Shape Parameter p6	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
p6 SubSys THW	0.06	0.07	0.02	0.55	-0.03	0.37	0.05	0.14	0.00	0.99	-0.01	0.68	-0.01	0.87
p7 SubSys THW	0.06	0.07	0.02	0.55	-0.03	0.37	0.05	0.14	0.00	0.99	-0.01	0.68	-0.01	0.87
p8 SubSys THW	0.06	0.07	0.02	0.55	-0.03	0.37	0.05	0.14	0.00	0.99	-0.01	0.68	-0.01	0.87
p3 Comp THW	0.02	0.62	-0.06	0.05	-0.02	0.58	0.22	0.00	0.02	0.44	-0.01	0.72	0.04	0.21
Risk Coeff Rel	0.00	0.95	0.02	0.59	-0.01	0.65	0.03	0.31	0.01	0.85	0.01	0.67	-0.05	0.12
Risk Coeff Dev Dur	0.01	0.78	0.01	0.74	0.03	0.39	0.01	0.83	0.01	0.86	-0.01	0.71	0.02	0.58
Risk Coeff Dev Cost	-0.02	0.60	0.00	0.91	0.03	0.39	-0.01	0.84	-0.04	0.22	0.02	0.54	-0.01	0.72
Comp / Test Cost [MMU]	-0.03	0.33	0.07	0.02	-0.03	0.31	-0.03	0.28	0.00	0.88	-0.04	0.24	0.02	0.43
Subsystem Test Cost [MMU]	0.03	0.31	0.00	0.89	-0.08	0.01	-0.01	0.84	-0.01	0.67	-0.03	0.36	0.00	0.95
System Test Cost [MMU]	-0.02	0.60	-0.01	0.87	0.01	0.64	0.01	0.82	0.03	0.28	-0.04	0.16	0.00	0.97
No. of Failures Sys TP 1	0.03	0.33	-0.04	0.27	0.00	0.97	0.01	0.65	0.01	0.76	0.02	0.56	-0.03	0.27
No. of Failures Sys TP 2	0.03	0.33	-0.03	0.37	-0.01	0.87	0.01	0.75	0.01	0.87	0.02	0.47	-0.03	0.30
No. of Failures Sys TP 3	0.04	0.25	-0.05	0.13	0.01	0.83	0.02	0.55	0.01	0.77	0.02	0.61	-0.04	0.23

Table 3-29:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix VIII-a

Output	EpiUn/Weibull Shape Parameter p7		EpiUn/Weibull Shape Parameter p8		EpiUn/Weibull Shape Parameter p9		LRECM/Dev Environment		LRECM/Mfg Environment		LRECM/Mfg MRL		LRECM/Design Scope	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
MoE Reliability	0.04	0.22	0.05	0.13	-0.04	0.21	0.05	0.12	0.03	0.42	0.03	0.30	0.03	0.31
MoE Affordability [MMU]	-0.01	0.77	-0.02	0.45	-0.03	0.42	-0.03	0.27	-0.03	0.29	0.03	0.34	0.24	0.00
MoE IOC [y]	0.01	0.67	0.04	0.18	0.01	0.85	-0.03	0.30	0.00	0.95	0.01	0.84	0.32	0.00
Service Life HW Tot	0.02	0.63	0.01	0.86	0.02	0.50	-0.02	0.51	-0.04	0.21	0.02	0.45	-0.02	0.55
Development Cost (D&D), [MMU]	0.04	0.17	0.01	0.76	0.06	0.06	-0.22	0.00	0.05	0.12	0.01	0.83	0.93	0.00
Development Cost (HW) [MMU]	-0.02	0.45	-0.03	0.29	-0.05	0.13	0.03	0.31	-0.05	0.10	0.03	0.32	-0.01	0.73
Development Cost (DD+HW) [MMU]	-0.01	0.75	-0.03	0.37	-0.02	0.53	-0.04	0.24	-0.03	0.30	0.03	0.29	0.25	0.00
Development Cost (Testing) [MMU]	0.01	0.70	0.04	0.26	-0.02	0.46	-0.02	0.58	0.00	0.92	0.01	0.84	-0.01	0.74
0.5-percentile (Median) System Level	0.04	0.21	0.05	0.14	-0.04	0.24	0.05	0.10	0.02	0.47	0.03	0.28	0.03	0.28
DDE&T Cost (Total Dev) [MMU]	-0.01	0.77	-0.02	0.45	-0.03	0.42	-0.03	0.27	-0.03	0.29	0.03	0.34	0.24	0.00
p-th Percentile Rel Sys	0.04	0.22	0.05	0.13	-0.04	0.21	0.05	0.12	0.03	0.42	0.03	0.30	0.03	0.31
Engine Level THW	-0.02	0.51	0.01	0.77	-0.05	0.15	0.01	0.83	-0.02	0.46	-0.01	0.82	0.00	0.98
TFU [MMU]	0.00	0.94	-0.08	0.02	-0.03	0.32	0.05	0.13	-0.07	0.03	0.07	0.04	0.00	0.89
weighted FF	0.01	0.72	-0.04	0.19	0.03	0.34	0.02	0.58	0.04	0.21	0.01	0.71	-0.18	0.00
BogeyEQL based Reliability	-0.05	0.08	-0.01	0.78	0.04	0.27	0.02	0.56	-0.02	0.49	0.01	0.69	0.00	0.90
AUC [MMU]	0.01	0.86	-0.07	0.04	-0.02	0.45	0.03	0.34	-0.07	0.03	0.06	0.06	0.01	0.83
p1 SubSys THW	-0.04	0.20	-0.03	0.41	0.01	0.87	0.05	0.14	-0.01	0.67	-0.03	0.35	-0.04	0.22
p2 SubSys THW	-0.04	0.20	-0.03	0.41	0.01	0.87	0.05	0.14	-0.01	0.67	-0.03	0.35	-0.04	0.22
Production Cost Tot [MMU]	0.02	0.47	-0.03	0.33	0.01	0.75	0.01	0.79	-0.07	0.02	0.05	0.09	-0.01	0.79
p3 SubSys THW	-0.04	0.20	-0.03	0.41	0.01	0.87	0.05	0.14	-0.01	0.67	-0.03	0.35	-0.04	0.22
Affordability [MMU]	0.02	0.62	-0.04	0.23	0.01	0.85	0.01	0.83	-0.08	0.02	0.07	0.04	0.04	0.20
p5 SubSys THW	-0.04	0.20	-0.03	0.41	0.01	0.87	0.05	0.14	-0.01	0.67	-0.03	0.35	-0.04	0.22

Table 3-30:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix VIII-b

Output	EpiUn/Weibull Shape Parameter p7		EpiUn/Weibull Shape Parameter p8		EpiUn/Weibull Shape Parameter p9		LRECM/Dev Environment		LRECM/Mfg Environment		LRECM/Mfg MRL		LRECM/Design Scope	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
p6 SubSys THW	-0.04	0.20	-0.03	0.41	0.01	0.87	0.05	0.14	-0.01	0.67	-0.03	0.35	-0.04	0.22
p7 SubSys THW	-0.04	0.20	-0.03	0.41	0.01	0.87	0.05	0.14	-0.01	0.67	-0.03	0.35	-0.04	0.22
p8 SubSys THW	-0.04	0.20	-0.03	0.41	0.01	0.87	0.05	0.14	-0.01	0.67	-0.03	0.35	-0.04	0.22
p3 Comp THW	-0.04	0.17	0.01	0.66	0.00	0.92	-0.01	0.82	0.01	0.68	0.00	0.91	0.04	0.26
Risk Coeff Rel	-0.02	0.45	-0.01	0.65	0.02	0.60	-0.06	0.07	-0.03	0.39	-0.01	0.83	-0.01	0.68
Risk Coeff Dev Dur	0.02	0.53	0.00	0.98	-0.02	0.52	0.03	0.28	-0.03	0.38	0.00	0.91	0.03	0.37
Risk Coeff Dev Cost	0.00	0.96	0.03	0.42	0.00	0.92	0.01	0.84	0.05	0.13	-0.03	0.27	-0.02	0.49
Comp / Test Cost [MMU]	0.04	0.22	0.03	0.37	-0.01	0.83	0.00	0.95	-0.04	0.17	0.00	0.88	0.07	0.02
Subsystem Test Cost [MMU]	-0.01	0.82	-0.05	0.13	0.03	0.39	0.04	0.22	0.02	0.50	0.00	0.89	-0.02	0.49
System Test Cost [MMU]	0.01	0.74	0.04	0.19	-0.02	0.44	-0.03	0.42	-0.01	0.81	0.01	0.73	-0.01	0.66
No. of Failures Sys TP 1	-0.01	0.73	0.02	0.44	0.02	0.60	-0.02	0.47	-0.01	0.79	-0.05	0.14	-0.04	0.22
No. of Failures Sys TP 2	-0.02	0.62	0.03	0.42	0.02	0.63	-0.02	0.53	0.00	0.93	-0.05	0.15	-0.04	0.26
No. of Failures Sys TP 3	-0.01	0.78	0.01	0.66	0.02	0.56	-0.03	0.35	-0.02	0.59	-0.04	0.17	-0.04	0.23

Table 3-31: Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix IX-a

Output	LRECM/ Team Experience rho	p-val	LRECM/ Producibility rho	p-val	LRECM/ Unit LC Factor rho	p-val	LRECM/ Overhead on TFU for Dev System Price rho	p-val	DT/ Average Fund Availability 2012 rho	p-val	LRECM/ Oxidizer Cost per kilogramm [MU per kg] rho	p-val	LRECM/ Fuel Cost per kilogramm [MU per kg] rho	p-val
MoE Reliability	-0.05	0.12	0.04	0.18	-0.03	0.41	0.02	0.57	-0.03	0.39	-0.01	0.66	-0.03	0.42
MoE Affordability [MMU]	0.06	0.08	0.47	0.00	0.01	0.81	0.12	0.00	0.01	0.87	-0.06	0.06	0.09	0.01
MoE IOC [y]	0.00	0.91	-0.06	0.05	0.03	0.40	0.01	0.87	0.09	0.00	-0.03	0.37	0.02	0.50
Service Life HW Tot	0.01	0.87	0.01	0.85	-0.03	0.42	0.00	0.97	0.00	0.88	-0.05	0.13	-0.03	0.30
Development Cost (D&D), [MMU]	0.25	0.00	-0.01	0.77	0.01	0.83	-0.01	0.82	-0.01	0.74	-0.01	0.65	0.03	0.33
Development Cost (HW) [MMU]	-0.03	0.40	0.53	0.00	0.00	0.98	0.12	0.00	0.00	0.97	-0.06	0.06	0.03	0.34
Development Cost (DD+HW) [MMU]	0.05	0.09	0.49	0.00	0.01	0.86	0.12	0.00	-0.01	0.70	-0.07	0.04	0.04	0.21
Development Cost (Testing) [MMU]	0.00	0.95	-0.05	0.12	-0.01	0.86	0.02	0.52	0.13	0.00	0.02	0.60	0.39	0.00
0.5-percentile (Median) System Level	-0.05	0.12	0.04	0.18	-0.03	0.37	0.02	0.56	-0.03	0.40	-0.01	0.68	-0.03	0.40
DDE&T Cost (Total Dev) [MMU]	0.06	0.08	0.47	0.00	0.01	0.81	0.12	0.00	0.01	0.87	-0.06	0.06	0.09	0.01
p-th Percentile Rel Sys	-0.05	0.12	0.04	0.18	-0.03	0.41	0.02	0.57	-0.03	0.39	-0.01	0.66	-0.03	0.42
Engine Level THW	-0.04	0.23	0.02	0.49	0.00	0.94	0.00	0.96	0.00	0.91	-0.03	0.29	0.02	0.47
TFU [MMU]	-0.02	0.56	1.00	0.00	0.02	0.54	0.05	0.14	0.01	0.77	-0.06	0.07	0.02	0.44
weighted FF	-0.03	0.39	-0.20	0.00	-0.01	0.70	-0.04	0.18	-0.05	0.10	0.04	0.18	-0.03	0.35
BogeyEQL based Reliability	0.05	0.09	0.00	0.99	-0.01	0.66	-0.01	0.68	-0.05	0.15	-0.02	0.58	0.02	0.57
AUC [MMU]	-0.01	0.69	0.97	0.00	0.23	0.00	0.03	0.31	0.02	0.56	-0.04	0.20	0.02	0.61
p1 SubSys THW	0.03	0.30	0.02	0.59	-0.01	0.64	-0.02	0.58	-0.02	0.43	-0.01	0.75	0.04	0.22
p2 SubSys THW	0.03	0.30	0.02	0.59	-0.01	0.64	-0.02	0.58	-0.02	0.43	-0.01	0.75	0.04	0.22
Production Cost Tot [MMU]	0.01	0.83	0.54	0.00	0.10	0.00	0.02	0.59	0.01	0.87	-0.06	0.05	-0.02	0.59
p3 SubSys THW	0.03	0.30	0.02	0.59	-0.01	0.64	-0.02	0.58	-0.02	0.43	-0.01	0.75	0.04	0.22
Affordability [MMU]	0.01	0.66	0.61	0.00	0.10	0.00	0.05	0.14	0.01	0.75	-0.07	0.03	0.01	0.87
p5 SubSys THW	0.03	0.30	0.02	0.59	-0.01	0.64	-0.02	0.58	-0.02	0.43	-0.01	0.75	0.04	0.22

Table 3-32: **Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix IX-b**

Output	LRECM/ Team Experience		LRECM/ Producibility		LRECM/ Unit LC Factor		LRECM/ Overhead on TFU for Dev System Price		DT/ Average Fund Availability 2012		LRECM /Oxidizer Cost per kilogram [MU per kg]		LRECM/ Fuel Cost per kilogram [MU per kg]	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val	rho	p-val
p6 SubSys THW	0.03	0.30	0.02	0.59	-0.01	0.64	-0.02	0.58	-0.02	0.43	-0.01	0.75	0.04	0.22
p7 SubSys THW	0.03	0.30	0.02	0.59	-0.01	0.64	-0.02	0.58	-0.02	0.43	-0.01	0.75	0.04	0.22
p8 SubSys THW	0.03	0.30	0.02	0.59	-0.01	0.64	-0.02	0.58	-0.02	0.43	-0.01	0.75	0.04	0.22
p3 Comp THW	-0.01	0.79	0.03	0.37	0.01	0.80	0.00	0.90	0.01	0.66	-0.02	0.46	0.02	0.55
Risk Coeff Rel	-0.03	0.40	-0.04	0.26	0.02	0.45	-0.03	0.34	0.00	0.92	-0.05	0.15	-0.04	0.24
Risk Coeff Dev Dur	0.01	0.74	-0.01	0.75	-0.01	0.81	0.02	0.46	-0.06	0.05	-0.05	0.15	0.01	0.65
Risk Coeff Dev Cost	0.01	0.85	0.01	0.81	-0.02	0.59	0.00	0.88	0.02	0.57	-0.01	0.73	0.00	1.00
Comp / Test Cost [MMU]	0.04	0.27	-0.01	0.86	-0.02	0.47	-0.01	0.80	0.08	0.01	0.00	0.95	0.00	0.95
Subsystem Test Cost [MMU]	-0.03	0.29	-0.01	0.73	0.00	0.97	0.00	0.91	0.01	0.70	0.06	0.04	0.08	0.01
System Test Cost [MMU]	0.00	0.91	-0.05	0.14	0.00	0.90	0.02	0.51	0.13	0.00	0.01	0.87	0.38	0.00
No. of Failures Sys TP 1	0.02	0.57	0.08	0.01	0.01	0.71	0.00	0.93	-0.03	0.43	-0.01	0.64	0.03	0.39
No. of Failures Sys TP 2	0.02	0.49	0.07	0.02	0.01	0.71	-0.01	0.84	-0.03	0.37	-0.01	0.74	0.03	0.35
No. of Failures Sys TP 3	0.01	0.76	0.08	0.01	0.02	0.59	0.01	0.84	-0.02	0.56	-0.02	0.56	0.03	0.32

Table 3-33:   Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix X-a

Output	TF/Consumables [MMU per test]		weight1 Rel		weight2 Aff		weight3 IOC	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val
MoE Reliability	0.02	0.54	0.01	0.77	0.01	0.81	-0.02	0.45
MoE Affordability [MMU]	0.03	0.29	-0.03	0.31	0.01	0.69	0.00	0.93
MoE IOC [y]	0.04	0.19	0.06	0.07	-0.06	0.08	0.00	0.98
Service Life HW Tot	-0.01	0.69	0.02	0.57	-0.02	0.46	0.01	0.78
Development Cost (D&D), [MMU]	0.03	0.30	-0.01	0.70	-0.05	0.12	0.05	0.15
Development Cost (HW) [MMU]	0.01	0.68	-0.02	0.45	0.02	0.51	-0.01	0.82
Development Cost (DD+HW) [MMU]	0.02	0.54	-0.04	0.25	0.01	0.74	0.01	0.76
Development Cost (Testing) [MMU]	0.13	0.00	0.04	0.21	0.00	0.97	-0.04	0.19
0.5-percentile (Median) System Level	0.01	0.65	0.01	0.84	0.01	0.78	-0.02	0.47
DDE&T Cost (Total Dev) [MMU]	0.03	0.29	-0.03	0.31	0.01	0.69	0.00	0.93
p-th Percentile Rel Sys	0.02	0.54	0.01	0.77	0.01	0.81	-0.02	0.45
Engine Level THW	0.01	0.68	-0.02	0.44	0.04	0.26	-0.02	0.47
TFU [MMU]	0.02	0.60	-0.01	0.78	-0.02	0.45	0.02	0.48
weighted FF	-0.05	0.09	0.02	0.44	-0.05	0.11	0.03	0.41
BogeyEQL based Reliability	0.00	0.92	-0.01	0.78	0.00	0.90	0.01	0.80
AUC [MMU]	0.01	0.71	-0.02	0.47	-0.02	0.60	0.03	0.36
p1 SubSys THW	0.00	0.89	-0.03	0.42	0.05	0.14	-0.02	0.51
p2 SubSys THW	0.00	0.89	-0.03	0.42	0.05	0.14	-0.02	0.51
Production Cost Tot [MMU]	-0.01	0.69	0.00	0.97	-0.03	0.36	0.03	0.35
p3 SubSys THW	0.00	0.89	-0.03	0.42	0.05	0.14	-0.02	0.51
Affordability [MMU]	0.00	0.96	-0.01	0.81	-0.03	0.36	0.03	0.31
p5 SubSys THW	0.00	0.89	-0.03	0.42	0.05	0.14	-0.02	0.51

Table 3-34: **Spearman's Rank Coefficient and associated Spearman Rank Hypothesis p-values – Matrix X-b**

Output	TF/Consumables [MMU per test]		weight1 Rel		weight2 Aff		weight3 IOC	
	rho	p-val	rho	p-val	rho	p-val	rho	p-val
p6 SubSys THW	0.00	0.89	-0.03	0.42	0.05	0.14	-0.02	0.51
p7 SubSys THW	0.00	0.89	-0.03	0.42	0.05	0.14	-0.02	0.51
p8 SubSys THW	0.00	0.89	-0.03	0.42	0.05	0.14	-0.02	0.51
p3 Comp THW	-0.02	0.61	-0.03	0.29	0.02	0.52	0.01	0.78
Risk Coeff Rel	0.01	0.67	0.00	0.90	0.01	0.71	0.01	0.87
Risk Coeff Dev Dur	0.06	0.08	0.00	0.99	-0.01	0.73	0.00	0.91
Risk Coeff Dev Cost	-0.02	0.51	0.01	0.83	-0.02	0.59	-0.01	0.65
Comp / Test Cost [MMU]	0.04	0.18	-0.02	0.45	0.01	0.82	0.03	0.38
Subsystem Test Cost [MMU]	0.05	0.15	-0.05	0.11	0.04	0.19	0.01	0.87
System Test Cost [MMU]	0.12	0.00	0.05	0.14	-0.01	0.77	-0.04	0.21
No. of Failures Sys TP 1	0.03	0.27	-0.03	0.32	0.04	0.21	-0.01	0.78
No. of Failures Sys TP 2	0.03	0.34	-0.03	0.42	0.03	0.27	-0.01	0.76
No. of Failures Sys TP 3	0.05	0.14	-0.05	0.15	0.04	0.18	0.00	0.91

3.4.6 Minimally Informative Priors versus Informative Priors

In Section 3.3.2, various minimally informative and informative priors were discussed. This section studies various prior distributions, as listed in Table 3-35, in order to substantiate the prior distributions selection for the Bayesian estimation approach that is used in the RAIV strategy. Note that the test case starts with the initial hot-fire test plan of the RS-68 liquid rocket engine as presented in [27]. Then, multiples of 2, 3, and 5 were applied to successively increase the number of hot-fire tests, i.e., 183, 366, 549, and 915, with an corresponding accumulated hot firing duration of 18,979, 37,958, 56,937, and 94,895 seconds, respectively. The simulation results are listed in Table 3-36 to Table 3-44.

Table 3-35: List of Prior Distributions of Interest

Distribution	Mix para. ϕ_{sys}	Shape para. α_{sys}	Shape para. β_{sys}
Jeffreys' prior	--.--	0.5	0.5
Beta (informative) [1]	--.--	40.0	0.5
Krolo – Type Ia [81] [2]	0.676	138.705	2.018
Krolo – Type Ib [81] [2][3]	0.676	138.705	0.224
Kleyner – Type Ia [82] [2]	0.676	138.705	2.018
Kleyner – Type Ib [82] [2][3]	0.676	138.705	0.224
Component mixture Type Ia – see Eq. (3.23) [2]	0.676	138.705	2.018
Component mixture Type Ib – see Eq. (3.23) [1]	0.676	40.0	0.5
Component mixture Type II – see Eq. (3.24) [2][3]	0.676	138.705	0.224

[1] Shape parameters are determined from information given in [80]. In [27], shape parameters of 38.8 and 0.68 were used, respectively.
[2] Shape parameters are determined from the methods of quantiles using Eqs. (3.18) to (3.20) and mix parameter from Eq. (3.25) using the testing profiles of RS-68 as given in [27].
[3] Shape parameter β is based on a competing risk assumption

Table 3-36:   Simulation Results – Jeffreys' Prior

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.8888	0.9413	0.9593	0.9749
Variance	0.0023	0.0007	0.0004	0.0001
Median	0.8951	0.9442	0.9622	0.9760
0.4 Percentile	0.8836	0.9382	0.9575	0.9734
0.1 Percentile	0.8225	0.9058	0.9335	0.9598

Table 3-37:   Simulation Results – Informative Beta Prior

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.9440	0.9609	0.9701	0.9797
Variance	0.0007	0.0003	0.0002	0.0001
Median	0.9480	0.9636	0.9722	0.9814
0.4 Percentile	0.9416	0.9589	0.9683	0.9790
0.1 Percentile	0.9100	0.9351	0.9523	0.9662

Table 3-38:   Simulation Results – "Krolo" Prior – Type I-a

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.8717	0.8970	0.9140	0.9345
Variance	0.0008	0.0005	0.0004	0.0002
Median	0.8730	0.8981	0.9156	0.9356
0.4 Percentile	0.8656	0.8913	0.9106	0.9312
0.1 Percentile	0.8324	0.8668	0.8881	0.9142

Table 3-39:   Simulation Results – "Krolo" Prior – Type I-b

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.9849	0.9876	0.9921	0.9926
Variance	0.0001	0.0001	0.0000	0.0000
Median	0.9875	0.9896	0.9933	0.9936
0.4 Percentile	0.9850	0.9872	0.9921	0.9925
0.1 Percentile	0.9707	0.9759	0.9849	0.9860

Table 3-40:   Simulation Results – "Kleyner" Prior – Type I-a

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.8989	0.9181	0.9283	0.9432
Variance	0.0006	0.0003	0.0002	0.0002
Median	0.9002	0.9202	0.9288	0.9440
0.4 Percentile	0.8956	0.9160	0.9256	0.9399
0.1 Percentile	0.8693	0.8937	0.9080	0.9266

Table 3-41:   Simulation Results – "Kleyner" Prior – Type I-b

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.9848	0.9897	0.9912	0.9933
Variance	0.0002	0.0001	0.0000	0.0000
Median	0.9887	0.9916	0.9926	0.9944
0.4 Percentile	0.9859	0.9897	0.9911	0.9934
0.1 Percentile	0.9720	0.9800	0.9833	0.9870

Table 3-42:   Simulation Results – Component Mixture Prior – Type I-a

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.9015	0.9186	0.9299	0.9476
Variance	0.0006	0.0004	0.0003	0.0002
Median	0.9037	0.9194	0.9307	0.9483
0.4 Percentile	0.8973	0.9152	0.9264	0.9454
0.1 Percentile	0.8715	0.8918	0.9082	0.9300

Table 3-43:   Simulation Results – Component Mixture Prior – Type I-b

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.9393	0.9602	0.9687	0.9790
Variance	0.0009	0.0003	0.0002	0.0001
Median	0.9437	0.9630	0.9707	0.9808
0.4 Percentile	0.9372	0.9582	0.9672	0.9781
0.1 Percentile	0.8976	0.9363	0.9500	0.9661

Table 3-44:   Simulation Results – Component Mixture Prior – Type II

Statistics	183|18,979 s	366|37,958 s	549|56,937 s	915|94,895 s
Mean	0.9835	0.9886	0.9909	0.9933
Variance	0.0003	0.0001	0.0000	0.0000
Median	0.9884	0.9910	0.9925	0.9942
0.4 Percentile	0.9855	0.9891	0.9908	0.9932
0.1 Percentile	0.9663	0.9770	0.9826	0.9872

Figure 3-34 depicts the median statistics from the simulation runs, actual mission reliability levels based on published data [117], and RAIV strategy validations using the hot-fire test plans that were executed in the SSME and F-1 development and reliability growth programs [186], and the RAIV strategy application to the RS-68 hot-fire test plan to qualify/certify the liquid rocket engine for flight [27].
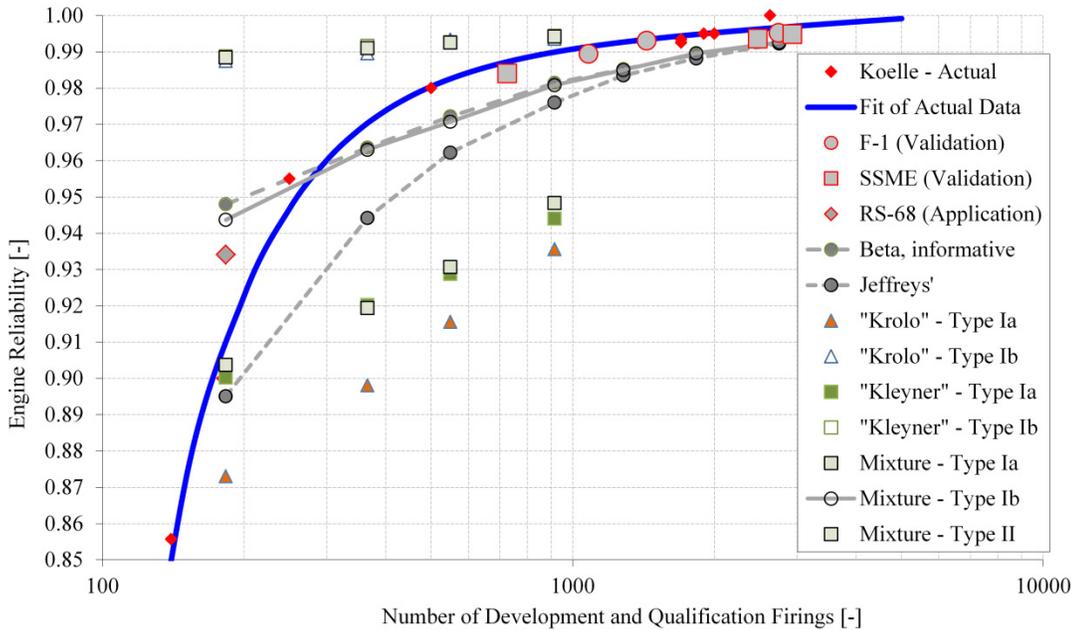
Figure 3-34: Comparison of Simulation Results to Actual Engine Mission Reliability Levels

By looking at Figure 3-34 and assessing the first four simulation runs, one could identify three groups. The first group of priors, i.e., strong informative priors (Krolo – Type Ib, Kleyner – Type Ib, and Mixture – Type II), dominate the posterior medians in terms of overestimation. Therefore, these priors are not adequate for the Bayesian estimation that is used in the RAIV strategy. The second group of priors, i.e., informative and mixture priors (Krolo – Type Ia, Kleyner – Type Ia, and Mixture – Type Ia), dominate the posterior medians in terms of underestimation. The third group of priors, i.e., Jeffreys' prior, a Beta informative, and the mixture of both (Mixture – Type Ib), allow generally the data to dominate the posterior medians. Hence, this general behavior is further investigated by additional simulation runs up to 2745 hot-fire tests. The results are listed in Table 3-45 to Table 3-47 and already displayed in Figure 3-34.

150

Table 3-45:   Simulation Results – Best Prior Candidates – Case 1

Statistics	Jeffreys' 1281|132,853 s	Beta informative 1281|132,853 s	Mixture Type Ib 1281|132,853 s
Mean	0.9820	0.9842	0.9840
Variance	0.0001	0.0001	0.0001
Median	0.9835	0.9852	0.9850
0.4 Percentile	0.9820	0.9837	0.9831
0.1 Percentile	0.9709	0.9739	0.9740

Table 3-46:   Simulation Results – Best Prior Candidates – Case 2

Statistics	Jeffreys' 1830|189,790 s	Beta informative 1830|189,790 s	Mixture Type Ib 1830|189,790 s
Mean	0.9872	0.9885	0.9885
Variance	0.0000	0.0000	0.0000
Median	0.9882	0.9893	0.9896
0.4 Percentile	0.9863	0.9880	0.9882
0.1 Percentile	0.9787	0.9816	0.9812

Table 3-47:   Simulation Results – Best Prior Candidates – Case 3

Statistics	Jeffreys' 2745|284,685 s	Beta informative 2745|284,685 s	Mixture Type Ib 2745|284,685 s
Mean	0.9916	0.9921	0.9921
Variance	0.0000	0.0000	0.0000
Median	0.9923	0.9926	0.9926
0.4 Percentile	0.9913	0.9917	0.9916
0.1 Percentile	0.9862	0.9872	0.9869

As shown Figure 3-34, the mixture prior seems to dominate the posterior median when the number of hot firings is small, but it converges to Jeffreys' prior when the number of hot firings increases. The difference between Jeffreys' prior and the informative Beta prior or the mixture prior when the number of hot firings is large is of no practical importance. However, the significant difference when the number of hot firings is small can be utilized with respect to the knowledge transfer factor ϕ_{sys} in combination with the mixture prior. In particular, if the knowledge transfer factor $\phi_{sys} = 0$ (no transfer of knowledge or additional failures in a reliability growth

tracking that were not initially planned in the TAAF cycle), then the estimated system reliability would be penalized; otherwise, if the knowledge transfer factor $\phi_{sys} = 1$, then the estimated system reliability follows the actual engine mission reliability. Therefore, the mixture prior Type Ib, i.e., a finite mixture distribution with Jeffreys' prior and an informative Beta distribution as distribution functions of the populations with the knowledge transfer factor ϕ_{sys} as mix parameters as given in Eq. (3.24) is selected in the frame of this research.

3.4.7 Impact of Failure Mechanisms Weighting Factor on the estimated System Reliability

The EQM definition [see Eq. (3.13)] requires the definition of a factor that weighs the two failure mechanisms. In that context, two limiting cases can be studied, i.e., the domination of the stress-increased failure mechanism ($\zeta_{ij} = 1$) or the domination of the strength-reduced failure mechanism ($\zeta_{ij} = 0$). The resulting median system level reliabilities for the SSME and the RS-68 are listed in Table 3-48.

Table 3-48: Impact of Failure Mechanism Weights on the System Reliability

	SSME	RS-68
Stress-increased failure mechanism only	0.9847	0.9547
Weighted failure mechanisms	0.9833	0.9544
Strength-reduced failure mechanism only	0.9814	0.9541

The differences are small but as pointed out in [124], the system level reliability is overestimated if only unsteady modes (stress-increased failure mechanisms) are considered and is underestimated if the unsteady modes are neglected. In addition, the sensitivity study (see Section 3.4.5) indicated already that the weighting factors for

the failure mechanisms are not of practical importance for the measure of effectiveness reliability and that the most influencing parameters are the decision variables and the parameters of the two-component mixture prior (see Table 3-15 through Table 3-34).

Chapter 4: Specific Problems and Discussions

This chapter presents the results of applying the RAIV strategy and the RISDM approach to four problems related to liquid rocket engine development and test planning. Section 4.1 describes the RAIV strategy and applies it to liquid rocket engine hot-fire test plans. Section 4.2 describes the application of the RISDM approach to optimize liquid rocket hot-fire tests plans. Section 4.3 describes the application of the RAIV strategy as reliability growth model. Finally, Section 4.4 describes the behavior of the genetic algorithm that is used in the RISDM approach. These results demonstrate the usefulness of the RAIV strategy and the RISDM approach.

4.1 Reliability-as-an-Independent-Variable Applied to Liquid Rocket Hot-fire Test Plans

Manufacturers lack an adequate method to balance affordability, reliability, and Initial Operational Capability (IOC). The reliability-as-an-independent-variable (RAIV) strategy is the solution proposed by expressing quantitatively the reliability trade space as ranges of a number of hardware sets and a number of hot-fire tests necessary to develop and qualify/certify a liquid rocket engine against a stated reliability requirement. Therefore, reliability-as-an-independent-variable becomes one of the key decision parameters in early tradeoff studies for liquid rocket engines because the reliability trade space directly influences the performance requirements and, as a result, the affordability and IOC. The overall solution approach of the RAIV

strategy is based on the Bayesian statistical framework using either the planned or actual number of hot-fire tests. The planned hot-fire test results may include test failures to simulate the typical design-fail-fix-test cycles present in liquid rocket engine development programs in order to provide the schedule and cost risk impacts for early tradeoff studies. The RAIV strategy is applied to the actual hot-fire test history of the F-1 liquid rocket engine, the space shuttle main engine (SSME), and the RS-68 liquid rocket engine. The results show adequate agreement between the estimated values and the actual flight engine reliability.

4.1.1 Introduction

Liquid rocket engines have always been one of the major affordability drivers of launch vehicles, but, in the past, national prestige or military requirements dominated the decisions about the development of a new launch vehicle. This paradigm has changed. Affordability, reliability, and IOC have equal importance in the decision-making process. Europe is currently facing this paradigm change by defining the requirements for an expendable next generation launcher in the frame of the ESA's Future Launchers Preparatory Program [23]. Various launch vehicle architectures were identified, ranging from a two-stage pure liquid rocket engine-based architecture to a three-stage launch vehicle with two solid propellant stages and a cryogenic upper-stage engine. Although innovative technologies are identified in all relevant areas, the focus will be on affordability in order to develop a launcher that is competitive on cost [187] but maintains the same the mission success reliability and other launch service factors as the current European launch vehicle (the Ariane 5) and other operational launch vehicles.

The affordability of expendable launch vehicles is largely determined by the development and production costs of their liquid rocket engines [1, 10]. The major part of the development cost is spent on development test hardware that is subjected to hot-fire tests in order to sufficiently demonstrate design maturity and robustness and to qualify/certify the liquid rocket engines for a successful flight operation [117]. The reliability-as-an-independent-variable (RAIV) strategy provides the framework for specifying qualification/certification hot-fire test attributes in terms of the number of tests, number of hardware sets, and total test duration that are allocated at the component, the subsystem, and the engine system level. The production cost is driven mainly by performance and reliability requirements that can be transferred into a manufacturing complexity expressed as a number of parts, precision of the parts, and selected materials. One of the main leverages on the development cost is the chosen verification strategy, with seeks to minimize the number of hardware sets by testing the mission requirements on a single hardware set multiple times but increases the production cost because of increased performance requirements, the selection of special materials, and the need for elevated manufacturing precisions in order to guarantee the longer life capability. Affordability of the launch vehicle would be incomplete without the consideration of vehicle operation and support, mission assurance, range cost, and insurance fees [188].

Therefore, finding the optimal choice in the conflicting trade spaces for performance, reliability, and affordability becomes a multiple-criterion decision-making (MCDM) problem. The trade spaces for affordability and performance are generated with parametric cost models and thermodynamic cycle codes. However, the

main shortcoming of the current MCDM solutions is the lack of an adequate modeling technique for the reliability trade space in terms of the number of hot-fire tests and number of hardware sets given a formal reliability requirement; the RAIV strategy addresses this shortcoming.

4.1.2 Background

Liquid rocket engine qualification or, synonymously, flight certification has always been a concern of space industry and agency alike because no industry or government-wide recognized standard exists. The approach by which the confidence is gained to fly includes the elements design methodology, analyses, component tests, subsystem tests, system development tests, and system qualification or certification tests. In short, the confidence-building process is dominated by an expensive and schedule-impacting hot-fire test program [16].

Historically, the hot-fire test program definitions experienced an evolution from a formal reliability demonstration to an aggressive cost minimization approach. Initially, liquid rocket engine development programs included a formal reliability demonstration requirement (e.g., F-1 or J-2) but they were discarded in favor of design verification specifications (DVSs) [e.g., space shuttle main engine (SSME)] due to prohibitively high hot-fire test costs [10]. The most recent approach is the objective-based variable test/time philosophy executed for qualifying the RS-68 liquid rocket engine that required the least amount of hot-fire tests and accumulated hot-fire test duration [118].

Although these different test program philosophies were applied for various liquid rocket engines with large performance differences, one may wonder why no

significant trend can be seen on the qualification/certification hot-fire test attributes as listed in Table 4-1 [9]. The numbers of tests required per hardware set are higher for the F-1 and J-2 compared with the SSME, which were all man rated, and the SSME is even reusable but subjected to different hot-fire test definitions, i.e., the formal reliability demonstration versus the DVS. The J-2 and RL10 are both cryogenic upper-stage liquid rocket engines, but hardware changes were allowed only for the J-2 and not for the RL10. Table 4-2 may reveal the only difference among the test attributes that is linked to the propellant combination used and the resulting internal loads present during engine operation; that is, more tests and, as a consequence, a higher accumulated test duration, which is expressed as a number of multiple mission durations, is placed on hardware sets for the propellant combination liquid oxygen (LOx)/liquid hydrogen (two- to fivefold) compared with the propellant combination LOx/kerosene or hypergolic storable propellants (more than tenfold). This identified difference may be biased by the lack of visibility on the extent of the prior component level or the development engine test history.

Table 4-1:   Qualification/certification Hot-fire Test Attributes

Test Attributes	F-1	J-2	RL10	LR87	LR91	SSME	RS-68[1]
Number of tests required	20	30	20	12	12	10	12
Total test duration required, s	2250	3750	4500	1992	2532	5000	1800
Number of samples	1	2	3	1	1	2	2
Hardware changes allowed	Yes	Yes	No	Yes	Yes	Yes	Yes
Fleetleader concept used	No	No	No	No	No	Yes	No
Overstress testing	No	No	Yes	No	No	Yes	Yes
* Values are based on the data given in [118]							

The surveys performed by Emdee [4, 5] and Pempie and Vernin [21] provide further details about the variety of current best practices by recommending the scope of hot-fire test programs and highlighting the lack of an industry or government standard or guideline. The recommendations vary from 400 hot-fire tests with 40,000 seconds

accumulated test duration spread over 15 hardware sets to 150 hot-fire tests with at least 50,000 seconds of accumulated test duration but without a statement about a required number of hardware sets.

Table 4-2:   Detailed Analysis of Qualification/Certification Hot-fire Test Attributes

Test Attributes	F-1	J-2	RL10	LR87	LR91	SSME	RS-68
Test per hardware	20	15	6.7	12	12	5	6
Test duration per hardware, s	2250	1875	1500	1992	2532	2500	900
Duration per test per hardware, s	112.5	125.0	225.0	166.0	211.0	500.0	150.0
Mission nom. time (max), s	165	500	700	165	225	520	250
Multiple of mission nom. time, s	13.6	3.8	2.1	12.1	11.3	4.8	3.6

Despite these two recommendations, Wood [118] reports that the RS-68 engine was subjected to 183 hot-fire tests with an accumulated test duration of only 18,945 seconds spread over eight new and four refurbished hardware sets before the maiden flight on the Delta IV launch vehicle. Greene [119] describes a similar hot-fire test plan for the J-2X in its nonhuman rated certification configuration requiring 182 hot-fire tests spread over six engine hardware sets. An extreme for an expendable liquid rocket engine might be the RD-0120, which was subjected to 793 tests with 163,000 seconds accumulated hot-fire duration spread over more than 90 hardware sets [189].

Although the space industry was innovative with hot-fire test program definitions ranging from a formal reliability demonstration to an objective-based variable test/time philosophy without a quantified reliability demonstration requirement at all, the U.S. Air Force Guidelines (RM2000) and the U.S. Department of Defense Total Quality Management Initiatives dictated that liquid rocket engine contractors shall elevate reliability to an equal status with performance and cost [122]. In response to these guidelines and initiatives, a Space Propulsion Integrated Reliability Team was founded in order to define a reliability demonstration technique

for the space transportation main engine (STME) [123]. The proposed strategy is based on the U.S. Army Materiel Systems Analysis Activity reliability growth model and the well-known binomial distribution in order to support a formal reliability by confidence demonstration. However, this reliability demonstration technique has not been applied to the RS-68 although it was an outgrowth of the STME study [118], most likely due to budget constraints. Consequently, the lack of an industry or government standard or guideline remains evident.

4.1.3 Reliability-as-an-Independent-Variable Strategy

The RAIV strategy is a solution to the lack of an industry or government standard by providing the ranges for the trade space in terms of the number of hardware sets and number of hot-fire tests to achieve both a stated reliability demonstration (test bogey that may correspond with the hardware reliability) and a reliability projection (mission reliability) level to assure mission success. It is based on the statistical treatment of multilevel data aggregation and bogey time testing principles applying the Bayesian framework to assure minimum hot-fire test plans. Physics-based enhancements are included in the statistical treatment of the hot-fire test data in order to reflect particularities of liquid rocket engine hot-fire test programs. The overall goal of the RAIV strategy is to generate the quantitative figures of the reliability trade space.

The inputs to the RAIV strategy include the reliability level that must be demonstrated (the reliability projection requirement), a series of function nodes to model the functional architecture of the liquid rocket engine, prior distributions of the success probabilities for each functional node at a component level reflecting the

existing experience, the duration of the different hot-fire tests, the mission duration, and expert opinions about the life capability of hardware. The outputs of the RAIV strategy are the number of hot-fire tests that should be done at the system, subsystem, and component levels and the number of hardware sets required to perform these tests.

The overall RAIV strategy is depicted in Figure 4-1. The main steps of the strategy are listed below.

1) To define the hot-fire test strategy, the functional architecture of liquid rocket engine is modeled as a series of functional nodes (if one main function fails, the system fails) not only to provide the mathematical framework to determine the success probability of each node, and finally the system-level reliability projection, but also to represent the hot-fire test strategy. The single functional nodes represent the component level, whereas the combined sets of functional nodes define subsystem- and system-level hot-fire tests.

2) To express hot-fire tests as mission equivalents, the notion of equivalent mission (EQM) is used to relate the cyclic and time-dependent failure mechanisms to the mission specification. In particular, the time-dependent failure mechanisms are accounted for by weighing tests that are shorter than the full mission duration. In this way, for each functional node, the numbers of tests and failures for the components associated with that node are used to determine the EQMs.
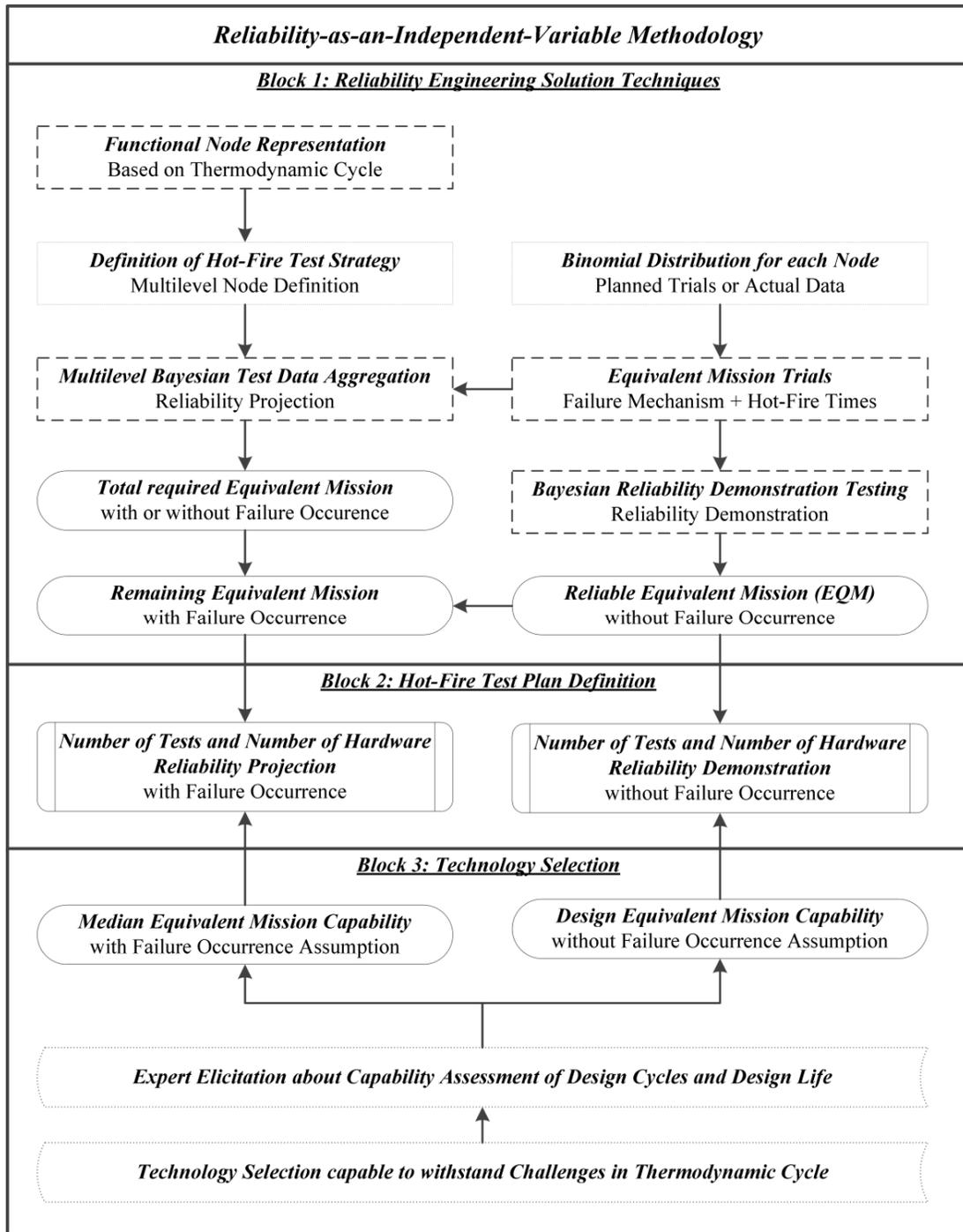
Figure 4-1:   RAIV Strategy

3) To estimate the reliability projection metric, a Markov chain Monte Carlo (MCMC) method is used to determine the posterior distributions of the success probabilities of the functional nodes at component level but uses all multilevel hot-fire test data that are obtained during development and qualification/certification testing, i.e., the results from component, subsystem, and system hot-fire tests. The functional node architecture at system level is then used to estimate the reliability projection metric using the results of the MCMC. The quantitative level of the reliability projection metric sizes the overall hot-fire test plan in terms of EQMs.

4) To estimate the reliability-by-confidence metric, the Bayesian reliability demonstration testing (BRDT) technique is used to determine the minimum equivalent design life of the hardware components that must be tested in order to demonstrate (with a given confidence) that the engine meets its hardware reliability requirement, under the assumption that there are no failures. The quantitative level of the reliability-by-confidence metric determines the hardware reliability.

5) To express hardware reliability as life capability, information about the ability of the hardware sets to survive the hot-fire tests is provided as expert opinions that are elicited to define the design number of cycles and design life. In addition, the associated failure mechanisms and failure modes are elicited based on the thermodynamic cycle of the liquid rocket engine. This information about the hardware reliability is converted into individual equivalent life (EQL) capability. The EQL uses the same basic definition as the EQM. Hence, it also relates the cyclic and time-dependent failure mechanisms to the mission specification but uses the design number of cycles and design life.

6) To determine the number of hardware sets, given the equivalent number of tests required and the EQL capability of the hardware sets, the number of hardware sets is estimated.

7) To optimize the hot-fire test plan subject to programmatic constraints and formal reliability requirements, the optimal hot-fire test plan specifies the smallest acceptable number of tests required at the component, subsystem, and system level and, as a consequence, the lowest number of required hardware sets given a certain life capability.

Functional Node Representation

The multilevel Bayesian test data aggregation (BTDA) technique requires the transfer of the physical liquid rocket engine cycle architecture into a node representation as a framework to aggregate mathematically the underlying hot-fire test strategy, i.e., the hot-fire tests either planned or performed at component, subsystem, and engine system levels [17, 19, 190]. The lowest level is defined by the structural relationship of the system components or subassemblies similar to the fault tree or reliability block diagram techniques.

However, this classical structural relationship was modified to a functional relationship because various liquid rocket engine piece parts or subassemblies are subjected to environment-based coupling factors that propagate a failure mechanism via identical internal environmental characteristics. Examples of subassemblies that have a common cause failure mode are 1) the main oxidizer valve, fuel preburner oxidizer valve, and oxidizer preburner oxidizer valve of the SSME; 2) the main oxidizer valve and oxidizer gas generator (GG) valve of the RS-68 or Vulcain 2; and
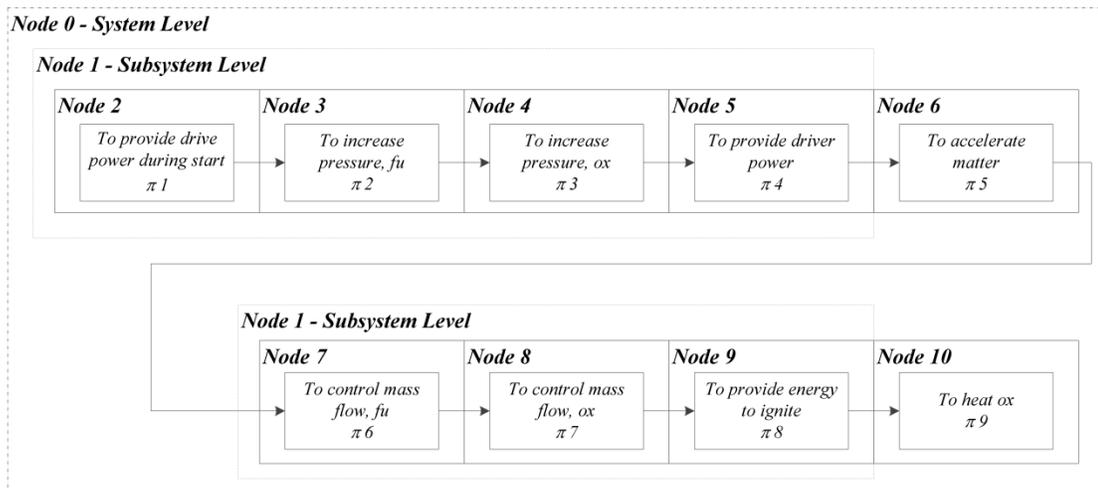
3) the coupling of boost pumps with main pumps performance. It is also important to notice that the functional node representation selects only components or subassemblies that are most pertinent to experience a failure mode during operation, i.e., turbomachinery, combustion devices, propellant valves, igniters, heat exchangers, etc. Smaller subassemblies (such as roll control, check valves, purge and pressurization lines, and electronic parts) are not included in this model because their reliability should be (nearly) 100%, which can be demonstrated with subassembly testing. If this is not true, then the RAIV strategy, which focuses on liquid rocket engine hot-fire test requirements, should be extended to incorporate the unreliable subassemblies and avoid overestimating the system reliability.

Figure 4-2 depicts the functional node representation of the hot-fire test strategy that was used for the RS-68 liquid rocket engine as described by Wood [118]. The engine system level is the node zero, the power-pack (PP) subsystem is the node one, and the components are the functional nodes two through 10. Note that "fu" refers to the fuel and "ox" refers to the oxidizer propellant route. The physical mapping to the functional nodes is given below.

Number of Trials Expressed as Equivalent Mission

The technology maturation and qualification/certification of liquid rocket engines include hot-fire tests of the rocket engine at system, subsystem, and component levels. These tests may be done at multiple durations. A hot-fire test group is a set of tests of the components associated with a functional node, where all the tests in this group have the same duration. Let subscript i denote the functional node and subscript j denote a duration group for that functional node. Associated with each

duration group is the number of cycles tested NFC_{ij}^{TP}; the number of failures r_{ij}; the test duration FD_{ij}^{TP}; the weighting factors for the two failure mechanisms, ζ_{ij} and $(1-\zeta_{ij})$; a weighting factor to account for hot-fire tests shorter than full mission duration w_{ij}; and an acceleration factor (AF) to account for different operational load points AF_{ij}^{TP}. As described in the following paragraphs, these data are used to determine EQM_{ij}^{TP}, the EQMs of these tests, and, EQM_i^{TP}, the EQMs of all of the tests for a functional node.

Node 0 - System Level

Node 1 - Subsystem Level

Node 2	Node 3	Node 4	Node 5	Node 6
To provide drive power during start $\pi\,1$	To increase pressure, fu $\pi\,2$	To increase pressure, ox $\pi\,3$	To provide driver power $\pi\,4$	To accelerate matter $\pi\,5$

Node 1 - Subsystem Level

Node 7	Node 8	Node 9	Node 10
To control mass flow, fu $\pi\,6$	To control mass flow, ox $\pi\,7$	To provide energy to ignite $\pi\,8$	To heat ox $\pi\,9$

Figure 4-2:   Functional Node Representation of the RS-68 Liquid Rocket Engine

The different hot-fire durations for the typical operation of liquid rocket engines are the consequences of the product life cycle, which include acceptance tests as well as the actual flight mission. A typical product life cycle for a liquid rocket engine includes the following hot-fire events: 1) acceptance hot-fire testing before the actual flight, 2) a possible engine ground start hold-down with launch commit criteria abort, and 3) a single flight mission duration (or several flight missions in case of a reusable

main stage engine) or multiple re-ignitions in case of upper-stage liquid rocket engines.

These hot-fire events are usually combined into a single main life cycle (MLC). Additional hot-fire tests are augmented to the product life cycle or, equivalently, MLC during the development and qualification/certification of liquid rocket engines. However, the augmentation of hot-fire tests is not infinite due to hardware degradation, and testing is stopped at the presence of a failure or even before. The test bogey is therefore the complete set of hot-fire test events that may consist of multiple MLCs and/or hot-fire events that are shorter than full mission duration. The test bogey can be chosen arbitrarily, but we suggest linking it to the reliable life capability of the hardware itself. The application of the test bogey is, however, deferred to later paragraphs of this section, because the different hot-fire events must be normalized first with respect to the mission and different hot firings as described next.

In each hot-fire operation, the hardware is degraded by the two fundamental failure mechanisms, stress-increased (cyclic) and strength-reduced (time dependent), which result in the failure mode wear, erosion, creep, and fatigue, including crack initiation and propagation, and thermal shock caused by cyclic high-temperature ranges as well as cyclic mechanical stress/strain amplitudes [71, 116].

The notion of EQM captures both the stress-increased and strength-reduced failure mechanisms caused by the cyclic startup and shutdown transients and the time-dependent material wearout during steady-state operations. The fundamental definition of the EQM is given in Eq. (4.1). The first term reflects the stress-increased

167

failure mechanism, and the second term reflects the strength-reduced failure mechanism, respectively:

$$EQM = \zeta \frac{NFC^{TP}}{NFC^{MP}} + \left(1 - \zeta\right)\frac{CFD^{TP}}{CFD^{MP}} \tag{4.1}$$

where ζ is the weighting factor for the two failure mechanisms, NFC^{TP} is the number of hot firing cycles associated to the testing profiles with the corresponding cumulative hot firing durations CFD^{TP}, and NFC^{MP} is the number of hot firing cycles associated to the mission profile with the corresponding cumulative hot firing durations CFD^{MP}.

The weighting factor ζ is assumed to be 0.5 in this study, but advanced physics-of-failure (POF) analysis models for the various subassemblies may determine more accurate values by varying the stress-increased and strength-reduced loading of the subassembly and component designs. One of these advanced POF analysis models is under final evaluation for the failure modes present in liquid rocket engine combustion chambers [116].

Startup and shutdown modes are more detrimental than the steady state operational mode of liquid rocket engines [124, 140]; therefore, some weighting inside the CFD^{TP} is used to account for these different effects. Worlund et al. [125] made available actual weighting factors for the liquid rocket engines J-2, F-1, H-1, and SSME, which were based on a failure probability model introduced by Lloyd and Lipow [124]. These data were partially used to describe the relationship as given in Eq. (4.2):

$$w_j = \frac{\ln(tp_j) - \beta_{med,0}}{\beta_{med,1}} \tag{4.2}$$

where $\ln(tp_j)$ is the natural logarithm of the hot-fire test proportion tp_j for the hot-fire group j, and $\beta_{med,0}$ as well as $\beta_{med,1}$ are the two median regression coefficients. The weighting factors may also be calculated using the Bayesian estimation for the parameters that define the likelihood function as given in Lloyd and Lipow [124].

If required, an AF for different operational load points may also be defined in order to account for accelerated life testing phenomena. However, more research is required in the field of advanced POF models for liquid rocket engine subassemblies and components in order to apply adequate rating factors in the planning stage of hot-fire test plans. The impact of the AF on the RAIV strategy can be seen in Section 4.1.4.

Introducing all extensions, the final EQM equation for a hot-fire test group j within a functional node i is given in Eq. (4.3):

$$EQM_{ij}^{TP} = \zeta_{ij} \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1 - \zeta_{ij}\right) \frac{NFC_{ij}^{TP} AF_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP}}{CFD^{MP}} \tag{4.3}$$

where ζ is the weighting factor for the two failure mechanisms, NFC_{ij}^{TP} is the number of cycles tested (one cycle consists of the startup and shutdown), NFC^{MP} is the MLC ignition quantity without overhaul of the system in between the missions, $\left(AF_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP}\right)$ is the rated and weighted test duration times the number of cycles tested NFC_{ij}^{TP}, and CFD^{MP} is the hot firings accumulated during the MLC.

The likelihood of the multilevel BTDA requires the aggregation of hot-fire test data of each functional node in terms of the equivalent number of total trials EQM_i^{TP} and equivalent number of successful trials $EQM_i^{TP\langle S \rangle}$. The number of total trials EQM_i^{TP} is given in Eq. (4.4):

$$EQM_i^{TP} = \sum_{j=1}^{J_i} EQM_{ij} \qquad (4.4)$$

where EQM_{ij} is the EQM as defined in Eq. (4.3). The number of equivalent successful trials $EQM_i^{TP\langle S \rangle}$ is given in Eq. (4.5):

$$\begin{aligned}
EQM_i^{TP\langle S \rangle} = EQM_i^{TP} \\
- \sum_{j=1}^{J_i} \left(\zeta_{ij} \cdot \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1 - \zeta_{ij}\right) \cdot \frac{NFC_{ij}^{TP} AF_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP\langle F \rangle}}{CFD^{MP}} \right)
\end{aligned} \qquad (4.5)$$

where EQM_i^{TP} is defined in Eq. (4.4), and the second term is equivalent to Eq. (4.3) but equated at the actual failure time that accounts for the different failure mechanisms, e.g., low and high cycle, wear, blanching, etc.

Equations (4.4) and (4.5) correspond to the number of trials and number of successes in an attribute sampling but normalized with the MLC. Both equations are used in the following section.

Multilevel Bayesian Test Data Aggregation Including Mathematical Solution

The multilevel BTDA serves two objectives: either to predict the reliability projection level during the hot-fire test planning process or to estimate the reliability projection level as metrics for the mission success probability during the actual hot-fire test program execution. The test data are planned or collected at various integration

levels, i.e., component, subsystem, and system using both the development and qualification hot-fire test events. The BTDA technique also provides a simulation framework to optimally allocate the hot-fire tests given a required reliability projection level subject to schedule and budget constraints.

The full Bayesian formulation of the multilevel BTDA technique is given as unscaled posterior in Eq. (4.6). The solution of Eq. (4.6) is, however, nontrivial because the mathematical relationship at the lowest level functional node decomposition is a function of the subsystems and system probabilities, i.e., $\pi_i = f\left(\pi_{C_x}\right)$:

$$\pi\left(\underline{\theta}\,|\,Data\right) \propto \prod_{i=1}^{I} \pi_i^{EQM_i^{TP\langle S\rangle}} \left(1-\pi_i\right)^{EQM_i^{TP}-EQM_i^{TP\langle S\rangle}} \prod_{i=1}^{I} \pi_i^0\left(\underline{\theta}\right) \qquad (4.6)$$

where $\pi\left(\underline{\theta}\,|\,Data\right)$ is the posterior of the parameter vector $\underline{\theta}$ given the Data, π_i is the individual lowest level functional node success probability, $\pi_i^0\left(\underline{\theta}\right)$ is the prior distribution of the individual lowest level functional node success probability, and *Data* is the multilevel data in terms of EQM EQM_i^{TP} as defined in Eq. (4.4) and equivalent successes $EQM_i^{TP\langle S\rangle}$ as defined in Eq. (4.5) of each functional node at component, subsystem, and system levels.

The difficulty of the multilevel BTDA implementation is linked to the numerical integration over the complete domain Θ, even with modern general-purpose multidimensional integration algorithms [191]. Instead, the MCMC method was used to generate samples from the unscaled target density using a one-variable-at-a-time Metropolis–Hastings (MH) algorithm [97, 100, 107, 108, 192-194]. The

171

algorithm cycles through all unknown parameters, one at a time, proposing new samples from an independent candidate density while holding the remaining parameters at their most recent values, i.e., at arbitrary initial values. The logit scale is used for the update of the samples from the candidate probability density function $q(\theta^*)$ as given in Eq. (4.7):

$$\theta_i^* = \frac{1}{e^{-F_X^{-1}(u)} + 1} \tag{4.7}$$

where $F_X^{-1}(u)$ is the equated inverse cumulative density function of $X \sim N\left(\text{logit}\,\theta_i^{(m)}, \sigma_i\right)$ at the random number u generated by $U \sim U(0,1)$. The standard deviation σ_i of the distribution function is a tunable constant that influences the one-variable-at-a-time acceptance rate of the acceptance probability α_i for new candidate values for each functional node probability π_i. The acceptance probability is given in Eq. (4.8):

$$\alpha\left(\theta_i^{(m)}, \theta_i^* \mid \underline{\theta}_{-i}\right) = \min\left\{1, \frac{\pi\left(\theta_i^* \mid Data\right)}{\pi\left(\theta_i^{(m)} \mid Data\right)} \cdot \frac{\theta_i^*\left(1 - \theta_i^*\right)}{\theta_i^{(m)}\left(1 - \theta_i^{(m)}\right)}\right\} \tag{4.8}$$

where $\pi\left(\theta_i^* \mid Data\right)$ is the unscaled target density (posterior) that is evaluated with the new candidate value θ_i^*, $\pi\left(\theta_i^{(m)} \mid Data\right)$ is the unscaled target density (posterior) that is evaluated at the previously accepted value $\theta_i^{(m)}$.

MCMC samples are not independent random samples; therefore, the burn-in time and the sample autocorrelation of the samples are a concern. The burn-in time is the number of steps in the MCMC needed to draw the samples from the long-run

distribution. Unfortunately, no mathematical treatment is given that determines the length of the burn-in period. As a remedy, the autocorrelation function is used to determine the sample autocorrelations and the lag by which the samples of the Markov chain must be thinned at in order to use independent draws. The standard deviation σ_i of the independent candidate distributions influence the sample autocorrelations and the acceptance rates of each Markov chain; therefore, the burn-in time is used to tune the standard deviations σ_i in such a way that the acceptance rates of each individual parameter are close to 0.35 [19, 107, 108].

Finally, the results of the MH MCMC for the individual functional node parameters π_i are used to calculate the subsystems and system success probability or reliability projection metrics such as the mean, the variance, or any other pth percentile.

The selection of the prior distributions for the functional node parameters π_i is crucial because only a small number of liquid rocket engine hot-fire test programs is available, providing only indirect information about the parameters to be estimated. In such a problem setting, the prior distribution becomes more important and sensitivity analyses should check the adequacy of the choice of prior distribution parameters. Several sets of prior distribution shape parameters were tested including the noninformative parameter settings $\alpha = 1$ and $\beta = 1$. The best set for the informative prior shape parameters were $\alpha = 38.3$ and $\beta = 0.7$ for the two sets of hot-fire test programs of the SSME and the RS-68, respectively. The sensitivity study for the selection of prior distribution parameter settings was also used to validate the

MH MCMC code by running the code several times with different initial values for the parameters to be estimated.

Bayesian Reliability Demonstration Testing

The main advantage of the BRDT technique is the reduction of test sample size [115]. The governing BRDT equation is derived using the Bayesian estimation of the failure fraction. The derivation starts with the classical Binomial distribution but modified with the EQM notion as given in Eq. (4.9):

$$L(Data \mid q) = \begin{pmatrix} EQM_{RbyC}^{TP} \\ r \end{pmatrix} q^r (1-q)^{EQM^{TP} - r}$$ (4.9)

where q is the failure fraction, EQM_{RbyC}^{MP} is the number of mission profile EQMs associated with the R-by-C requirement, and r is the number of observed failures during the hot-fire test plan. Note that the number of failures r is usually assumed to be zero in the Bayesian success testing under an exponential distribution assumption. Here, the number of failures is, however, kept in the remaining derivation because it can be used in sensitivity studies for test planning purposes using a planned number of failures or to account for actual failure cases if erroneous assumptions about the hardware reliability were initially made in the hot-fire test planning process.

The prior distribution in the classical Bayesian setup of attribute life test data is based on the Beta distribution as defined in Eq. (4.10):

$$f(q; \alpha, \beta) = \frac{q^{\alpha-1} (1-q)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}$$ (4.10)

where q is the failure fraction (the parameter to be estimated in the Bayes theorem),

α and β are the shape parameters of the Beta distribution, and $\mathrm{B}(\alpha, \beta)$ is the

solution of the Eulerian integral of the first kind:

$$\int_0^1 x^p (1-x)^q \, dx$$

An empirical Bayes approach was used to estimate the parameter settings for

the shape parameters α and β using the data given in McFadden and Shen [80]. The

procedures described by Martz and Waller [70, 171] or by Modarres et al. [71] were

applied that lead to the same parameter estimates, but the latter one is mathematically

more appealing and is given in Eq. (4.11):

$$p_{pr} = \frac{x_0}{n_0} \quad \text{and} \quad n_0 = \frac{1 - p_{pr}}{k^2 p_{pr}} - 1 \quad \text{where} \quad n_0 > x_0 \geq 0$$

$$\alpha = x_0 \quad \text{and} \quad \beta = n_0 - x_0$$

(4.11)

where p_{pr} is the prior mean, x_0 are the successes, n_0 are the trials, k is the

coefficient of variation, and α and β are the shape parameters of the Beta

distribution. The estimated shape parameters α and β that correspond to the mean

as well as the 0.05 and 0.95 percentiles of liquid rocket engine reliability are listed in

Table 4-3.

Table 4-3:   Shape Parameter for the Beta Prior Distribution in the BRDT
**   Plan**

	0.05 percentile	Mean	0.95 percentile
Shape parameter α	21	39	42
Shape parameter β	0.6	0.5	1.2

The posterior distribution percentiles of the failure fraction q are related to the

binomial distribution as given in Eq. (4.12):

$$\Pr\left(q \le q_u\right) = \int_0^{q_u} \pi\left(q; \alpha, \beta \mid Data\right) dq = C \tag{4.12}$$

where q_U is the upper percentile of the posterior distribution, $\pi\left(q; \alpha, \beta \mid Data\right)$ is the posterior distribution of the failure fraction q, and C is the level of confidence (credibility bound).

The analytical solution of the posterior distribution percentiles of the failure fraction q is given in Eq. (4.13):

$$\frac{\mathrm{B}_{q_U}\left(\alpha + r, \beta + EQM_{RbyC}^{TP} - r\right) \Gamma\left(\alpha + \beta + EQM_{RbyC}^{TP}\right)}{\Gamma\left(\beta + EQM_{RbyC}^{TP} - r\right) \Gamma\left(\alpha + r\right)} = C \tag{4.13}$$

where EQM_{RbyC}^{TP} is the EQM without occurrence of failures to meet the reliability-by-confidence (R-by-C) requirement, r is the number of equivalent failures set to zero in the BRDT, α and β are the Beta distribution shape parameters, C is the credibility bound, $\mathrm{B}_{q_U}\left(\cdot\right)$ is the incomplete beta function, and

$$\Gamma\left(z\right) = \int_0^\infty t^{z-1} e^{-t} dt$$

is the Gamma function.

The equivalency of Eq. (4.13) with the well-known frequentist binomial model $\left(1 - C\right) = R^n$ may not be obvious, but if Eq. (4.13) is rewritten using a vague prior (parameters α and β are set to 1), and assuming a zero failure success testing, the Bayesian-like binomial model can be stated as given in Eq. (4.14):

$$Bi\left(0, q, n+1\right) = \left(1-q\right)^{n+1} = \left(1-C\right) \Leftrightarrow R^{n+1} = \left(1-C\right) \tag{4.14}$$

where $Bi(0,q,n+1)$ is the binomial probability density function including the Bayesian adjustment of the vague prior by the quantity $n+1$ instead of only n in the frequentist framework and C as the confidence level.

Life Capability of Hardware Sets

The RAIV strategy uses the notion of EQM to capture the two stress-increased and strength-reduced failure mechanisms into a single metrics. The resulting failure modes are the result of accumulated damages during the various hot-fire runs as response to the internal thermofluid-mechanical challenges. The proper physical design of the parts and subassemblies of liquid rocket engines must withstand these challenges, which are expressed as design cycles and design life. Typical values are listed in Table 4-4, but one of the main deficits of the reported values is the lack of an associated reliability statement [9].

Table 4-4:   Engine Design and Mission Requirements

Engine Name	Design Cycles	Design Life, s	Missions	Mission Starts	Mission Nominal Time, s	MP Cycles	MP FD	EQL_{RbyC}^{MP}
SSME	55	27,000	55	1	520	4	680	26.7
F-1	20	2,250	1	1	165	3	215	8.3
J-2	30	3,750	1	1	380	3	480	8.8
				2[a]	150[a]	4	600	6.8
					350[a]		450	
RL10	20	4,500	1	2	700	4	890	5.1
LR87	12	1,980	1	1	165	3	215	6.4
LR91	12	2,700	1	1	225	3	295	6.7

[a] First hot-fire and restart

These two design metrics are transferred into the single metric EQL with an associated reliability level similar to the notion of EQM in order to use it in the frame of the RAIV strategy. It is important to note that the bogey EQL (EQL_{RbyC}^{MP}) is a

metric that is based on the assumption that no failure occurred up to the equivalent bogey number of cycles and bogey life that may also correspond with the design number of the cycles and design life. The promoted approach for future liquid rocket engines would be the equality of the bogey test requirements with the design number of cycles and design life. The computed EQL_{RbyC}^{MP} is also listed in Table 4-4, assuming the given MLC in terms of the number of cycles and accumulated HFTD. To transfer the bogey number of cycles and bogey life into a single EQL notion, the following two assumptions were made: (1) The stress-increased failure mechanism is modeled by a Poisson distribution and (2) The strength-reduced failure mechanism follows a Weibull distribution.

The Poisson distribution is a proper choice for cyclic loads since it describes a random discrete variable with no upper bound. The Weibull distribution governs the time to occurrence of the weakest link of many competing failure processes. Typical piece parts or subassemblies of liquid rocket engines that dominate the time to failure or cycles to failure occurrence are the turbine(s), bearings, or combustion chamber liner.

The life capability definition requires the two reliability measures in terms of bogey number of cycles and bogey life as well as the median number of cycles and median life. The bogey reliability measure is the number of cycles or time for which the reliability will be R (hot-fire testing without failure occurrence), whereas the median reliability measure corresponds to the 0.5 percentile of the underlying failure distribution (hot-fire testing is performed until a failure occurred).

The life capability uses the same functional structure as the EQM already introduced in Eq. (4.1) but with relevant modifications linked to the bogey number of cycles, the bogey life, and the 0.5 percentiles. The bogey EQL (EQL_{RbyC}^{MP}) is given in Eq. (4.15), and the median EQL ($\widetilde{EQL}^{TP}$) is given in Eq. (4.16):

$$EQL_{RbyC}^{MP} = \xi \frac{c_{RbyC}^{MP}}{NFC^{MP}} + (1-\xi)\frac{t_{RbyC}^{MP}}{CFD^{MP}} \tag{4.15}$$

where ξ is the weighting factor of the capacity to withstand the challenges that trigger the two failure mechanisms, c_{RbyC}^{MP} is the number of reliable cycles, t_{RbyC}^{MP} is the reliable time, and NFC^{MP} is the number of hot firing cycles associated to the mission profile with the corresponding cumulative hot firing durations CFD^{MP}.

$$\widetilde{EQL}^{TP} = \xi \frac{\tilde{\lambda}^{TP}}{NFC^{MP}} + (1-\xi)\frac{\tilde{t}^{TP}}{CFD^{MP}} \tag{4.16}$$

where ξ is the weighting factor of the capacity to withstand the challenges that trigger the two failure mechanisms, $\tilde{\lambda}^{TP}$ is the median number of cycles to failure, $\tilde{t}^{TP}$ is the median life, and NFC^{MP} is the number of hot firing cycles associated to the mission profile with the corresponding cumulative hot firing durations CFD^{MP}. The median number of cycles to failure $\tilde{\lambda}^{TP}$ is given by Eq. (4.17), and the median life $\tilde{t}^{TP}$ is given by Eq. (4.18):

$$\Pr\left(NFC^{TP} \leq c_{RbyC}^{MP}\right) = R\left(c_{RbyC}^{MP}\right) = 1 - \frac{\Gamma\left(1 + \lfloor c_{RbyC}^{MP} \rfloor, \bar{\lambda}^{TP}\right)}{\Gamma\left(1 + \lfloor c_{RbyC}^{MP} \rfloor\right)} \tag{4.17}$$

where $\Pr\left(NFC^{TP} \le c_{RbyC}^{MP} \right)$ is the probability of failure associated with the test bogey, $R\left(c_{RbyC}^{MP} \right)$ is the reliable cycles, $\overline{\lambda}^{TP}$ is the mean of the Poisson distribution, $\lfloor \cdot \rfloor$ is the floor function, and

$$\Gamma(z) = \int_{0}^{\infty} t^{z-1} \cdot e^{-t} dt$$

is the Gamma function. The search parameter is the mean of the Poisson distribution until the probability statement is true.

The inconsistency of using the mean instead of the median for the number of cycles does not impact the overall methodology because the Poisson distribution can be approximated with the normal distribution if the mean is above nine, for which the mean and the median will be indistinguishable.

$$\tilde{t}^{TP} = t_{RbyC}^{MP} \left[\frac{\ln(2)}{-\ln\left(R_{RbyC}^{MP} \right)} \right]^{\frac{1}{\beta}} \tag{4.18}$$

where t_{RbyC}^{MP} is the reliable life, R_{RbyC}^{MP} is the reliability associated with the R-by-C requirement, and β is the shape parameter of the Weibull distribution. The median time to failure was preferred over the classical mean time to failure because the median is more representative in terms of central tendency for highly skewed failure distribution, i.e., Weibull distributions with shape parameters less than three, as is the case for most of the weakest link piece parts or subassemblies present in liquid rocket engines. It should be noted that the weakest link assumption may also be used to estimate ranges for the individual hardware set requirements for each piece part or subassembly in order to adequately plan for hardware manufacturing during the design maturity demonstration and subsequent qualification/certification.

The life capability is usually derived by the mission requirements and is based on first engineering judgments, simplified engineering life time models, or on advanced POF models (recalling Table 4-4 for the used levels in the past). It is, however, important to use credible and realistic bogey capabilities in order to estimate the real hardware needs (see Section 4.1.4 for the initial SSME design cycle and design life assumptions).

Number of Hardware Sets

The number of hardware sets needed to complete the RAIV strategy hot-fire test scope is calculated using the hardware reliability necessary to support the total required EQM (EQM^{TP}) based on the multilevel BTDA technique as given in Eq. (4.19):

$$EQM^{TP} = EQM_{RbyC}^{MP} + EQM_{rem}^{TP} \Leftrightarrow HW^{TP} = HW_{RbyC}^{MP} + HW_{rem}^{TP} \qquad (4.19)$$

where EQM_{RbyC}^{MP} is the required EQM to support the BRDT, and EQM_{rem}^{TP} is the remaining EQM needed to complete the overall RAIV strategy hot-fire test scope defined by Eq. (4.6).

Equation (4.19) can be modified with the corresponding life capability in order to define the required number of hardware sets as given in Eq. (4.20):

$$HW^{TP} = \frac{EQM_{RbyC}^{MP}}{EQL_{RbyC}^{MP}} + \frac{EQM_{rem}^{TP}}{\widetilde{EQL}^{TP}} \qquad (4.20)$$

where EQM_{RbyC}^{MP} is the required EQM for the BRDT with corresponding bogey EQL (EQL_{RbyC}^{MP}), and EQM_{rem}^{TP} is the remaining EQM to complete the overall RAIV

strategy hot-fire test scope with corresponding median EQL ($\widetilde{EQL}^{TP}$) based on the 0.5 percentile.

Integrated Multiple Criteria Decision-Making Model

In general, hot-fire test planning is a MCDM problem. The criteria are the number of hardware sets, the number of hot-fire tests including the associated firing durations, the development duration, and the development cost. The RAIV strategy seeks to minimize the number of hot-fire tests subject to constraints on the development duration and cost. One of the possible solution strategies for the MCDM problem is the application of multiobjective optimization using evolutionary algorithms. Among the various evolutionary algorithms, the most popular type is the genetic algorithm, which searches the decision variable space by generating random populations of n strings using the operations of reproduction, crossover, and mutation. The distinction between feasible and infeasible solutions is determined by the penalty function approach that penalizes a soft or hard constraint violation [153, 195].

In Section 4.1.4, the impact on key hot-fire test plan metrics was analyzed for the RS-68 test case by varying the reliability projection targets.

4.1.4 Numerical Examples

The application and demonstration of the RAIV strategy with artificial hot-fire test data would lack credibility in the space industry. Therefore, the numerical examples used for the validation of the RAIV strategy are based on the hot-fire test histories of the F-1 liquid rocket engine, the SSME, and the RS-68 liquid rocket engine. They reflect the three different test program philosophies of formal reliability

demonstration, DVS, and the objective-based variable test/time, respectively. The numerical examples follow the main seven steps as introduced in Section 4.1.3.

Define Hot-Fire Test Strategy

The RAIV strategy is started with the definition of the functional node representation. The test histories of the F-1, SSME, and RS-68 liquid rocket engines were used to deduce the hot-fire test strategy. The F-1 hot-fire test history deduction is based on the data given in an immediate release by Rocketdyne [121], which stated that the number of hot-fire tests was 1081, and 278 tests were for 150 seconds or longer. No information is given on the accumulated HFTDs. The SSME hot-fire test history featured 726 hot-fire tests with 110,253 seconds of accumulated HFTD [120]. The RS-68 was qualified with 183 hot-fire tests and 18,945 seconds [118]. Based on these data, the hot-fire test strategies were deduced for the F-1, SSME, and RS-68, and they were expressed as functional nodes with the associated physical components as given in Table 4-5 for the F-1, as given in Table 4-6 for the SSME, and as given in Table 4-7 for the RS-68.

Table 4-5:   Functional Nodes of the F-1 Mapped to Physical Components

Functional node	Physical component
To provide ignition power π_1	Ignition system components
To increase pressure π_2	Single shaft turbopump arrangement (including gear)
To provide drive power π_3	Gas Generator (GG)
To accelerate matter π_4	Thrust Chamber Assembly (TCA)
To control mass flow, fuel side π_5	Valves on fuel-side
To control mass flow, oxidizer side π_6	Valves on oxidizer side

Table 4-6:   Functional Nodes of the SSME Mapped to Physical Components

Functional node	Physical component
To increase pressure, fuel side π_1	Boost and turbopump, fuel side
To increase pressure, oxidizer side π_2	Boost and turbopump, oxidizer side
To provide drive power, fuel side π_3	Preburner to drive turbine, fuel side
To provide drive power, oxidizer side π_4	Preburner to drive turbine, oxidizer side
To accelerate matter π_5	Thrust Chamber Assembly (flight nozzle extension)
To control mass flow, fuel side π_6	Main fuel valve
To control mass flow, oxidizer side π_7	Main oxidizer valve, preburner oxidizer valves
To provide energy to ignite π_8	Igniters for preburners and thrust chamber assembly
To heat oxidizer π_9	Heat exchanger to pressurize tank

Table 4-7:   Functional Nodes of the RS-68 Mapped to Physical Components

Functional node	Physical component
To provide drive power during start, π_1	Starter
To increase pressure, fuel side π_2	Turbopump, fuel side
To increase pressure, oxidizer side π_3	Turbopump, oxygen side
To provide drive power, π_4	GG to drive the fuel and oxygen pumps
To accelerate matter π_5	TCA
To control mass flow, fuel side π_6	GG and TCA Valves, fuel
To control mass flow, oxidizer side π_7	GG and TCA Valves, ox
To provide energy to ignite π_8	Igniters for GG and TCA
To heat oxidizer π_9	Heat exchanger to pressurize tank

Express Hot-Fire Tests as Mission Equivalents

The functional nodes define the hot-fire testing levels, such as component, subsystem, and system levels. The SSME test history provided more details about the system-level hot-fire tests in terms of hot-fire testing groups with different HFTDs using the data given by Biggs [120]. The F-1 and RS-68 data lack this kind of information, but the data were derived as follows. The F-1 hot-fire testing groups, with the corresponding EQMs, are based on the matching of the weighting factor for hot-fire

184

tests that are shorter than full mission duration that were given in Worland et al. [125] and a Bayesian solution for the parameters of the likelihood function of the model introduced by Lloyd and Lipow [124]. The resulting accumulated hot-fire test time is about 111,000 seconds, with the average hot firing of around 100 seconds that can be compared with the data given in Emdee [5], which result in the average hot firing of roughly 90 seconds. Likewise, the RS-68 hot-fire testing groups are based on a test allocation that resulted in the accumulated hot firing that is given by Wood [118]. The weighting factors for the hot-fire tests that were shorter than full mission duration were also calculated with the Bayesian solution for the parameters of the likelihood function of the model introduced by Lloyd and Lipow [124].

In addition, the objective-based variable test/time philosophy applied to the RS-68 includes the principles of accelerated life testing that require the application of an AF.

The derived hot-fire test strategies for the F-1, the SSME that includes an integrated subsystem test bed (ISTB) testing, and the RS-68 that includes GG component-level and PP subsystem-level testing are given in Table 4-8, Table 4-9, and Table 4-10 (RS-68 with AF of one) as well as in Table 4-11 (RS-68 with AF of five), respectively. The assumption that the AF equals five is given to investigate the impact on the number of hardware sets and the resulting reliability projection level. Using Eqs. (4.3), (4.4), and (4.5), the EQMs and the number of successful trials were determined as required inputs for (4.6).

Table 4-8:   Multilevel BTDA Scope: F-1

	r_{ij}	EQM_{ij}	NFC_{ij}^{TP}	FD_{ij}^{TP}, s	w_{ij}	ζ	$\zeta-1$
Node 0 – System							
Group 1 [15 s]	1	5.4	30	450	0.44	0.50	0.50
Group 2 [50 s]	1	11.3	50	2500	0.66	0.50	0.50
Group 3 [80 s]	1	90.3	323	25840	0.78	0.50	0.50
Group 4 [100 s]	1	127.6	400	40000	0.84	0.50	0.50
Group 5 [150 s]	2	107.4	250	37500	0.96	0.50	0.50
Group 6 [165 s]	0	13.1	28	4620	1.00	0.50	0.50

Test Scope Aggregation at System Integration Levels

	EQM_i^S	EQM_i	NFC_i^{TP}	$\sum FD^{TP}$, s
Node 0 – Engine	353.5	355.1	1081	110910

Table 4-9:   Multilevel BTDA Scope: SSME

	r_{ij}	EQM_{ij}	NFC_{ij}^{TP}	FD_{ij}^{TP}, s	w_{ij}	ζ	$\zeta-1$
Node 1 – ISTB							
Group 1 [100 s]	0	197.1	1000	100000	0.75	0.50	0.50
Node 0 – System							
Group 1 [2 s]	0	3.4	27	54	0.15	0.50	0.50
Group 2 [21 s]	0	14.1	107	2247	0.51	0.50	0.50
Group 3 [97 s]	3	31.1	184	17848	0.74	0.50	0.50
Group 4 [158 s]	4	26.9	132	20856	0.82	0.50	0.50
Group 5 [183 s]	4	26.5	121	22143	0.84	0.50	0.50
Group 6 [283 s]	3	36.0	128	36224	0.91	0.50	0.50
Group 7 [400 s]	0	7.5	21	8400	0.96	0.50	0.50
Group 8 [520 s]	0	2.6	6	3120	1.00	0.50	0.50

Test Scope Aggregation at System Integration Levels

	EQM_i^S	EQM_i	NFC_i^{TP}	$\sum FD^{TP}$, s
Node 1 – ISTB	197.1	197.1	1000	100000
Node 0 – Engine	147.5	148.1	726	110892

Table 4-10: Multilevel BTDA Scope (AF = 1): RS-68

	r_{ij}	EQM_{ij}	NFC_{ij}^{TP}	FD_{ij}^{TP}, s	w_{ij}	ζ	$\zeta-1$
Node 5 – GG							
Group 1 [50 s]	2	12.5	62	3100	0.04	0.50	0.50
Node 1 – PP							
Group 1 [100 s]	1	1.8	6	600	0.42	0.50	0.50
Node 0 – System							
Group 1 [28 s]	3	13.1	78	2195	0.04	0.50	0.50
Group 2 [136 s]	3	4.4	18	2450	0.53	0.50	0.50
Group 3 [139 s]	3	6.9	28	3900	0.53	0.50	0.50
Group 4 [163 s]	3	6.5	24	3900	0.59	0.50	0.50
Group 5 [173 s]	3	4.2	15	2600	0.59	0.50	0.50
Group 6 [195 s]	3	5.8	20	3900	0.59	0.50	0.50

Test Scope Aggregation at System Integration Levels

	EQM_i^S	EQM_i	NFC_i^{TP}	$\sum FD^{TP}$, s
Node 5 – GG	12.2	12.5	62	3100
Node 1 – PP	1.8	1.8	6	600
Node 0 – Engine	37.0	40.9	183	18945

Table 4-11: Multilevel BTDA Scope (AF = 5): RS-68

	r_{ij}	EQM_{ij}	NFC_{ij}^{TP}	FD_{ij}^{TP}, s	w_{ij}	ζ	$\zeta-1$
Node 5 – GG							
Group 1 [50 s]	2	16.7	62	3100	0.04	0.50	0.50
Node 1 – PP							
Group 1 [100 s]	1	1.9	6	600	0.42	0.50	0.50
Node 0 – System							
Group 1 [28 s]	3	20.3	78	2195	0.04	0.50	0.50
Group 2 [136 s]	3	5.7	18	2450	0.53	0.50	0.50
Group 3 [139 s]	3	8.9	28	3900	0.53	0.50	0.50
Group 4 [163 s]	3	7.9	24	3900	0.59	0.50	0.50
Group 5 [173 s]	3	5.0	15	2600	0.59	0.50	0.50
Group 6 [195 s]	3	7.3	20	3900	0.59	0.50	0.50

Test Scope Aggregation at System Integration Levels

	EQM_i^S	EQM_i	NFC_i^{TP}	$\sum FD^{TP}$, s
Node 5 – GG	16.2	16.7	62	3100
Node 1 – PP	0.3	1.9	6	600
Node 0 – Engine	53.5	55.1	183	18945

Estimate the Reliability Projection Metric

The EQMs and the number of successful trials given in Table 4-8, Table 4-9 and Table 4-10 are used in the multilevel BTDA using Eq. (4.6) to estimate the system-level reliability projection. The resulting reliability projection levels for the F-1, SSME, and RS-68 are listed in Table 4-12.

Table 4-12:   Reliability Projection Levels using the RAIV Strategy

Engine Designation	Mean	Lower Bound 0.05 percentile	Upper Bound 0.95 percentile
F-1	0.9894	0.9826	0.9964
SSME	0.9825	0.9730	0.9922
RS-68 (AF = 1)	0.9227	0.8866	0.9644
RS-68 (AF = 5)	0.9454	0.9162	0.9734

The average reliability projection levels for the F-1 and SSME of 0.9894 and 0.9825 may be compared with the formal reliability demonstration level of 0.99 at 50% [10] and the reported reliability level of 0.984 [117], respectively. In addition, Koelle [117] reported the conductance of 1437 hot-fire tests with a reliability level of 0.993 that may be compared with the RAIV-based projected reliability level of 0.9919 (Note that the average HFTD of around 100 seconds was assumed as well). No reliability has been reported for the RS-68, but the RAIV-based reliability projection levels, ranging from 0.9227 (using no AF) to 0.9454 (assuming an AF of five for all engine-level hot-fire tests), may be compared with levels of 0.92 (one flight anomaly) to 0.96 (zero flight anomaly), which were calculated with a first-level Bayesian estimate of the mean predicted reliability using the number of RS-68 liquid rocket engines used on the medium and heavy Delta IV launch vehicle until 2011 [196].

Table 4-12 should not suggest the conclusion that the RS-68 liquid rocket engine is an unreliable propulsion system. The risk of observing a launch failure

might be higher for the RS-68 since not all failure modes may have been discovered during the low number of hot-fire tests performed during the development. An intensive production quality inspection program and post-maiden-flight hot-fire testing will reduce the risks of a flight failure and increase the reliability projection, but at the expense of higher production cost than most likely initially foreseen. Flight hardware is usually subjected to a myriad of inspections and several acceptance tests at various system integration levels.

Estimate the Reliability-by-Confidence Metric

The R-by-C metric is used as input for the hardware reliability requirements, which influences the number of hardware sets required for the overall hot-fire test strategy. Equation (4.13) is used to determine the EQMs without the occurrence of a failure using the reliability projection level, which was calculated in the previous step. The confidence level is usually set to classical values of 50, 60 or 90%. In this study, the confidence level was set to 50% for the F-1 engine [10], to 60% for the SSME, and to 90% was used for the RS-68 engine [123].

Express Hardware Reliability as Life Capability

The hardware life capability is expressed by means of bogey or design cycles and bogey or design life. POF models, covariate models, or expert opinions can be used to provide credible figures. Table 4-13 lists the bogey or design cycles and bogey or design life for the F-1 [5] and SSME [9]. The bogey cycles and bogey life for the RS-68 were defined through the RAIV strategy. Only realistic hardware reliability levels should be stated during the requirement development process, as will be seen for the SSME in the next step of the RAIV strategy.

Determine the Number of Hardware Sets

Based on the R-by-C metric, where the reliability level is equal to the level of the reliability projection metric as given in Table 4-12, the confidence levels of 50, 60, and 90%, the life capability, and the number of hardware sets are determined with Eq. (4.20). The results in terms of average, minimum, and maximum numbers of hardware sets are given in Table 4-13 assuming Weibull shape parameters of 3 ± 0.5 and 4 ± 0.5 for the median lifetime estimation. The estimation of the EQM_{RbyC}^{MP} was performed with an informative prior for the SSME engine because of the ISTB, whereas the estimation for the F-1 and RS-68 engines used noninformed priors because both engines were state-of-the-art in terms of thrust size.

Table 4-13:   Total Number of Hardware Sets

Engine Designation	R-by-C	HW	Design Life	Design Cycles	$\beta=3\pm0.5$	$\beta=4\pm0.5$
F-1 (1081 hot-fire tests)	0.9839 at 0.5	56[1]	2250	20	$20 \le 22 \le 24$	$24 \le 26 \le 28$
F-1 (2740 hot-fire tests)	0.9952 at 0.5	56[1]	2250	20	$39 \le 46 \le 51$	$51 \le 55 \le 59$
SSME, specified life capability[2]	0.9825 at 0.6	20	27000[3]	55[3]	$4 \le 5 \le 6$	$6 \le 6 \le 7$
SSME, realistic test bogey[2]	0.9825 at 0.6	20	5000[4]	20[4]	$17 \le 21 \le 23$	$23 \le 26 \le 27$
RS-68	0.9454 at 0.9	8 + 4[5]	4000[6]	15[6]	$9 \le 10 \le 10$	$10 \le 10 \le 10$

[1] reported in Meisl (1986) and Emdee (2001) but spread over 2,740 hot-fire tests
[2] same hot-fire test plan assumed
[3] original design life and cycles requirement
[4] realistic life time and cycle numbers derived from Williams (1993)
[5] 4 engine hardware sets were refurbished [Wood (2002)]
[6] estimated test bogey life and test bogey cycles based on the RAIV strategy

The findings shown in Table 4-13 may suggest the use of Weibull shape parameters of 3 ± 0.5 for LOx/liquid hydrogen and 4 ± 0.5 for LOx/kerosene liquid rocket engines based on the estimated number of hardware sets using the RAIV strategy and

the corresponding reported values. Certainly, further investigations are needed to make final conclusions. The more important aspect of the results listed in Table 4-13 is, however, linked to the problem of unrealistic test bogey capability assumptions, as was the case for the SSME. Based on the initial or specified life capability requirements (55 cycles and 27,000 seconds), only five to six hardware sets would have been required for the complete development program using the RAIV strategy for the hardware estimation. However, the actual number of hardware sets was as high as 20 [8]. A similar level of hardware sets can be estimated with the RAIV strategy using the more realistic test bogey capability of 20 cycles and 5000 seconds. This set of test bogey capability, for the weakest components, is in fact more realistic using the figures reported in the generic deviation approval request limits [30]. Therefore, the SSME example demonstrates that any unrealistic test bogey capability assumption, when used in tradeoff studies, may result in infeasible hot-fire test plan definitions and may cause strong program cost overruns and schedule slippage.

Test Plan Optimization

The scenario investigated in this study assesses how changes to a stated reliability projection target value affect the key hot-fire test plan metrics: the number of hardware sets, the number of hot-fire tests, the development duration, and the development cost (no overall affordability optimization is addressed in this scenario). Therefore, the setup of the MCDM is subject only to programmatic hard constraints; that is, the development cost and the development time should not exceed twice their baseline values. The budget metric is determined with the cost tool NASA/Air Force Costing Model (NAFCOM®) in combination with a specific hot-fire test cost model

191

using the results of the RAIV strategy. The duration (schedule) metric is defined by a typical resource allocation for the design and development (DD) phase using the DD cost estimate of NAFCOM®. It is further assumed that 2.5 years of engine-level testing is accomplished within the resource allocation defined schedule with a yearly cadence of 30 tests on two test facilities. A schedule penalty function is defined to account for an elongated or expedited schedule due to the different hot-fire test numbers as a result of the different reliability projection targets.

Six values of the reliability projection target (from 0.92 to 0.96) were considered. For each value, the RAIV strategy determined the optimal hot-fire test plan. The results (presented in Figure 4-3) highlight quantitatively the expected tendencies of the claims presented in Section 4.1.2. Short development times and associated low development costs can be achieved only with limited hot-fire testing and at the expense of the confidence-building process. The limited number of hot-fire tests also impacts the number of hardware sets needed and, as a consequence, the development cost.

The recommendations for test plans ranging from 150 to 400 hot firings must be seen in conjunction with the reliability projection level that must be demonstrated before the first launch. The RS-68 test case results in a 50% increase for both the number of hot-fire tests and number of hardware sets, a 25% increase in development duration, and a 35% increase in normalized development cost if the reliability projection level is raised from the initial level of 0.933 to 0.95 (Note that the reliability level indicated is based on the case where no AF is used in the RAIV strategy).
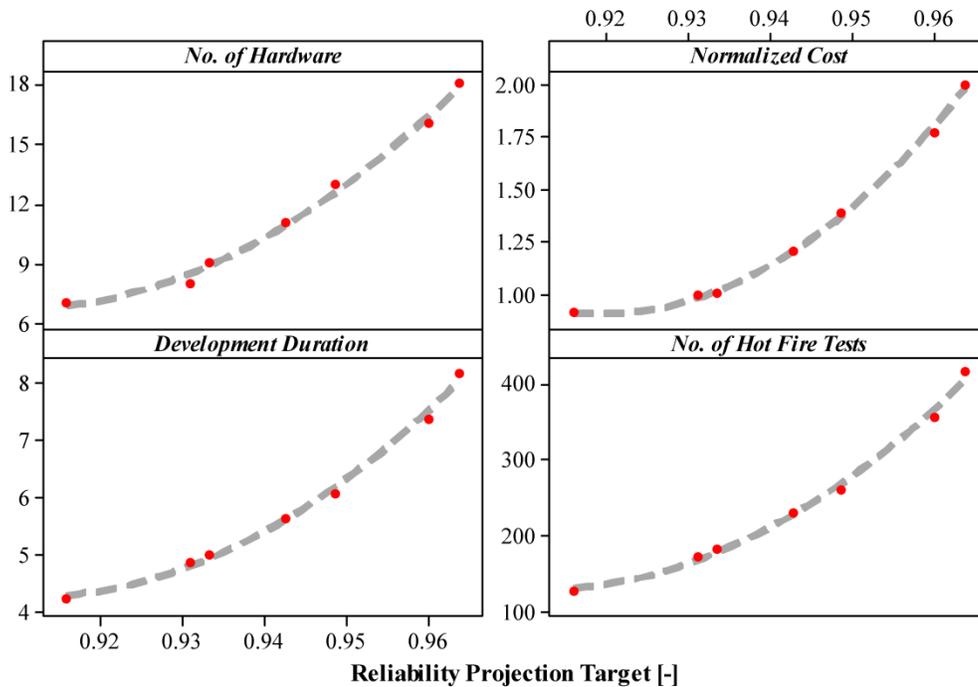
Figure 4-3: **Key Test Plan Metrics for various Reliability Projection Targets: RS-68 Test Case**

4.1.5 Conclusion

The presented RAIV strategy features unique characteristics currently not publicly available to the liquid rocket engine space industry for early tradeoff studies by providing quantitative reliability trade spaces for the number of hardware sets and the number of hot-fire tests needed to assure mission success and to demonstrate design maturity using multilevel planned hot-fire test data. In addition, the RAIV strategy can be used to define test bogeys that are associated with a reliability requirement that may also be used as a design requirement. One additional strength of the RAIV strategy is the inclusion of envisaged failures in the planning process of hot-fire test plans in order to simulate the typical design-fail-fix-test cycles present in liquid rocket engine developments. Therefore, program managers and systems engineers are

equipped with an adequate simulation framework to credibly balance performance, reliability, and affordability by combining the RAIV strategy with thermodynamic cycle models and parametric cost models. Although the RAIV strategy was demonstrated using the liquid rocket engine hot-fire test histories of the F-1, the SSME, and the RS-68 that were based on the different hot-fire test strategies formal reliability demonstration, DVS, and objective-based variable test/time, the overall acceptance of the approach depends on a future application of the methodology to a new liquid rocket engine program.

4.2 A Reliability as an Independent Variable Methodology for Optimizing Test Planning for Liquid Rocket Engines

The hot-fire test strategy for liquid rocket engines has always been a concern of space industry and agency alike because no recognized standard exists. Previous hot-fire test plans focused on the verification of performance requirements but did not explicitly include reliability as a dimensioning variable. The stakeholders are, however, concerned about a hot-fire test strategy that balances affordability, reliability, and Initial Operational Capability (IOC). A multiple criteria test planning model is presented that provides a framework to optimize the hot-fire test strategy with respect to stakeholder concerns. The Staged Combustion Rocket Engine Demonstrator, a program of the European Space Agency, is used as an example to support the claim that a reduced thrust scale demonstrator is cost beneficial for a subsequent flight engine development. Scalability aspects of major subsystems are considered in the prior information definition inside the Bayesian framework. The

model is also applied to assess the impact of an increase of the demonstrated reliability level on the development duration (IOC) and affordability.

4.2.1 Liquid Rocket Engine Test Planning

The selection of a hot-fire test plan for liquid rocket engines is a concern for the space industry and the European Space Agency because there exists no recognized standard that defines quantitatively the scope of hot-fire test plans. The current best practice is a blend of art and science that tries to define test plans that will verify performance requirements and demonstrate safety margins against known failure modes. The scope of initial test plans is defined by meeting the stated IOC and the available budget. Updates of test plans are made during the development to adjust the schedule constraints and the remaining budget. The predicted mission success probability is then a result of the executed hot-fire test plan. However, the key stakeholders – the space agency, the member states, and launch operators – are concerned about the predicted reliability, the time required for the development including the hot-fire testing to meet the IOC, and the cost of the development including hot-fire testing (''affordability'') in the early program planning stage. The scope definition of a test plan is one of the key drivers for the stakeholder concerns; therefore, the selection of an optimized hot-fire test plan becomes a multiple criteria decision-making (MCDM) problem in which the numbers of planned hot-fire tests at various system integration levels are the decision variables.

The multiple criteria test planning problem (MCTPP) is formulated as an optimization problem with elements from utility theory and normative target-based decision making. The number of hot-fire tests determines the reliability, defines the

195

development duration, and drives the affordability. The MCTPP formulation seeks to maximize a linear combination of the utilities of these values. We will solve this problem using an evolutionary algorithm that searches for the optimal hot-fire test plan.

The MCTPP is demonstrated in the context of ESA's Future Launcher Preparatory Programme (FLPP) [197]. Hot-fire test plans are found for two scenarios: (1) a reduced thrust scale engine demonstrator precedes the flight engine development and (2) a flight engine development is executed from scratch (without a demonstrator).

4.2.2 Problem Formulation

The decision variables of the MCTPP are the number of hot-fire tests. The objective function is a multiattribute utility function that relates the decision variables to the stakeholder's areas of concern: reliability, schedule, and affordability, which are all functions of the number of hot-fire tests.

Decision Variables

The decision variables of the MCTPP are the number of planned hot-fire tests allocated at the different system integration levels, i.e. component, subsystem, and system level. For example, for the LE-7A liquid rocket engine, there are nine types of tests that must be considered (see Table 4-14). The key component tests are the preburner test and the igniter test. The key subsystem tests are the fuel turbomachinery test, the oxidizer turbomachinery test, and the combustion test. There

196

are also four different system tests (which have different durations). The specific number of tests of each type must be determined, so there are nine decision variables.

For each test type, the specific number of tests is bounded below by the minimum number required to verify the performance requirements, optimize the start-up and shut down sequences, demonstrate margin against known failure modes, and attain an adequate level of demonstrated reliability to assure mission success subject to schedule and budget constraints. In addition, the specific number of tests is bounded above such that the number of required hot-fire tests for engine reliability certification is placed on engine system level (Nota Bene: The bounds provided in Table 4-14 are given only as example). Therefore, various hot-fire test strategies can be defined to demonstrate these basic test objectives. However, test facility capabilities and physical hardware degradation phenomena impose constraints to the allocation of the hot-fire tests that defines the hot-fire test strategy.

Table 4-14:   Hot Fire Test Strategy for LE-7A

System Integration Level	Min. no. of tests	Max. no. of tests	Hot-fire test time (s)
Component			
Preburner	20	60	10
Igniter	20	80	2
Subsystem			
Fuel turbomachinery	40	100	60
Ox turbomachinery	40	100	60
Combustion devices	40	100	10
System			
Test duration 1	5	50	3
Test duration 2	5	50	30
Test duration 3	5	200	150
Test duration 4	5	200	300

The component and subsystem test facilities lack the capability of providing adequate testing boundary conditions that allow the operation of the tested hardware at full

rated conditions. At system level, the full rated conditions are achieved but the test facility may lack the capability of providing the required amount of propellants to support the operation of the full mission duration. Both the limitations are superimposed by the fact that start-ups and shut downs are more detrimental than the simple accumulation of hot-fire test time. Therefore, a framework is needed to account not only for the various test facility limitations but also for the hardware degradation phenomena.

The proposed framework uses a functional node representation of the physical architecture of a liquid rocket engine and the notion of mission equivalents. The details about these two elements of the framework are described using the LE-7A architecture.

Functional Node Representation

The functional node representation of a physical architecture of a liquid rocket engine not only describes the structural relation of components known from the fault tree (FT) or reliability block diagram (RBD) techniques but also defines the fundamental hot-fire test strategy [27].

The LE-7A liquid rocket engine architecture (see [198]) is used to explain a possible fundamental hot-fire test strategy. The main components of the LE-7A, which are most likely pertinent to main failure modes, are the turbomachinery on fuel and oxidizer side, the preburner, the thrust chamber assembly, the two ignition systems, the control valve on the fuel side (MFV), the control valves on the oxidizer side (MOV and POV), the mixture ratio setting device, and the heat exchanger.

Based on the definition of components with pertinent failure modes, the functional node representation can be defined (see Figure 4-4). All of the main functions are in series (if one function fails the system fails). This node representation is node 0 and is used to aggregate all engine level hot-fire tests. It should be noticed that not all subassemblies or components of the liquid rocket engine are included in the functional node representation because the reliability levels of the ''missing'' components are considered to be unity or almost unity and therefore do not affect the reliability analysis. In case a specific subassembly or component is failure mode susceptible, it can be easily included in the node representation.
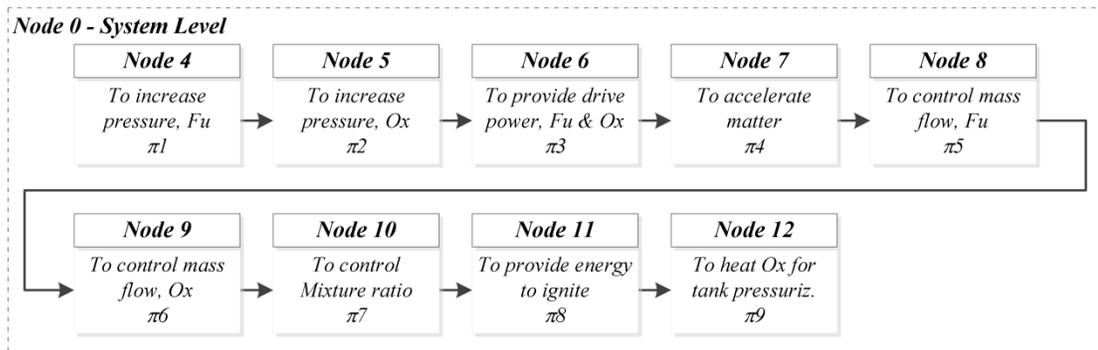
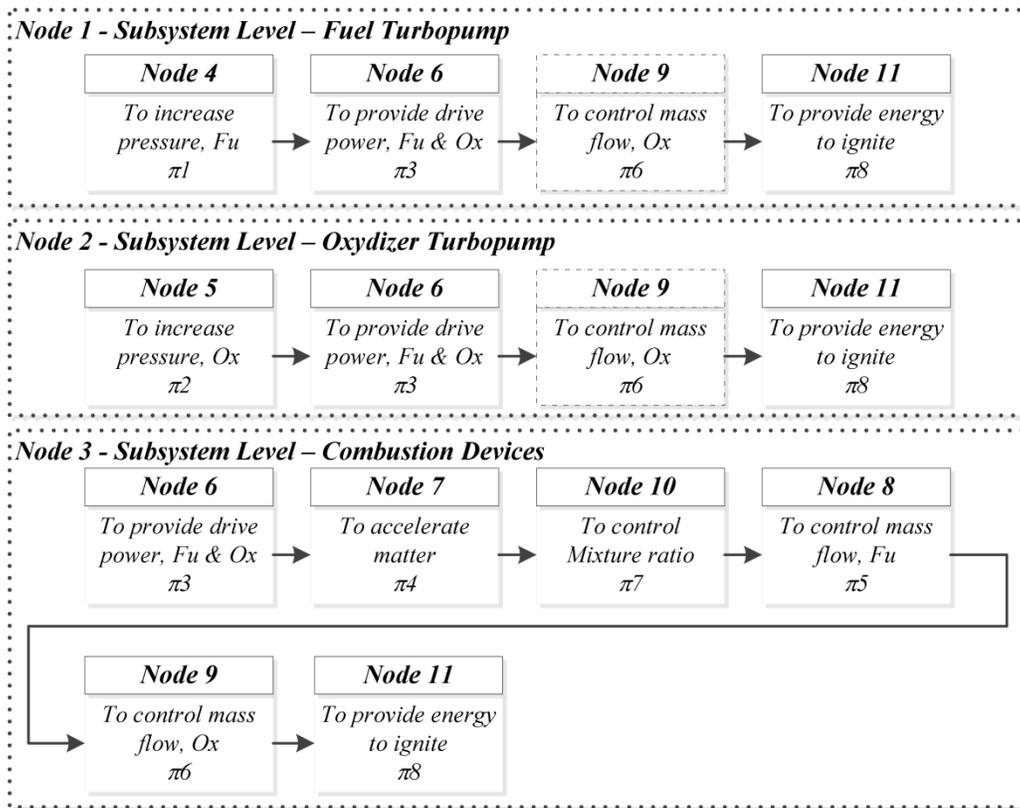**Figure 4-4:   Node 0: Engine Level – Functional Node Representation**

Once the engine level functional node representation is defined, the fundamental hot-fire test strategy at lower system level can be established. Fundamental in that sense means that subsystem level hot-fire test configurations at combustion device and turbomachinery level can be defined as shown in Figure 4-5.

Mission Equivalents

Based on the fundamental hot-fire test strategy definition, through the functional node representation, the mission equivalents are needed to relate the planned hot-fire tests at the various system integration levels to the mission requirement as well as to

capture the two fundamental stress-increased and strength-reduced failure mechanisms into a single metric, the equivalent mission (see Section 3.1).

The mission requirement not only includes the actual flight but also any other hot-fire tests aggregated throughout the product life cycle. The notion of main life cycle (MLC) is used to normalize the hot-fire test events which may consist of a single or multiple acceptance hot-fire test(s) before the actual flight, a possible engine ground start hold-down with launch commit criteria abort, and the single flight mission (or several flight missions in case of a reusable main stage engine) or multiple reignitions in case of upper stage liquid rocket engines.

Figure 4-5: **Nodes 1 and 2: Subsystem level – Functional Node Representation**

During the design maturation and qualification, additional hot-fire tests are added to the MLC. Such tests may include multiples of a nominal MLC and those that are a

fraction of full mission duration. Each hot-fire test contributes to the degradation of the hardware due to the stress-increased and strength-reduced failure mechanisms that are present in every liquid rocket engine piece part or subassembly. Equation (4.21) captures mathematically the two fundamental failure mechanisms and normalizes them with the hot-fire events of the MLC; hence, the notion of equivalent mission (EQM).

$$EQM_{ij}^{TP} = \zeta_{ij} \cdot \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1 - \zeta_{ij}\right) \cdot \frac{NFC_{ij}^{TP} AF_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP}}{CFD^{MP}} \tag{4.21}$$

The first term accounts for the stress-increased failure mechanism, and the second term accounts for the strength-reduced failure mechanism. The second term includes also the weighing of planned hot-fire tests which are shorter than full mission duration.

Therefore, the hot-fire tests can be performed with different hot-fire test durations which is reflected in the index j. The various system integration levels are defined through the index i, a group of hot-fire tests. The number of hot-fire tests in each hot-fire test group is defined by J_i. The total number of equivalent missions in each hot-fire test group i is given in Eq. (4.22).

$$EQM_i^{TP} = \sum_{j=1}^{J_i} EQM_{ij}^{TP} \tag{4.22}$$

Equation (4.23) accounts for planned hot-fire test failures in each hot-fire test group i to reflect the typical design-fail-fix-test cycles present in liquid rocket engine developments. The second term of Eq. (4.23) is based on Eq. (4.21) but is measured at the failure time.

$$EQM_i^{TP\langle S \rangle} = EQM_i^{TP}$$

$$-\sum_{j=1}^{J_i}\left(\zeta_{ij} \cdot \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1-\zeta_{ij}\right) \cdot \frac{NFC_{ij}^{TP} \, AF_{ij}^{TP} \, w_{ij}^{TP} \, FD_{ij}^{TP\langle F \rangle}}{CFD^{MP}}\right) \quad (4.23)$$

Equations (4.22) and (4.23) are used in Section 4.2.2, which describes a methodology to estimate the projected mission success probability based on the number of planned hot-fire tests that are allocated at the various system integration levels.

Measures of Effectiveness for the Areas of Concern

The measure of effectiveness for each area of concern is a function of the number of hot-fire tests. The measure of effectiveness for reliability is determined by means of the reliability as independent variable (RAIV) strategy, the measure of effectiveness for the schedule is effort level driven in terms of work force and test plan scope, and the measure of effectiveness for the budget is based on cost models that partially depend on the test plan scope, respectively. These measures of effectiveness are later used to compute the score value of the utility functions that are implemented in the MCTPP formulation.

Reliability

The RAIV methodology estimates the projected mission success probability based on the number of hot-fire tests planned. As the number of hot-fire tests increases, the reliability measure of effectiveness and, as a consequence, the reliability utility score increases. The unique features of RAIV are the multi-level aggregation of hot-fire test results (planned or actual), i.e. results may be obtained at component, subsystem,

and/or system level using the functional node representation and the pooling of test results with various hot-fire test durations using the notion of mission equivalents.

The fundamental mathematical expression of RAIV is given in Eq. (4.24). It is based on the Bayesian formulation to estimate parameters (probability of success π_i) given a set of data (the number of hot-fire tests).

$$\pi\left(\underline{\theta} \mid Data\right) \propto \prod_{i=1}^{I} \pi_i^{EQM_i^{TP\langle S \rangle}} \left(1-\pi_i\right)^{EQM_i^{TP}-EQM_i^{TP\langle S \rangle}} \prod_{i=1}^{I} \pi_i^0\left(\underline{\theta}\right) \qquad (4.24)$$

The first product expresses the hot-fire test strategy defined by the equivalent number of planned hot-fire tests EQM_i^{TP} including possible test failures $EQM_i^{TP\langle F \rangle}$ at the various functional node levels. The second product defines the prior knowledge of the parameters to be estimated in the Bayesian framework. Each individual function node may feature a different level of prior knowledge due to scalability constraints, e.g. a turbomachinery is limited in terms of scalability from a small to a much larger thrust scale if compared to a thrust chamber.

The solutions for the functional node reliability levels are used to calculate the mean, the variance or any other pth percentile of the projected engine level mission success probability.

Development Duration (IOC)

The measure of effectiveness for the IOC is effort driven as well as by the time which is needed to perform the hot-fire tests to attain the reliability target estimated with the RAIV strategy.

The effort which is needed to design and develop the hardware is estimated with the NASA/Air Force Cost Model (NAFCOM®). The test occupation simply

depends on the number of hot-fire tests allocated to the various integration levels and the number of test facilities, the test cadence per week, a yearly maintenance period, and the mounting and dismounting periods.

Based on empirical evidences given in Koelle [117], a quantile regression equation for the development period in years is defined which relates the cost for the design and development divided by the work force yearly cost [the first part of Eq. (4.25)]. The second part of Eq. (4.25) is simply the addition of the overall test duration also given in years which is determined by the test occupation model described next.

$$DP = \left[-6.62 + 1.35\ln\left(\frac{DD}{WY}\right) \right] + TO \qquad (4.25)$$

The second term of Eq. (4.25) links the measure of effectiveness for the development duration with the decision variable number of hot-fire tests. The test operational assumptions such as the number of test facilities, the test cadence per week, a yearly maintenance period in weeks, and the mounting and dismounting periods in weeks define the minimum test occupation. Eq. (4.26) defines the simple test occupation model used.

$$TO = \frac{HFT_{tot}}{TR\left(52 - M - MD\right)TF_{tot}} \qquad (4.26)$$

It should also be noted that Eq. (4.25) is not considering any schedule penalty term due the lack of a proper funding profile. It is assumed that an adequate funding profile exists.
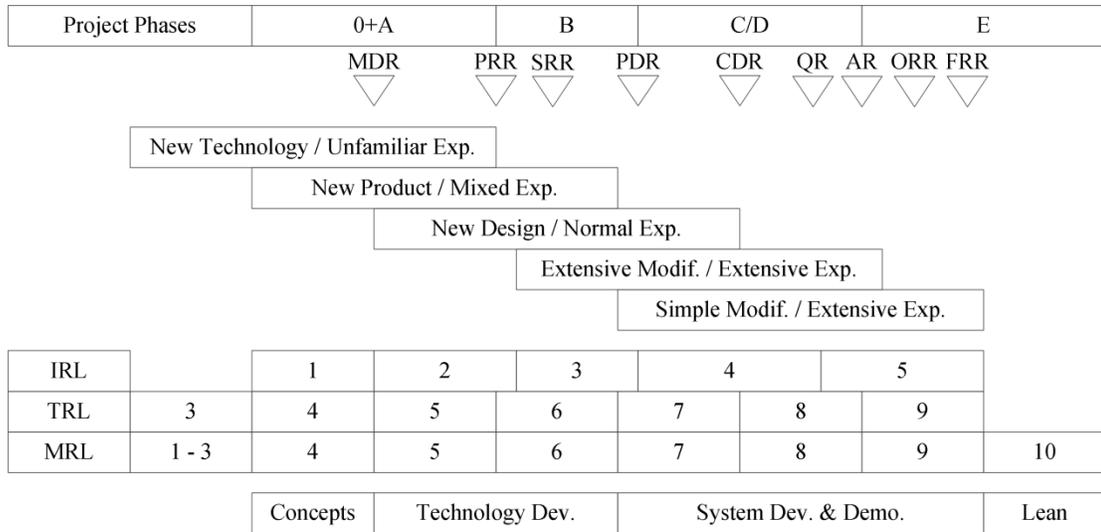
Budget (Affordability)

The measure of effectiveness for the area of concern affordability is based on two cost models: NAFCOM® and the effort-driven test facility operation cost model defined herein. The purchasing power parity principle is used to transfer the U.S. to the European productivity level in order to obtain an adequate European level for the price estimations obtained from NAFCOM® [13].

Available European engine development programmatic evidences were used to anchor/validate the two cost models for a European multi-national environment.

Design and Development and System Test Hardware Cost Model

The NAFCOM® tool is used to estimate the design and development (D&D) cost as well as the System Test Hardware (STH) cost. The D&D cost includes all the specifications and requirements, engineering drawings as well as program management and configuration control efforts that are required to achieve the built-to baseline for the definition of the STH. It includes also design rework which may become necessary after the hot-fire test conductance and evaluation.

The NAFCOM® effort-driven input variables for the D&D cost estimate are the development environment, the manufacturing environment, the Manufacturing Readiness Level (MRL), the design scope, and the design team experience. However, a correlation exists between the design effort and the team's experience, as pointed out by Sherman [199], i.e. a high design effort is also linked with a low team experience level and vice versa. In addition, the links between this correlation, the Technology Readiness Level (TRL), the MRL, the Integration Readiness Level (IRL), and project phases exist and are highlighted in Figure 4-6 [112, 163-165].

Project Phases	0+A	B	C/D	E
	MDR	PRR SRR	PDR CDR QR AR ORR	FRR

New Technology / Unfamiliar Exp.

New Product / Mixed Exp.

New Design / Normal Exp.

Extensive Modif. / Extensive Exp.

Simple Modif. / Extensive Exp.

IRL		1	2	3	4	5	
TRL	3	4	5	6	7	8	9	
MRL	1 - 3	4	5	6	7	8	9	10

Concepts	Technology Dev.	System Dev. & Demo.	Lean

Figure 4-6:   Effort-driven Cost Model Input Variables in Relation to IRL, TRL, MRL, and IRL

The STH cost is estimated based on the theoretical first unit (TFU) cost but includes a 25% overhead applied to reflect a prototype design approach. No learning curve effect is considered for the STH cost estimation. The total number of STH sets needed to complete the overall hot-fire test plan is given in Eq. (4.27) and is based on elements defined by the RAIV strategy [27].

$$HW^{TP} = \frac{EQM_{RbyC}^{MP}}{EQL_{RbyC}^{MP}} + \frac{EQM_{rem}^{TP}}{\widetilde{EQL}^{TP}} \tag{4.27}$$

Equation (4.27) uses the results obtained from Eq. (4.24) in terms of total number of equivalent missions required to attain the specified reliability level and relates it to the life capability of the piece parts or subassemblies of the liquid rocket engine components. The first term relates the number of equivalent missions without the occurrence of failures to the hardware reliability (reliable number of cycles and reliable life time). The second term completes the overall test plan by testing the remaining number of equivalent missions needed to attain the specified level of

206

reliability and relates this number to the medians of the underlying hardware reliability distribution functions describing the two fundamental failure mechanisms, i.e. the Poisson and Weibull distributions. Equation (4.28), a Bayesian formulation to estimate the percentile of a binomial distribution, is used to estimate the mission equivalents needed in the reliability by confidence (R by C) success-testing scheme, i.e. the EQM_{RbyC}^{MP}.

$$\frac{B_{q_U}\left(\alpha+r,\beta+EQM_{RbyC}^{TP}-r\right)\Gamma\left(\alpha+\beta+EQM_{RbyC}^{TP}\right)}{\Gamma\left(\beta+EQM_{RbyC}^{TP}-r\right)\Gamma\left(\alpha+r\right)}=C \qquad (4.28)$$

The percentile or failure fraction q_U is equal to the estimated reliability level. The confidence level C is specified by the customer; typically 60 or 90%. The parameters α and β reflect the prior knowledge about the engine reliability levels either based on the data given in McFadden and Shen [80] or user specific information.

The hardware reliability is defined by specifying the reliable number of cycles c_{RbyC}^{MP} and reliable life time t_{RbyC}^{MP} but is transferred into the EQM notion using Eq. (4.29). The parameter ξ is used to weigh the two failure mechanisms.

$$EQL_{RbyC}^{MP}=\xi\frac{c_{RbyC}^{MP}}{NFC^{MP}}+\left(1-\xi\right)\frac{t_{RbyC}^{MP}}{CFD^{MP}} \qquad (4.29)$$

The remaining hot-fire tests, in terms of equivalent missions needed in Eq. (4.27), are calculated with Eq. (4.30).

$$EQM_{rem}^{TP}=EQM^{TP}-EQM_{RbyC}^{MP} \qquad (4.30)$$

Similarly to Eq. (4.29), Eq. (4.31) is used to transfer the medians of the Poisson and Weibull distribution into the EQM notion which is also needed in Eq. (4.27).

$$\widetilde{EQL}^{TP} = \xi \frac{\tilde{\lambda}^{TP}}{NFC^{MP}} + (1-\xi)\frac{\tilde{t}^{TP}}{CFD^{MP}}$$

(4.31)

Equations (4.32) and (4.33) are used to calculate the medians of the Poisson and Weibull distribution, which are required in Eq. (4.31), based on the assumed reliable number of cycles c_{RbyC}^{MP} and reliable life time t_{RbyC}^{MP} of the piece parts or subassemblies.

$$\Pr\left(NFC^{TP} \le c_{RbyC}^{MP}\right) = R\left(c_{RbyC}^{MP}\right) = 1 - \frac{\Gamma\left(1+\left\lfloor c_{RbyC}^{MP}\right\rfloor, \overline{\lambda}^{TP}\right)}{\Gamma\left(1+\left\lfloor c_{RbyC}^{MP}\right\rfloor\right)}$$

(4.32)

$$\tilde{t}^{TP} = t_{RbyC}^{MP}\left[\frac{\ln(2)}{-\ln\left(R_{RbyC}^{MP}\right)}\right]^{\frac{1}{\beta}}$$

(4.33)

Test Operational Cost Model

The test operational cost model is also effort-driven, i.e. the test occupation is determined based on assumptions concerning engine mounting, test rate, and test facility operation using empirical data. The test operational cost model estimates the cost based on the values of the decision variables (the number of hot-fire tests). Once the test occupation in years is determined using Eq. (4.26), the yearly cost for a work force year is used to estimate the cost associated to the test conductance. Although minor in magnitude, the propellant cost is also considered which may become more significant in the future if the current tendency of the price increase remains evident for the hydrogen propellant.

Utility Functions and Normative Decision-Making

The utility function and normative decision making are used to define the objective function as well as to divide the search space in terms of the decision variable number of hot-fire tests into feasible and infeasible regions.

Utility Function

For each area of concern, the measure of effectiveness of a test plan is converted into a utility score. The stakeholder has target values for each measure of effectiveness, which could be used to define a simple step utility function in which any measure of effectiveness that meets the target receives a value of one, and any measure of effectiveness that does not receives a value of zero. However, this type of step function makes optimization difficult because it penalizes all poor performance solutions equally and does not reflect adequately the customer value in case a solution is above the target but is still acceptable with a lower value. Thus, we sought a utility function that would be equivalent in some sense. This will be discussed more in the next subsection.

For reliability, we used the monotonically increasing utility function given in Eq. (4.34). For schedule and affordability, we use the monotonically decreasing function given in Eq. (4.35).

$$UF = h^{MI}\left(g, \gamma^{Eff}, LB, UB\right) = \begin{cases} \dfrac{1 - e^{-\gamma^{Eff}(g-LB)}}{1 - e^{-\gamma^{Eff}(UB-LB)}} & \gamma^{Eff} \neq 0 \\[2ex] \dfrac{g-LB}{UB-LB} & otherwise \end{cases} \tag{4.34}$$

$$UF = h^{MD}\left(g, \gamma^{Eff}, LB, UB\right) = \begin{cases} \dfrac{1 - e^{-\gamma^{Eff}(UB-g)}}{1 - e^{-\gamma^{Eff}(UB-LB)}} & \gamma^{Eff} \neq 0 \\[2ex] \dfrac{UB - g}{UB - LB} & otherwise \end{cases} \tag{4.35}$$

The range of the measure of effectiveness g is defined by the stakeholder's least preferred and most preferred values for the particular area of concern. The least preferred value evaluates to a score of zero, whereas the most preferred value evaluates to a score of one in order to maintain uniformity over the various areas of concern domains [200].

The utility assigned to an intermediate value of the measure of effectiveness is determined by the utility function. The shape of the utility function is determined by the risk aversion coefficient γ^{Eff}. For each area of concern, this parameter is set so that the utility function has an aspiration equivalent equal to the stakeholder's target for that measure of effectiveness.

Based on the three individual exponential utility functions, the objective function of the MCTPP is the weighted linear combination of the three exponential utility functions. The weights are provided by the stakeholder based on his preferences about the tradeoffs between the three areas of concerns.

Normative Target-based Decision-Making

The selection of an adequate value for the risk aversion coefficient γ^{Eff} is based on the normative target-based decision making framework because stakeholders are usually not in a position to directly express a value. Instead, stakeholders define their preferences in terms of a target for each area of concern, e.g. the reliability level should be at least 0.95, the development duration (schedule) should be at most eight

years, and the budget (affordability) should be no more than 1.00 (normalized cost), respectively.

We wish to define a utility function for each area of concern that reflects the customer target and is equal to the expected value of the utility function.

From normative target-based decision making theory, we know that there exists a unique effective risk aversion coefficient γ^{Eff} for any stated aspiration-equivalent (target) and probability distribution (likelihood) that results in the same expected utility and aspiration-equivalent of a particular utility function [55, 56]. That is, we can find the appropriate value of the risk aversion coefficient γ^{Eff} by finding the value that satisfies the equality of Eq. (4.36).

$$F(\hat{g}) = \int_{LB}^{UB} \frac{\gamma^{Eff} e^{-\gamma^{Eff} g}}{e^{-\gamma LB} - e^{-\gamma UB}} F(g) dg \qquad (4.36)$$

The cumulative density function $F(\hat{g})$, which expresses the uncertainty of the degree of attainment of the target for each area of concern, is evaluated at the target value $\hat{g}$ (aspiration-equivalent) and set equal to the product of the derivative of the utility function and the cumulative density function (expected utility). The integration limits are defined through the range of the particular area of concern.

In particular, for each area of concern, the stakeholder can provide a probability distribution $F(g)$ for the measure of effectiveness that captures the general uncertainty associated with that Measure of effectiveness. This distribution (over the range $\{LB,UB\}$ for this measure of effectiveness) may be based on the performance of previous development programs or expert opinion. Among the

various distributions, the general Beta, the Uniform or a truncated Lognormal are the preferred ones.

The first two moments, mean and variance, are used to find the general Beta distribution parameters given the range $\{LB, UB\}$. The parameters for the truncated lognormal are found using the bounds of the range $\{LB, UB\}$ as the 5th and the 95th percentile, respectively. In case for the exponential utility function and the use of the general Beta distribution to reflect the uncertainty about the measure of effectiveness, the solution for the risk aversion coefficient γ^{Eff} is found by applying first the integration by parts technique to simplify the integral such that a closed form solution is obtained. In a second step, Brent method is used to solve finally for the risk aversion coefficient γ^{Eff}. Note that this has to be performed appropriately for each area of concern.

4.2.3 Application of the Multiple Criteria Test Planning Problem

The hot-fire test strategies are determined for two scenarios of interest in the context of FLPP: (1) a flight engine development after a successful completion of a demonstrator project at reduced thrust scale and (2) a flight engine development without a prior execution of a demonstrator project. These two scenarios were chosen in order to study the claim that the execution of a prior demonstrator project is cost beneficial for the subsequent flight engine development especially in case of considerable involvement of new technology maturation.

The MCTPPs were solved with a genetic algorithm that is implemented in Palisade's Evolver® [156]. Each run took about three hours on an Intel Duo Core

CPU with 2.40 GHz with an optimization run time setting of 0.01% change of the fitness function within the last 100 trials. The used parameter settings are already given in Table 3-3.

The parameter that drives the overall run time is linked to the solution of the reliability measure of effectiveness, which requires a Markov chain Monte Carlo (MCMC). In order to optimize the MCMC sampling from the posterior, a one-variable-at-a-time with independent candidate density Metropolis–Hastings algorithm was selected which uses already the burn-in samples to tune the independent candidate density properties such that the required acceptance rate of 35% is obtained. The time required to run a single MCMC for nine parameters takes about one minutes with 1000 burn-in samples and chain lengths of 10,000 samples.

Key Liquid Rocket Engine Requirements

The key liquid rocket engine requirements are determined in early design trade-off studies performed at launch vehicle level. The launch vehicle optimizations vary the thrust level, the nozzle area ratio, and the combustion chamber pressure level to obtain optimal solutions for lifting the given payload weight into a particular orbit. An optimum exists between the gross lift off weight of the vehicle and the thrust level of the propulsive system. This optimum should correlate with minimum launcher affordability. Geometric constraints of the launch vehicle limit the nozzle area ratio, and higher levels of the combustion chamber pressure increase the sea-level performance. The mission profile defines the mission durations of the propulsion system(s).

The launch vehicle optimizations are not finalized within the FLPP but the following key liquid rocket engine assumptions were made to perform the study (see Table 4-15). The reduced thrust scale is set to 1400kN for the demonstrator. In addition, the liquid rocket engine architecture is similar to LE-7A which allows the reuse of the fundamental hot-fire test strategy as already defined in Figure 4-4 for the engine system level and Figure 4-5 for the subsystem level.

Table 4-15:   Key Performance Requirements

Performance characteristics	Values
Combustion chamber pressure, bar	150
Vacuum thrust, kN	2,300
Main life cycle (Mission profile)	
Acceptance test, s	150
Acceptance test, s	150
Hold-down, launch commit, s	10
Mission duration, s	300
Number of ignitions, -	4
Reliable cycle at 0.98 reliability, -	5
Reliable life at 0.98 reliability, s	5

Stakeholder Preference

The stakeholder preferences about the three areas of concern affordability, reliability, and IOC were elicited. The main outcomes are listed in Table 4-16. The budget (affordability) figures are proprietary data and are given only as normalized values. In both scenarios, an IOC in 2025 is required.

Based on the customer responses, the three aspiration equivalent exponential utility functions were determined using the techniques presented in Section 4.2.2.

The stakeholder preferences for the three areas of concern influence the search for an optimal test plan because they determine the three utility functions that are included in the fitness function used as the objective function of the MCTPP.

214

Table 4-16:   Customer Preferences

Trade space	Min	Target	Max	Mode	Weights	Remarks
Reliability, -	0.90	0.95	0.995	0.98	0.50	The higher the better
Budget, -	0.67	1.00	1.42	1.17	0.35	Defined by the authors
Development duration, y	7	8	12	10	0.15	Defined by the authors

Measure of Effectiveness Settings

Reliability

The required inputs for calculating the reliability measure of effectiveness are the MLC, the weights for the two failure mechanisms (ζ and $1-\zeta$), the weights for hot-fire test durations which are shorter than full mission duration (w_{ij}^{TP}), the number of anticipated hot-fire test failures, and prior information about the component reliabilities. The following paragraphs provide details for these input parameters. All remaining model parameters are calculated internally by the model setup using the mathematical expressions given in Section 4.2.2.

The MLC is already defined in Table 4-15. The weights ζ and $1-\zeta$ for the two failure mechanism depend on the planned hot-fire test durations and are based on previous European engine development programs (see Section 4.1). The weights w_{ij}^{TP} for planned hot-fire tests which are shorter than full mission duration are based on a quantile regression using data from previous cryogenic liquid rocket engine programs (see Section 4.1). The numbers of anticipated hot-fire test failures are set to zero in all scenarios. The prior information about the component reliabilities depend on the scenarios. Two cases are discussed next.

No prior information is available because Europe has never demonstrated the mastery of a cryogenic staged combustion liquid rocket engine. Therefore, a non-informative (uniform) prior distribution is assumed for the reduced thrust scale demonstrator engine in scenario I as well as for the flight engine development in scenario II.

Prior information is, however, available for the flight engine development after an assumed successful execution of the demonstrator project in scenario I. The data given in McFadden and Shen [80] is used to estimate the prior distribution parameters [27].

Development Duration (IOC)

The required inputs for calculating the IOC measure of effectiveness are limited to the assumptions concerning the number of available test facilities, weekly test cadence, maintenance periods, and mounting and dismounting activities. All remaining model parameters are calculated internally by the model setup using the mathematical expressions given in Section 4.2.2. It should be recalled that the presented model setup does not include any schedule penalty due to the lack of an adequate funding profile.

There are two engine test facilities available in Europe. Both were assumed to be operational for the flight engine development. The demonstrator engine is tested only on one test facility. The component and subsystem test facilities are limited to one for turbomachinery tests and one for combustion devices hot-fire tests. The weekly test cadence is set to 0.6 which may seem to be low but was set to that level to account for possible testing interferences with other hot-fire test facilities. The non-

testing periods due to maintenance and mounting/dismounting activities were set to four months per year for engine level test facilities. No impact was considered for component and subsystem test facilities.

Affordability

The required inputs for calculating the affordability measure of effectiveness are linked to the settings for the design and development cost and the test facility operation cost. All remaining model parameters are calculated internally by the model setup using the mathematical expressions given in Section 4.2.2.

Table 4-17 lists the input parameters for the design and development as well as the TFU cost needed to estimate a single STH cost. The total STH cost is a multiple of the single STH based on the number of required hardware sets defined by the technique discussed in Section 4.2.2. The inputs for the test facility operation cost were already defined in the discussion above.

Table 4-17:   NAFCOM Settings used to assess the Scenarios

Model parameter	Scenario I: Demonstrator	Scenario I: Flight engine	Scenario II: Flight engine
Dev. environment	CAD	CAD	CAD
Manu. environment	Semi-automated	Semi-automated	Semi-automated
MRL	Similar/modified	New	New
Design scope	New technology	New design	New technology
Team experience	Unfamiliar	Normal	Unfamiliar
Engine cycle	SC-Single PB	SC-Single PB	SC-Single PB
MCC pressure, bar	150	150	150
Vacuum thrust, kN	1400	2300	2300

Scenario Assessment

Scenario I: Demonstrator and Flight Engine Development

Reduced Thrust Scale Demonstrator Engine

No customer preference consideration is needed because the programmatic elements are defined by means of the requirements with regards to an IOC for the subsequent flight engine in 2025 and a limited testing scope of 30 hot-fire tests spread over two engine hardware sets. On engine level, four hot-fire test groups were defined, i.e. the 3 seconds tests are used as start-up verification tests, the 30 seconds tests as ramp-up tests, the 150 seconds tests as an intermediate test step, and the 300 seconds tests as full duration tests. The component and subsystem testing scope were defined by systems engineering best practices. An additional hardware was assumed for the component and subsystem level tests. The results in terms of key programmatic elements and test plan characteristics are listed in Table 4-18.

Table 4-18:   Hot-Fire Test Plan Defining Characteristics for Demonstrator

System Integration Level	Number of tests	HTF time, s	Accumulated test time, s
Component			
Preburner	20	10	200
Igniter	35	2	70
Subsystem			
Fuel turbomachinery	20	60	1,200
Ox turbomachinery	20	60	1,200
Combustion devices	20	10	200
System			
Test duration 1	5	3	15
Test duration 2	5	30	150
Test duration 3	10	150	1,500
Test duration 4	10	300	3,000

Key programmatic elements and test plan characteristics:
Total number of hot-fire tests (system level): 30, Number of hardware sets: 3, Reliability projection level: 62.8%, Reliability projection level at 90% confidence: 50.0%, Total duration (schedule): to be finished in 2017/2018, Total budget: 0.770

<u>Subsequent Flight Engine Development</u>

The customer preferences were considered when solving the MCTPP for the flight engine development after the successful completion of the reduced thrust scale demonstrator project. The same numbers of hot-fire test groups as defined for the demonstrator were kept on engine level for the flight engine. The lower bounds for the number of tests for each hot-fire test group is set by the minimum number of hardware sets and the associated MLC, i.e. five in this scenario. The upper bounds are set to 300 for each hot-fire test group. The results in terms of key programmatic elements and test plan characteristics are listed in Table 4-19. The customer targets in terms of development duration and development budget were met. The demonstrated reliability target is marginally not met.

Table 4-19:   Optimized Hot-fire Test Plan defining Characteristics – Flight Engine after Demonstrator

System Integration Level	Number of tests	HTF time, s	Accumulated test time, s
Component			
Preburner	40	10	400
Igniter	35	2	70
Subsystem			
Fuel turbomachinery	160	60	9,600
Ox turbomachinery	160	60	9,600
Combustion devices	210	10	2,100
System			
Test duration 1	5	3	15
Test duration 2	30	30	900
Test duration 3	33	150	4,950
Test duration 4	90	300	27,000

Key programmatic elements and test plan characteristics:
Total number of hot-fire tests (system level): 158, Number of hardware sets: 5, Reliability projection level: 94.4%, Reliability projection level at 90% confidence: 92.2%, Total duration (schedule): 7.5 years, Total budget: 0.737

Scenario II: Flight Engine Development without Demonstrator

The customer preferences were also considered when solving the MCTPP for the flight engine development without a prior execution of a demonstrator project. The results in terms of key programmatic elements and test plan characteristics are listed in Table 4-20. The numbers of hot-fire test groups were increased to six in this scenario to provide an additional degree of freedom for the optimal hot-fire test allocation. The lower and upper bounds for the number of tests for each hot-fire test group is set in a similar way as it was done for scenario I. The minimum number of hot-fire tests is, however, set to 11 which corresponds to the number of hardware sets needed in scenario II. The customer targets in terms of demonstrated reliability, development duration, and development budget were not met. However, the demonstrated reliability level is only marginally not met as it was the cased in the flight development of scenario I. Both demonstrated reliability levels obtained in scenario I and II are at about the same level which allows an easy comparison.

Comparison of Results with Previous Liquid Rocket Engine Programs

Before the two scenarios are compared, the results of the MCTPP are reflected against previous liquid rocket engine key programmatic elements and test plan characteristics.

Koelle [117] provides a figure about the empirical relation of engine reliability versus the number of development and qualification hot firings (see Figure 4-7). The figure was expanded with additional flight engines, results from Section 4.1, and the results obtained from the two scenario assessments. The model results follow rather well the empirically determined relation.

**Table 4-20:** **Optimized Hot-fire Test Plan defining Characteristics – Flight Engine only**

System Integration Level	Number of tests	HTF time, s	Accumulated test time, s
Component			
Preburner	40	10	400
Igniter	35	2	70
Subsystem			
Fuel turbomachinery	160	60	9600
Ox turbomachinery	160	60	9600
Combustion devices	210	10	2100
System			
Test duration 1	5	3	15
Test duration 2	30	30	900
Test duration 3	33	150	4950
Test duration 4	90	300	27000

Key programmatic elements and test plan characteristics:
Total number of hot-fire tests (system level): 281, Number of hardware sets: 11, Reliability projection level: 94.3%, Reliability projection level at 90% confidence: 91.9%, Total duration (schedule): 11.1 years, Total budget: 1.781

The number of hardware sets required in the two scenarios, five and 11, correspond also well with previous experiences if one considers the planned number of tests and the assumed hardware reliability level. Evidences of similar hardware set utilizations for developments are given in Emdee [4]. Section 4.1 further highlights the impact on too stringent hardware reliability requirements on the overall hot-fire test plan credibility. Based on this information, the assumed hardware reliability levels as given in Table 4-15 are reasonable.

The development duration (IOC) results for the two scenarios fit also well with previous evidence given in Emdee [4], e.g. LE-7 with 282 hot-fire tests required 11 years and Vulcain 1 with 278 hot-fire tests ten years, respectively.

Therefore, the results obtained for the two scenarios by solving the MCTPP can be seen as credible based on the comparison of the key programmatic and test plan characteristics with evidences from previous liquid rocket engine programs.
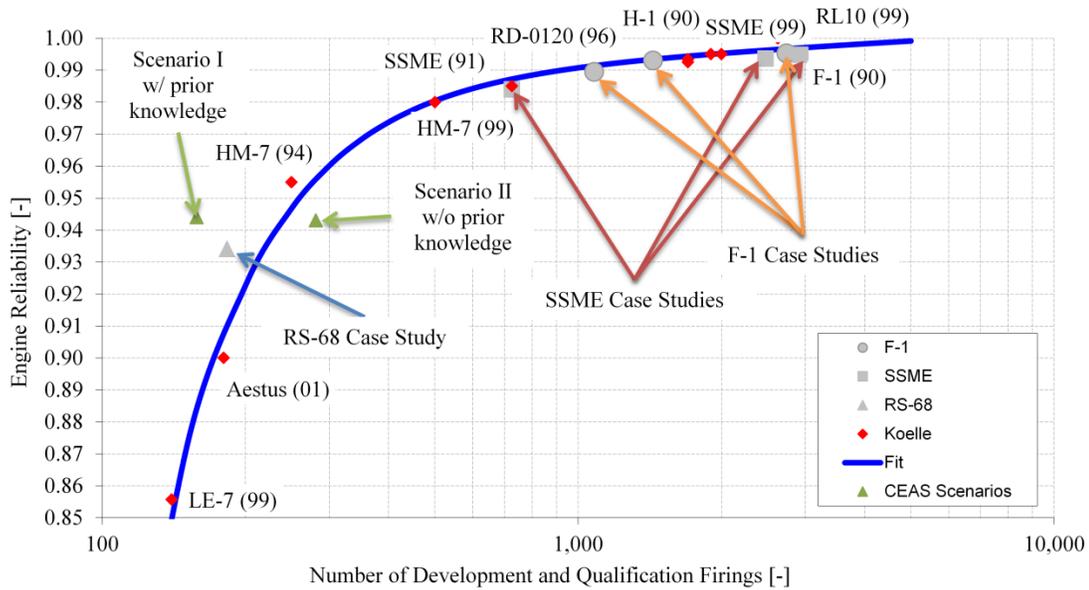

Figure 4-7: **Engine Reliability versus Number of Development and Qualification Hot Firings**

Cost Advantages of a Demonstrator Project

The claim that a prior demonstrator project is cost beneficial for the flight engine development can be confirmed by assessing the results obtained from the two scenarios as summarized in Figure 4-8. The customer targets for the reliability level and the development duration are also included for ease of comparison.

By looking at Figure 4-8, the longer development duration for scenario I should not raise any concern by the stakeholders because the budget for a demonstrator project is limited and as a consequence the work force level allocated to such a project which directly impacts the development duration. In addition, the IOC in 2025 is met even with this longer development duration.
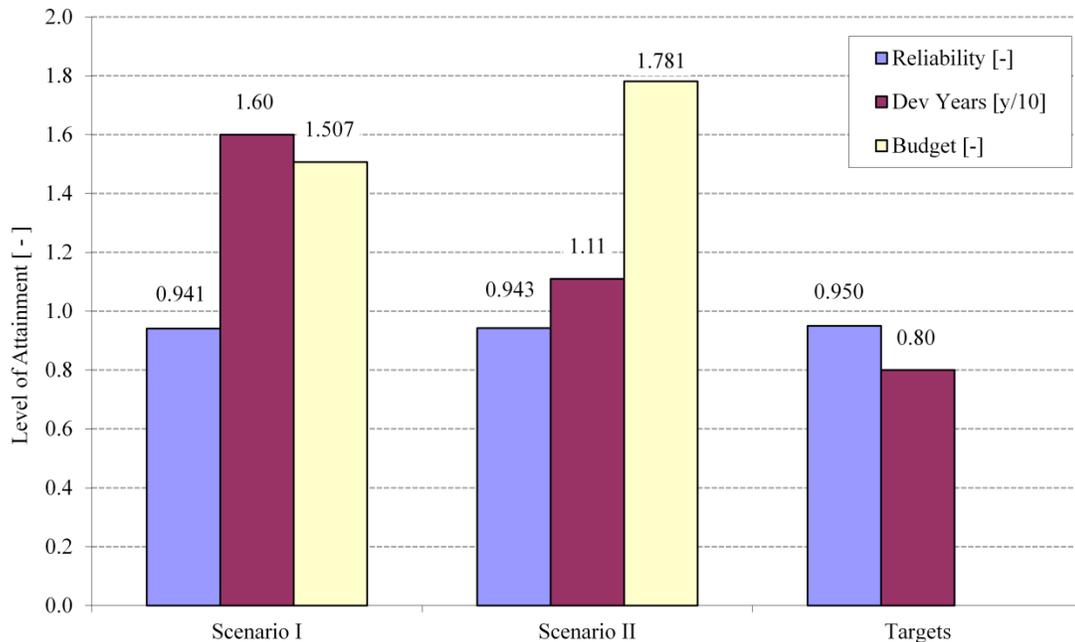
Figure 4-8: **Comparison of Scenarios and Customer Targets**

The Cost of Reliability

The cost of reliability is also a long lasting question in the space industry and by the European Space Agency. The MCTPP setup provides the proper framework for answering this question with quantitative facts. Figure 4-9 shows the impact of an increase in the demonstrated reliability level on the schedule and affordability. The flight engine development of scenario I is included as reference.

By looking at Figure 4-9, the effect on the number of hot-fire tests on engine level, the number of engine hardware sets, development duration (schedule), and affordability (development cost) of an increase of the reliability from 0.95 to 0.98 (roughly 3%) can be assessed. The number of hot-fire tests on engine level is increased by 260%, the number of hardware sets by 320%, the development duration by 150%, and the affordability by 270%, respectively. The number of hardware sets that are needed can be significantly reduced in case of an enhanced life capability of

the piece parts and subassemblies is given but at the expense of an increase in the production cost for later flight utilization. The development duration may be significantly reduced by erecting additional test facilities for engine level tests but at the expense of an increase in development cost.
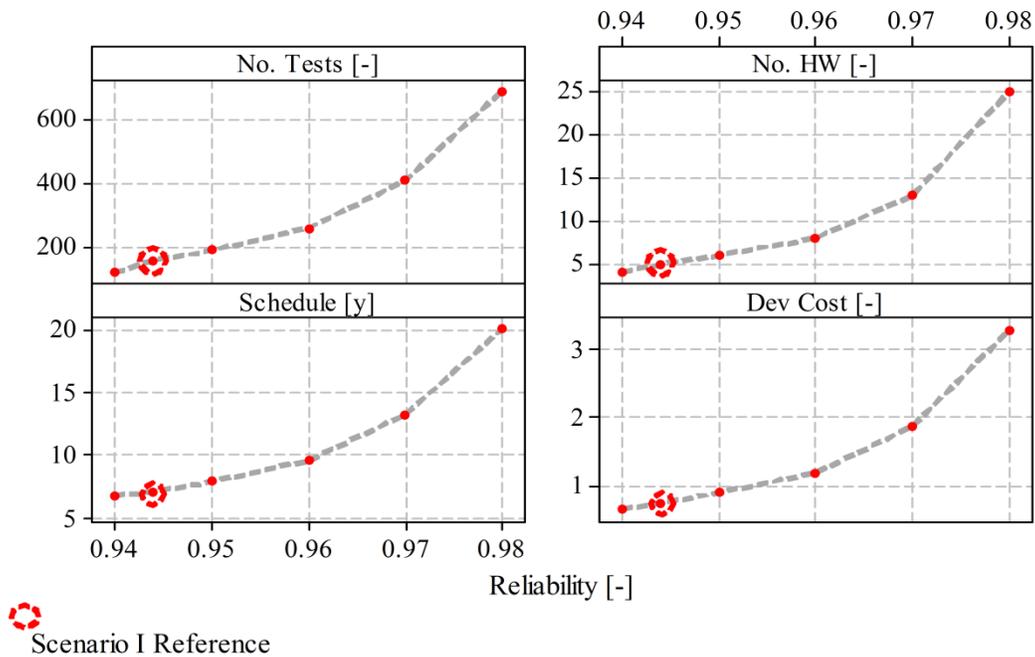
Figure 4-9: **Impact of Reliability Level on Development Schedule and Cost for Flight Engines**

4.2.4 Conclusion

The MCTPP presented here supports early design tradeoff studies by providing quantitative relationships between the hot-fire test plan and reliability, schedule, and affordability performance measures. Moreover, the model allows one to find the best hot-fire test strategy that meets customer targets for these performance measures. (The best test strategy has the smallest number of tests and hardware sets.)

In addition, the study substantiated the claim that a prior test bed or demonstrator project reduces the development cost of the actual flight engine in case

there is a substantial technology maturation need. Scalability aspects for the technology maturation at lower scale are adequately accounted for the different components and subsystems through the prior in the Bayesian framework.

The sensitivity of the development schedule and development cost to an increased level of reliability is quantitatively confirmed as well.

Of course, optimal plans increase the likelihood of success but do not guarantee it. The actual flight mission success is still subject to good workmanship, brilliant engineers, and luck.

4.3   Planning, Tracking, and Projecting Reliability Growth:

A Bayesian Approach

Liquid rocket engine reliability growth modeling is a blend of art and science because of data scarcity and heterogeneity, which result from the limited number of engine development programs as well as testing profiles that are much different from the actual mission profile. In particular, hot-fire tests are shorter than full mission duration due to test facility limitations and some of them are performed at extreme load points to demonstrate robustness and design margin.

As a response to modern liquid rocket engine hot-fire testing profiles, which require a new reliability growth modeling approach, this section presents a new, fully Bayesian estimation based methodology that estimates the system reliability without the MTBF metrics; instead, it takes into account all component, subsystem, and system level hot-fire test data. The Bayesian estimation provides naturally the framework that is needed to apply the methodology in the three areas of reliability

growth: planning, tracking, and projection because pseudo, actual, and the combination of both pseudo and actual hot firings test data can be used to estimate the system level reliability.

The methodology is applied to planning, tracking, and projecting reliability growth and illustrated using an example. In the example, a system reliability target must be demonstrated in a TAAF program. The system reliability target defines the scope of the hot-fire test plan for the reliability growth planning using pseudo numbers for the planned hot-fire tests. At each occurrence of a failure, the methodology is used in the context of reliability growth tracking, i.e. the attained system level reliability is estimated. The test plan is updated to reflect the need for additional tests to meet the system reliability target. Reliability growth projection is easily performed using either specific projection models or the prior distribution that features a knowledge factor to model the specified level of fix effectiveness.

4.3.1 Reliability Growth

Reliability growth is typically attained through a formal TAAF program that discovers and corrects design deficits. Reliability growth models are used for test planning, tracking reliability throughout the program, and projecting the reliability when the tests are completed. The two most widely used reliability growth models are the empirical Duane and the analytical Crow/AMSAA, which both use the MTBF to estimate the reliability growth rate. The MTBF is calculated from the total accumulated test time divided by the total number of failures without considering the operational loads, durations, and sequences of the applied stresses, which highly affect the failure rate and as a consequence the MTBF metric [201]. Therefore,

ignoring the applied stresses makes the Duane and Crow/AMSAA models questionable for cases in which the testing profiles differ, in terms of applied stresses, significantly from the stated mission profile [31, 149].

Modern liquid rocket engine hot-fire testing profiles belong to such cases because the testing profile is a potpourri of tests that are shorter than full mission duration and tests performed at extreme load points to demonstrate robustness and design margins. Therefore, neither the Duane nor the Crow/AMSAA data analysis may be any longer best practice as the following brief discussion highlights.

Historically, liquid rocket engine hot-fire testing profiles were used to comply with a formal reliability demonstration as it was the case for the F-1 and J-2 engines. These hot-fire testing profiles followed adequately well the operational loads, and, as a consequence, the Golovin and empirical Duane models were successfully applied [147].

However, formal reliability demonstration hot-fire testing profiles are lengthy and cost prohibitive, which led to the DVS approach that was applied to the SSME certification. The Crow/AMSAA model, one of the two reliability growth models used, initially estimated an increase of the MTBF (indicating reliability growth), but the system reliability declined towards the end of the testing profiles although overall testing experience would have suggested an increase in the system reliability [30].

Most recently, an objective based variable test/time philosophy was used to qualify the RS-68 liquid rocket engine while lowering the development cost and reducing the development schedule. To achieve these objectives, the hot-fire testing profile included extreme load points to demonstrate robustness and design margin

[118]. Based on the SSME experience, the RS-68 engine testing profile should have been even more difficult to analyze with the Duane and Crow/AMSAA models and to estimate a system reliability that is based on the MTBF metric.

4.3.2   Methodology

The methodology is based on the Bayesian aggregation of multilevel binomial test data [93] but is extended with the notion of equivalent mission to account for the operational loads, durations, and sequences of the applied stresses that are present in the specific testing profiles but are unlike those in the mission profile [27].

The Bayesian aggregation of multilevel binomial test data uses a functional network that is based on the principles of the reliability block diagram technique [202]. The functional network serves two purposes: (1) It defines the fundamental test strategy that defines also the hot-fire test configurations at the component, subsystem, and engine system levels and (2) it is used to derive the governing likelihood function that combines simultaneously all available multilevel hot-fire test data. It should be noted that the functional component level nodes correspond to individual physical components or to a CCCG of the actual physical system architecture. Figure 4-10 depicts an example of such a functional network.
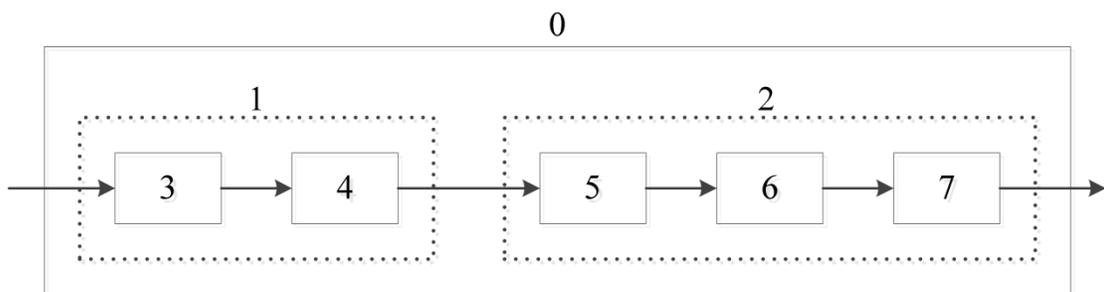
Figure 4-10:   Functional Network

The methodology begins with the set of prior distributions about the reliability of each functional node. Each prior is a modified Beta distribution with three parameters: α_i and β_i, which can be derived from previous engine reliability data as given in [80], and the knowledge factor (or relevance factor) ϕ_i, which measures the level of transformation of similar designs into new product designs and is derived from methods defined in [81, 82]. It can be determined qualitatively or quantitatively with methods described in [81, 83]. Thus, the prior for node i is the following distribution:

$$\pi_{0,i} = f\left(\pi_i \mid \alpha_i, \beta_i, \phi_i\right) = \frac{\pi_i^{\alpha_i\phi_i - 1}\left(1 - \pi_i\right)^{(\beta_i - 1)\phi_i}}{\mathrm{B}\left[\alpha_i\phi_i, (\beta_i - 1)\phi_i + 1\right]} \tag{4.37}$$

In addition, the methodology requires for each functional node the number of equivalent trials, EQM_{ij}^{TP}, and the number of equivalent successes, $EQM_i^{TP\langle S\rangle}$. The notion of equivalent mission is introduced because it captures the two fundamental failure mechanisms (characterized as stress-increased and strength-reduced) that are present in liquid rocket engine piece parts and subassemblies. The number of equivalent trials, EQM_{ij}^{TP}, is calculated as follows:

$$EQM_{ij}^{TP} = \zeta_{ij} \cdot \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1 - \zeta_{ij}\right) \cdot \frac{D_{ij}^{TP}}{CFD^{MP}} \tag{4.38}$$

The first term relates the stress-increased (cyclic) and the second term the strength-reduced (time-dependent) failure mechanism, respectively. Both terms are weighted and relate the specific testing profiles to the mission profile.

These quantities are derived from the characteristics of the testing profiles as follows. For the number of equivalent trials, the testing profile duration D_{ij}^{TP} depends

upon the test duration, an acceleration factor, which is introduced to model the extreme load points, and a weighting factor accounts for the hot-fire tests that are shorter than full mission duration. Note that these different testing profiles at functional node i are accounted for by defining specific hot firings j. The acceleration factor, AF_{ij}^{TP}, is based on the acceleration testing theory [169] and is not further discussed. The weighting factor, w_{ij}^{TP}, is based on a likelihood function that models the union of two mutually exclusive events: (1) a failure that takes place during the start-up and steady state operation (ordinary failure) and (2) a failure that takes place during the shutdown operation [124].

$$D_{ij}^{TP} = AF_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP} \tag{4.39}$$

Because the individual hot-fire test durations are usually different within each functional node i, which is reflected through subscript j, we use the following to calculate EQM_i^{TP}:

$$EQM_i^{TP} = \sum_{j=1}^{J_i} \left(\zeta_{ij} \cdot \frac{NFC_{ij}^{TP}}{NFC^{MP}} + \left(1 - \zeta_{ij}\right) \cdot \frac{NFC_{ij}^{TP} AF_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP}}{CFD^{MP}} \right) \tag{4.40}$$

For the number of equivalent successes at node i, the testing profile duration FD_{ij}^{TP} depends upon the actual failure time $FD_{ij}^{TP\langle F \rangle}$, the acceleration factor, and the weighting factor:

$$D_{ij}^{TP\langle F \rangle} = AF_{ij}^{TP} w_{ij}^{TP} FD_{ij}^{TP\langle F \rangle} \tag{4.41}$$

Then, the number of equivalent successes at node i is derived using an equation similar to equation (4.40):

$$EQM_i^{TP\langle S\rangle} = EQM_i^{TP} -$$

$$\sum_{j=1}^{J_i}\left(\zeta_{ij}\cdot\frac{NFC_{ij}^{TP}}{NFC^{MP}}+\left(1-\zeta_{ij}\right)\cdot\frac{NFC_{ij}^{TP}AF_{ij}^{TP}w_{ij}^{TP}FD_{ij}^{TP\langle F\rangle}}{CFD^{MP}}\right) \quad\quad (4.42)$$

After these quantities are derived, the Bayesian estimation uses Bayes' Theorem to define an unscaled posterior distribution for the parameters that must be estimated. The unscaled posterior distribution is defined through a likelihood function which models the data and a set of prior distributions for the parameters of the model (the likelihood function) that is given as

$$\pi\left(\underline{\theta}\mid Data\right)\propto\prod_{i=1}^{I}\pi_i^{EQM_i^{TP\langle S\rangle}}\left(1-\pi_i\right)^{EQM_i^{TP}-EQM_i^{TP\langle S\rangle}}\prod_{i=1}^{I}\pi_i^0\left(\underline{\theta}\right) \quad\quad (4.43)$$

The parameter vector, $\underline{\theta}$, of this unscaled posterior distribution is estimated with a one-variable-at-a-time MH algorithm. Important metrics of this solution strategy are the acceptance rate of the acceptance probability as well as the autocorrelation and convergence of the Markov chain of the proposed candidates. The candidates are drawn on a logit scale for which the proper acceptance rate is around 0.35. In order to obtain that rate, the burn-in period of the Markov chain is used to tune the standard deviation of the candidate density function. The autocorrelation function is used to obtain the lag at which the Markov chain is thinned. Finally, the convergence of the accepted Markov chain was visually inspected by means of trace plots.

The combined likelihood function of equation (4.43) is found as follows: The fundamental test strategy defines the test configurations that are expressed in terms of nodes. Using the example depicted in Figure 4-10, the system level is node 0, the two subsystem nodes would be 1 and 2, and the functional component level nodes are 3 to

7. The subsystem node 1 and 2 reliabilities are expressed as $\pi_1 = \pi_3\pi_4$ and $\pi_2 = \pi_5\pi_6\pi_7$. The system level node 0 reliability is given as $\pi_0 = \pi_1\pi_2$ or equivalently as $\pi_0 = \pi_3\pi_4\pi_5\pi_6\pi_7$. Finally, these functional component, subsystem, and system level reliabilities are inserted in equation (4.43) to combine simultaneously all level test data.

The probabilities of the mutually exclusive events that define the weighting factor, w_{ij}^{TP}, for the different testing profiles are also found by applying Bayes' Theorem to the likelihood function and a prior distribution for the model parameters. The likelihood function that describes the mutually exclusive events is based on a quasi-multinomial distribution. Uniform distributions are used as prior.

Figure 4-11 depicts empirical evidence for the weighting factors for different liquid rocket engines using the data given in [148]. The figure includes also the weighting factors that are used in the illustrative example described in 4.3.3.
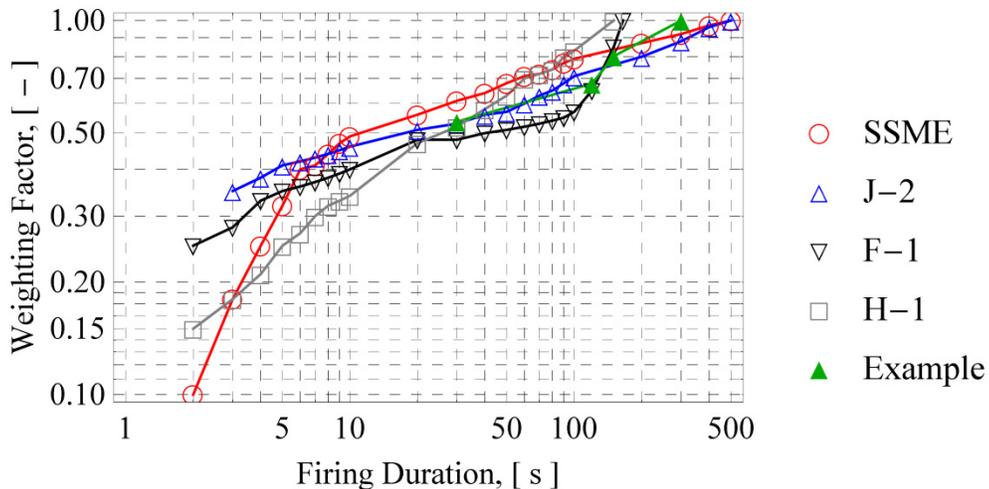
Figure 4-11:   Weighting Factor versus Hot Fire Test Duration

232

The equivalent trials, EQM_i^{TP}, can be related to an equivalent life for the hardware components in order to estimate the number of hardware sets required to complete the specific testing profile. The reliable equivalent life is given as:

$$EQL_{RbyC}^{MP} = \xi \frac{c_{RbyC}^{MP}}{NFC^{MP}} + (1-\xi) \frac{t_{RbyC}^{MP}}{CFD^{MP}}$$

(4.44)

The definitions of the reliable cycle, c_{RbyC}^{MP}, and the reliable time, t_{RbyC}^{MP}, may be either based on physics-of-failure models if available or on expert elicitation.

It should be noted that the structure of the reliable equivalent life is the same as for the equivalent mission. Therefore, the number of required hardware sets can be estimated with

$$HW_{RbyC}^{MP} = \frac{EQM_{RbyC}^{MP}}{EQL_{RbyC}^{MP}}$$

(4.45)

Equation (4.45) can be applied using the overall number of equivalent trials or the equivalent trials that are associated with the relevant functional node i level.

The Bayesian estimation methodology is applied next to an illustrative example that describes the application in the context of reliability growth: planning and tracking. The area reliability growth projection is not explicitly demonstrated but once the system reliability is estimated various projection models can be applied [148].

4.3.3 Illustrative Example

As an illustrative example, we consider a hypothetical liquid rocket engine TAAF program that includes a contractual reliability growth objective (system reliability target) for a cryogenic Gas Generator main stage engine.

233

The physical system architecture is similar to the RS-68 or Vulcain 2 liquid rocket engine. Therefore, the physical architecture can be described with nine functional component nodes in series.

The thrust class of the new engine is a significant increase compared to previous designs but our a priori knowledge is that the design authority has mastered a staged combustion engine at lower thrust scale. Based on this, we decided to use a knowledge factor ϕ_i of 0.80 for the functional component level node priors with distribution parameters $\alpha_i = 38$ and $\beta_i = 0.7$.

Furthermore, the stated engine mission profile consists of a 100 seconds acceptance test, a 10 seconds engine ground start hold-down with launch commit criteria abort, and a 300 seconds flight mission. The contractor and agency selected specific testing profiles (hot-fire test plan), which includes component level, subsystem, and system level tests. Table 4-21 lists these testing profiles in terms of number of tests, hot-fire test duration, and acceleration factor to indicate the severity of the hot-fire test conditions.

Table 4-21:   Testing Profile

Node	No. of Tests	Hot-fire Duration, s	Acceleration Factor
Gas Generator	60	50	1
Powerpack	10	100	1
Engine, Group 1	70	30	1
Engine, Group 2	50	120	1
Engine, Group 3a	35	150	1
Engine, Group 3b	35	150	5
Engine, Group 4a	20	300	1
Engine, Group 4b	20	300	5
Total / Accumulated	230	30600	

Based on this data, setting the weight $\zeta = 0.5$, and the application of Eq. (4.43), the average system level reliability estimate is 0.956.

We now consider the impact of failures. We will consider a scenario in which three failures occur (see Table 4-22). The failures are fully defined by means of the hot-fire test order number, the failure time, and the affected physical component.

Table 4-22: Assumed Failure Metrics

Node	No. of Tests	Failure Time, s	Component
Engine, Group 1	45	150	Turbopump, ox
Engine, Group 2	100	300	Gas Generator
Engine, Group 3a	150	300	Turbopump, fu

In this scenario, the TAAF program has started, the first couple of hot-fire tests are successful, and then the failures occur. At each failure event, the following updating procedure is performed:

- the likelihood function for the weighting factor, w_{ij}, is updated with the failure event and the Bayesian estimation calculates new weights that are used in Eq. (4.39) and Eq. (4.41),

- Eq. (4.40) and (4.42) are equated using the new weights and the actual failure event time,

- the a priori knowledge is considered as non-existing for the failed component that modifies the prior distribution, and

- the recalculation of the functional component level reliabilities using Eq. (4.43) in order to update the system level reliability.

Table 4-23 lists the resulting system level reliability estimates at each failure occurrence and demonstrates the application of the methodology in the context of reliability growth tracking.

Table 4-23:   Reliability Growth Tracking

Tracking steps	Test number	Reliability level
Failure 1	45	0.831
Failure 2	100	0.861
Failure 3	150	0.879

The next step in our TAAF program scenario is the definition of the remaining hot-fire test effort given the failure occurrence in order to attain the contracted system reliability target (reliability growth planning). Either of two assumptions can be made: (1) no additional failures will occur during the remaining hot-fire tests or (2) additional failures will occur and the number of the additional failures is estimated using reliability growth projection models. This work considers only the first case and updates the reliability growth planning hot-fire test scope at each time when an assumed failure occurred. Table 4-24 lists the consequences in terms of additional hot-fire tests and as a delta (difference) from the initial hot-fire test plan to attain the contracted system reliability target, i.e., 0.956. Figure 4-12 depicts the described scenario graphically.

Table 4-24:   Test Scope Consequences

Events	Additional hot-fire tests	Delta from initial test plan
Failure 1	20	20
Failure 2	30	50
Failure 3	25	75

The practical importance to both contractors and the space agency should be noted because the methodology not only estimates the attained or planned system reliability

to assure the mission success but also provides the hot-fire test scope during the requirements definition and after a failure occurrence. Thus, the presented Bayesian methodology in the context of reliability growth is also a valuable management tool for program managers.
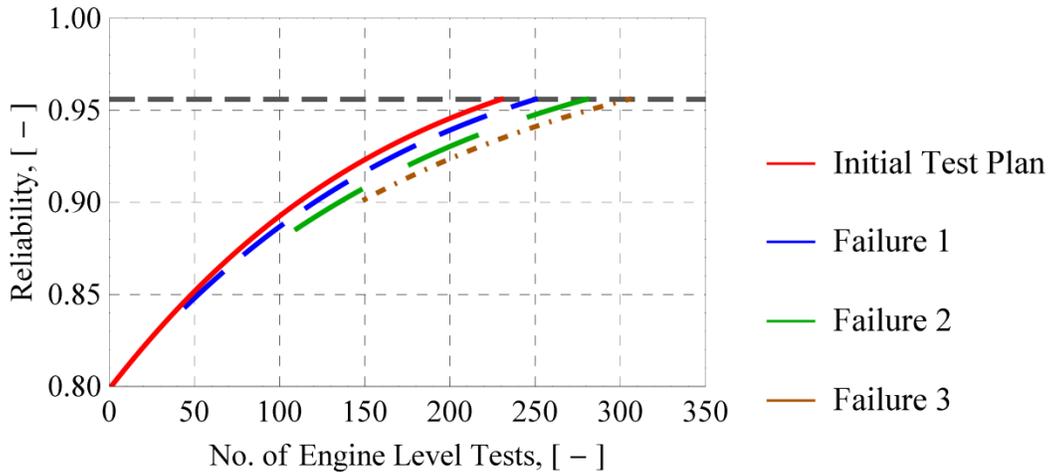
Figure 4-12: Reliability Growth Planning and Tracking

4.3.4 Conclusion

This section presented a new, fully Bayesian estimation based methodology that provides a true alternative to the empirical Duane and analytical Crow/AMSAA models. The key features that distinguish the proposed methodology from the classical models are the aggregation of multilevel test data, the neutralization of the differences of the specific testing profile to the mission profile, the inclusion of a priori knowledge, and the capability to apply it to all three main areas of reliability growth: planning, tracking, and projection.

The illustrative example demonstrated the practical use of the proposed methodology by quantifying the impact of failures on the estimated system reliability in the context of reliability growth planning, tracking, and projection. The illustrative

example also highlighted the importance of the methodology as a risk management tool by providing quantitative figures for the hot-fire test scope definition that drives both the development cost and development schedule.

4.4 Preference-based Risk-informed satisficed Decision-Making with Epistemic Uncertainty

Motivated by the problem of developing and certifying a liquid rocket engine, this section describes a multiobjective optimization approach that incorporates user preferences about the objectives (expressed as both targets and relative weights) and epistemic uncertainty about design problem parameters. The proposed approach supports program management decisions that involve the correlated objectives of affordability, reliability, and initial operational capability and include technical, financial, and schedule program risks.

Section 4.4.1 describes some general considerations about the problem. Section 4.4.2 describes the preference-based risk-informed decision-making problem formulation including the specific model details. Section 4.4.3 presents the main results, and Section 4.4.4 summarizes the main findings.

4.4.1 Introduction

Managing the development of a new product involves decision-making with multiple, usually conflicting, and correlated objectives that include program risks and epistemic uncertainty. A single optimal solution is not attainable with respect to all of the objectives, but Pareto-optimal solutions exist. In addition, the decision-makers are

satisficers, as the theory of bounded rationality proposes [24-26], who seek solutions from a limited set of alternatives.

This section presents a risk-informed satisficed decision-making method for a new liquid rocket engine development. The programmatic elements (the objectives) are affordability (cost), demonstrated reliability, and Initial Operational Capability (IOC) (development duration), and the decision-maker has a set of targets for and uncertainty about each objective. In addition, the decision-maker provides a set of minimum product characteristics such as the vacuum thrust and the main combustion chamber pressure for the selected thermodynamic cycle architecture [3].

In this context, the decision variables describe the test plan that will verify the product's inherent reliability. These include hot-fire tests at component, subsystem, and system level. It is also well-known that the test-analyze-and-fix (TAAF) cycle failure assumptions strongly influence program decisions because not only is an impact given for reliability but also for affordability and IOC. In order to predict the number of TAAF cycle failures, the product characteristics are used to define the newness of the liquid rocket engine system.

In the problem specific context, it may seem straightforward to apply evolutionary multiobjective optimization (EMO) algorithms with Pareto dominance-based fitness evaluation but it is not because the objectives are not only conflicting (the classical case) but also correlate among each other and incorporate uncertainty. Studies about the impact of correlated objective functions have already shown that the application of elitist multiobjective non-dominated sorting genetic algorithms such as the NSGA-II or the SPEA either will find only solutions around the center of the

Pareto front (proposed remedies are the incorporation of preference as described in [203-205]), or will generate similar sets of Pareto-optimal solutions compared to dimensionality reduction approaches, in the most extreme case a single-objective genetic algorithm (SOGA) or the multiobjective evolutionary algorithm based on decomposition (MOEA/D) [155], but at the expense of computational time (which is demonstrated in Section 4.4.3 using the NSGA-II) [206, 207]. The inclusion of uncertainty in the objective functions is usually modeled as noise by adding an error term that is generated by a statistical distribution [208-212]. However, in program management related decision-making such an approach is impractical because the decision-maker will not be able to define an adequate statistical distribution a priori; the impact of the epistemic uncertainty depends upon the design solution and cannot be modeled as noise.

Our proposed approach, a preference-based risk-informed decision-making problem formulation, is based on a SOGA using the weighted sum approach [153, 154] which addresses well the shortcomings of the present mainstream EMO solution strategies because the approach not only is computationally more efficient but also incorporates the decision-maker's preferences, targets, and the uncertainty about the objectives. Note that the targets and uncertainty define the decision-maker's risk attitude for each of the objectives using utility-probability duality [55]. The weights are determined by means of the preference programming method to include already the decision-maker's uncertainty about the weights in the weighting elicitation process [213], but it will be seen that the inclusion is not of first order importance and that other methods could have been used such as ranking methods, rating methods,

weighted sum approaches or the concept of hypothetical equivalents and inequivalents but with the limitation that these methods result in single weight estimates [214]. The fitness function is based on truncated exponential utility functions [200, 215] that not only normalize disparately-scaled objective spaces but also allow for a fitness evaluation on the score values which measures' of effectiveness depend on the decision variables and on the decision-maker's risk attitude for each objective space. The measures of effectiveness are determined by specific affordability, reliability engineering, and IOC models, respectively.

4.4.2 Satisficing Problem Formulation

Before considering the specific problem that motivated this research, we will present the general approach. Consider a design optimization problem with M performance measures. Let $\bar{x} = (x_1, \ldots, x_I)$ be the vector of I decision variables. Each variable has a lower bound $x_i^{(L)}$ and an upper bound $x_i^{(U)}$.

For each of the M performance measures, let g_m be the value of the performance measure, which is determined by the evaluation function $h_m(\bar{x})$, and let LB_m and UB_m be the lower and upper bounds. These bounds express the decision-maker's beliefs about the possible range of the performance measure. Let $\hat{g}_m$ be the decision-maker's target for this performance measure. This value, as explained later in this section, determines γ_m^{Eff}, the effective risk coefficient that defines the shape of the utility function UF_m. Finally, let w_m be the weight associated with the performance measure.

The optimization problem can then be formulated as the sum of weighted normalized utility functions as follows:

$$\text{Maximize} \qquad FF = \sum_{m=1}^{M} w_m UF_m$$

$$UF_m = h_m\left(g_m, \gamma_m^{Eff}, LB_m, UB_m\right) \quad m = 1, 2, ..., M;$$

$$g_m = f_m(\bar{x}) \qquad\qquad m = 1, 2, ..., M \qquad (4.46)$$

$$\text{subject to} \qquad g_m \geq LB_m \qquad\qquad m = 1, 2, ..., M;$$

$$g_m \geq \left(-UB_m\right) \qquad\qquad m = 1, 2, ..., M;$$

$$x_i^{(L)} \leq x_i \leq x_i^{(U)} \qquad\qquad i = 1, 2, ..., I.$$

The concept of utility functions with the associated risk attitudes is discussed in Keeney and Raiffa [215]. In this approach, two types of truncated exponential utility functions are used: a monotonically increasing function, denoted with the superscript MI, for objectives (like reliability) that should be maximized, and a monotonically decreasing function, denoted with the superscript MD, for objectives that should be minimized (like affordability and IOC). The general expressions, in which UF is a function of g, are the following equations:

$$UF = h^{MI}\left(g, \gamma^{Eff}, LB, UB\right) = \begin{cases} \dfrac{1 - e^{-\gamma^{Eff}(g - LB)}}{1 - e^{-\gamma^{Eff}(UB - LB)}} & \gamma^{Eff} \neq 0 \\[4mm] \dfrac{g - LB}{UB - LB} & otherwise \end{cases}$$

$$\qquad (4.47)$$

$$UF = h^{MD}\left(g, \gamma^{Eff}, LB, UB\right) = \begin{cases} \dfrac{1 - e^{-\gamma^{Eff}(UB - g)}}{1 - e^{-\gamma^{Eff}(UB - LB)}} & \gamma^{Eff} \neq 0 \\[4mm] \dfrac{UB - g}{UB - LB} & otherwise \end{cases}$$

where γ^{Eff} is the effective risk coefficient that defines the shape of the utility function and expresses the decision-maker's risk attitude. The impact of different risk-attitudes is discussed in Section 4.4.3.

Given the target $\hat{g}_m$, the bounds LB_m and UB_m, and $F_m(\cdot)$, the distribution of the decision-maker's uncertainty about this performance measure, the effective risk coefficient γ_m^{Eff} can be found using utility-probability duality [55]. In particular, γ_m^{Eff} is the value that generates the function UF_m that satisfies the following equation. The adaptation for monotonically decreasing utility functions is simply by symmetry (see Section 3.2.1).

$$F_m(\hat{g}_m) = \int_{LB_m}^{UB_m} F(g_m) \frac{UF_m}{dg_m} dg_m \qquad (4.48)$$

Figure 4-13 depicts examples of utility functions that convey the decision-maker's risk attitudes risk-averse, risk-neutral, or risk-seeking. Note that the resulting risk attitudes are in good agreement with the prospect theory given in Kahneman and Tversky [48].
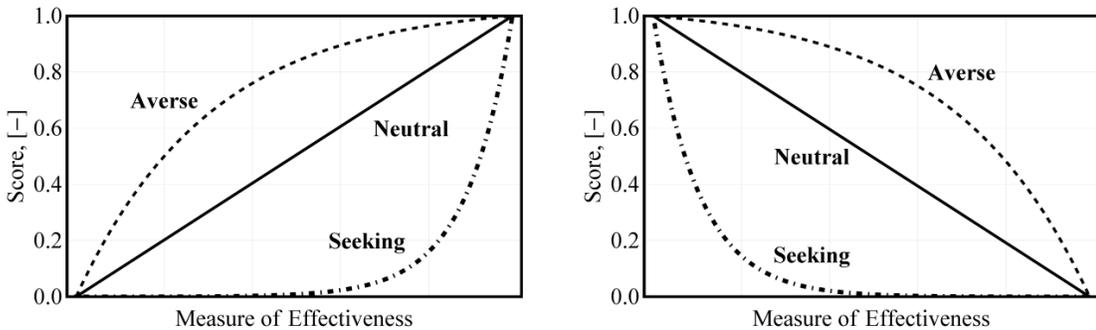
Figure 4-13: Example of Utility Functions for different Risk Coefficient Settings

Models to Determine the Measures of Effectiveness for the Objectives

This section describes the performance measures for the liquid rocket engine development and certification application and how they depend upon the design variables (the number of hot-fire tests). Because deriving these relationships is not

the purpose of this work, the reader is referred to other sources for details that have been omitted.

Affordability

Minimizing the cost of developing a liquid rocket engine is an important objective. In the application considered here, its affordability is measured by the development cost C^{Dev}. This cost can be estimated using the liquid rocket engine cost model (LRECM) [6], which is implemented in the NASA/Air Force Cost Model (NAFCOM®) [13], in combination with a specific effort-driven hot-fire test model using the results of the reliability-as-an-independent-variable (RAIV) strategy (see Section 3.3.2). The cost model used in this work was discussed in detail in Section 3.3.1.

Reliability-As-an-Independent-Variable Strategy

The number of hot-fire tests, which corresponds to the decision variables ($\bar{x}$) of the genetic algorithm, is used to determine the number of equivalent missions that the liquid rocket engine undergoes during testing (see Section 4.1). This is used to estimate the objective demonstrated reliability at liquid rocket engine system level and the number of hardware sets (which are used in the affordability model). In addition, it is used to predict the number of TAAF cycle failures using the knowledge transfer factor ϕ that reflects the newness of the liquid rocket engine. The RAIV strategy presented in Section 3.3.2 was used in this study.

Initial Operational Capability

Minimizing the initial operational capability (IOC), which is equivalent to the development duration D^{Dev}, is an important objective that depends upon the number

of hot-fire tests. This performance measure depends on the design maturity (TRL), the design process maturity (experience of the team), the reliability-by-credibility (confidence) requirement that determines the hot-fire test plan, the hot-fire test cadence, the number of test facilities, and the yearly funding level. Therefore, the Schedule Estimating Relation (SER) for the IOC can be expressed as the sum of the design and development duration D^{DD} and the test facility occupation duration D^{TP}. Section 0 presented the details of how these durations are estimated.

A Typical Liquid Rocket Engine Development Program Tradeoff Decision

This section describes the development and certification optimization for a particular liquid rocket engine development scenario. A typical liquid rocket engine development program tradeoff decision is concerned about selecting the best alternative among various design solutions considering the three objectives affordability, reliability, and IOC. The decision-maker defines the targets, the uncertainty, and the weights for the objectives. Next, the space transportation system requirements, mainly thrust, vacuum specific impulse, propellant combination, propellant mixture ratio, and geometric constraints, are transferred into liquid rocket engine requirements that define the possible set of design alternatives.

The system that is the subject in this study, assuming an early program phase, is the U.S. liquid rocket engine RS-68. Details about the actual project performance are given in Wood [118].

Decision-maker's Targets, Uncertainty Bounds, and Weights

The decision-maker provides the objective functions' targets and weights. In addition, the decision-maker may express the ranges for the objectives to express their

uncertainty about their targets. Table 4-25 lists the decision-maker's responses (note that normalized figures are given for the affordability). Based on these inputs, the corresponding effective risk coefficients γ^{Eff} can be determined; these and the associated risk attitudes are given in Table 4-26. In addition, the decision-maker is also asked to provide the product and processes characteristics that are the required inputs for the LRECM.

Table 4-25:   Decision-maker's Uncertainty Bounds, Targets, and Weights

Objectives	Min	Max	0.05 LB	0.95 UB	Targets	Weights
Affordability, MU	0	2	0.930	1.350	1.200	0.20
Reliability, -	0	1	0.9663	0.9974	0.958	0.65
IOC, y	0	13	7.50	12.00	10.25	0.15

Table 4-26:   Decision-maker's Risk Attitudes

Objectives	γ^{Eff}	Risk Attitude
Affordability	7.1	Risk averse
Reliability	0.017	Risk neutral
IOC	29.1	Risk averse

The values given for the objective reliability need more explanation because it may seem odd that the target is outside the lower and upper bounds. A distinction must be made between the mission reliability for the new liquid rocket engine once it is in operation and the reliability demonstration target that is set by the decision-maker due to the cost prohibitive design verification hot-fire test plan. Note that the impact of the target value on the risk attitude is deferred until Section 4.4.3.

Description of a Liquid Rocket Engine Design Alternative

The engine components, from a main function point of view and that are classically at risk, are depicted in Figure 4-14. Note that the node notion indicates also the testing strategy for liquid rocket engines, i.e., node 0 identifies the components that are used

for the system level hot-fire tests whereas node 1 defines the subset of the components that are considered for the subsystem level hot-fire tests. Node 6 and Node 10 are not included in the system level test configuration because they do not contribute to the main test objective, i.e., the turbomachinery. On component level, only node 5 is, however, considered for hot-fire tests in order to mitigate the technical risks of combustion instability for a gas generator component.
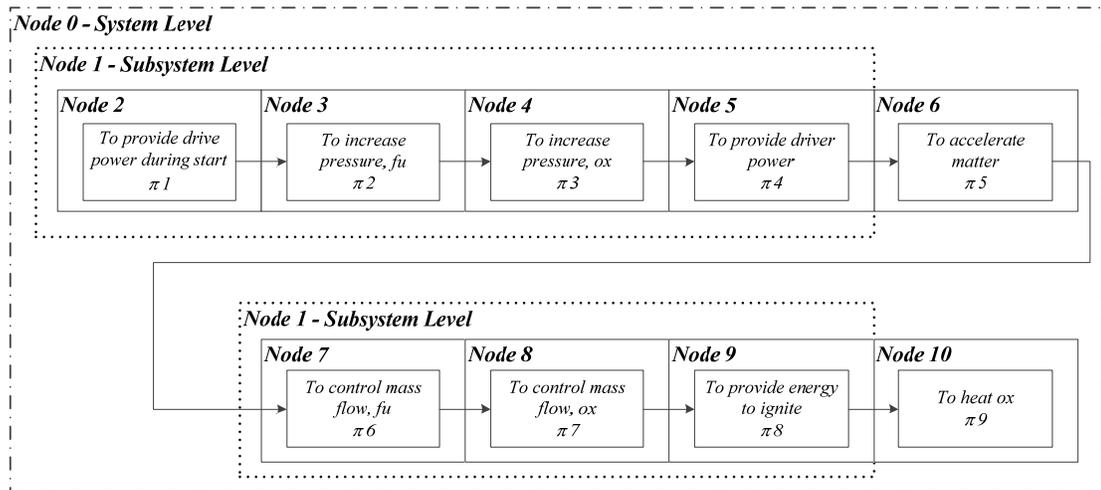
Figure 4-14: Functional Representation of the U.S. Liquid Rocket Engine RS-68

Determination of Knowledge Transfer Factor to predict the technical Program Risks

The product characteristics, namely vacuum thrust and combustion chamber pressure, are used to determine the system level knowledge transfer factor ϕ_{sys} for the RS-68 liquid rocket engine. The Space Shuttle Main Engine (SSME) is considered as the prior experience, and the knowledge transfer factor is set to 0.676. Table 4-27 lists the results. The predicted number of TAAF cycle failures (3) equals the number of main failure modes reported in Wood [118]: shortfall of turbopump power, fatigue life of turbine blisks, and damping of turbine blisks.

247

Table 4-27:   Knowledge Transfer Factor for the RS-68

	Thrust, kN	Pressure, bar	Propellants
SSME (old system)	2279	206.4	LOx/LH2
RS-68 (new system)	3370	97	LOx/LH2
Knowledge transfer factor		0.676	
Projected TAAF cycle failures		3	

Correlations among the Objectives

The correlation structures of the three objectives are determined by the specific models for the affordability, reliability, and IOC which will influence the behavior of the EMO algorithms [206].

In program management, the correlation among the objectives affordability and IOC is inherent (because more effort requires more time and costs more), and the consideration of joint confidence intervals for the assessment of the project budget and the associated project schedule is specifically requested by NASA [216]. In that context, Book [217] suggests an empirically determined correlation coefficient of 0.2, whereas Harmon [218] advocates a value of 0.45, which was derived using the Bayesian estimation method. If we assume that the joint distribution is normally distributed with an estimated correlation of 0.28, then Figure 4-15 depicts the joint confidence intervals at the 0.90, 0.95, and 0.99 confidence levels for the given U.S. liquid rocket engine RS-68 development program.
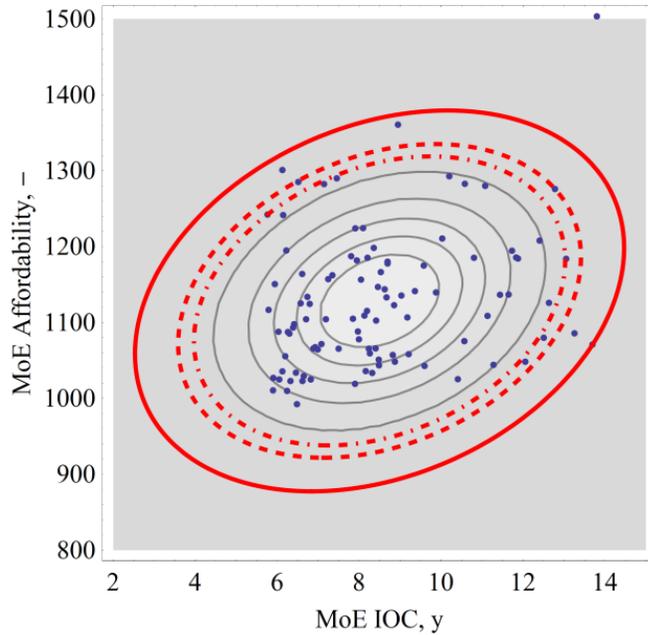
Figure 4-15: Joint Confidence Intervals

4.4.3 Satisficing Results considering Objective Weights, Decision-maker's Uncertainty, and Program Risks

Impact of Objective Weights on Satisficing Results

It is well-known that the weights of the composite fitness function influence the single optimal solution [155]. In addition, a SOGA will fail to find Pareto-optimal (non-dominated) solutions when the set of non-dominated solutions is non-convex [153-155]. Therefore, we performed the comparison of a SOGA using Palisade's Evolver® software [156] against the well-known and frequently used NSGA-II using the SolveXL® software [157] to show not only that the SOGA outperforms the NSGA-II but also that the set of non-dominated solutions is convex. Table 4-28 lists the parameter settings for the SOGA and NSGA-II. Noticeable is the run-time that is tenfold for the NSGA-II in comparison with the SOGA using the same computer hardware.

249

Table 4-28:   Parameters of the SOGA and NSGA-II

	Evolver®	NSGA-II
Population size	50	40
No. of generations	Progress based	12
Cross-over probability	0.5	0.5
Cross-over type	Arithmetic	Uniform random
Selector	Weighted average	Crowded tournament
Mutation probability	0.1	0.1
Mutator	Cauchy mutation	Simple by Gene
Utility function weights	Equally weighted	Not applicable
Run-time, h	2 to 3	35

The Pareto-optimal solutions of the NSGA-II are also used to show that the set of non-dominated solutions is convex and that the weights, as listed in Table 4-29, can be used to force the SOGA to explore this set, as depicted in Figure 4-16. Using the notation that is given in Table 4-29, Case-III is able to explore the upper bounds whereas Case-VII and Case-XI cover the lower bounds of the objective spaces. It also seems that the NSGA-II fails to converge in particular toward the Pareto-optimal front for Case-VI, Case-VIII, and Case-X which is either due to the correlated objectives or the result of both a small population size and a low number of generations.

Table 4-29:   Weights used to define the Fitness Function

|	Affordability w_1	Reliability w_2	IOC w_3
Case-I	1	0	0
Case-II	1/2	1/2	0
Case-III	0	1	0
Case-IV	2/3	1/6	1/6
Case-V	1/3	1/3	1/3
Case-VI	0	2/3	1/3
Case-VII	1/2	0	1/2
Case-VIII	0	1/2	1/2
Case-IX	1/3	0	2/3
Case-X	0	1/3	2/3
Case-XI	0	0	1
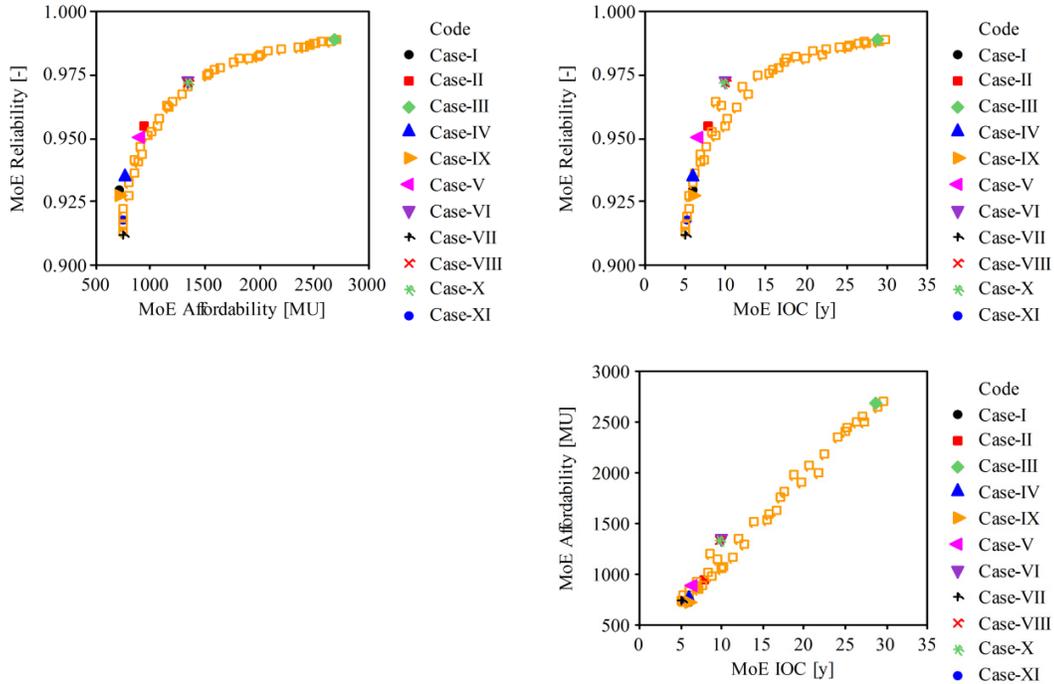
Figure 4-16: **The satisficed solutions found using the SOGA for the eleven cases (which use different weights) and the Pareto-optimal front found by the NSGA-II**

Impact of the Decision-maker's Risk Attitude

The decision-maker's risk attitude (expressed as a target) influences the utility (score value) for a given performance measure and the overall fitness evaluation of a solution. Studies have shown that the most sensitive risk-attitude is associated with the objective reliability (see Figure 3-2).

Historically, liquid rocket engines were hot-fire tested until an inherent reliability of 0.900 to 0.995 was demonstrated [117], and there was a tendency to target a level of around 0.956 for new liquid rocket engine developments (see Section 4.1). Therefore, the two reliability targets of 0.926 and 0.986 are selected to study the impact on the satisficing. The resulting effective risk coefficients are $\gamma^{Eff} = 5.8$ and $\gamma^{Eff} = -43.7$ characterizing the decision-maker's risk attitudes as risk averse and risk

251

seeking, respectively. (The original reliability target of 0.958 corresponds to a risk-neutral attitude.)   The weights from Case-V ($w_1 = w_2 = w_3 = \frac{1}{3}$) are used, and the SOGA for these two additional cases was run.   Figure 4-17 depicts the resulting single solutions for the three risk attitudes: risk-averse, risk-neutral (the original Case-V solution), and risk-seeking. Note that the other results from previous section are also included for ease of reference.
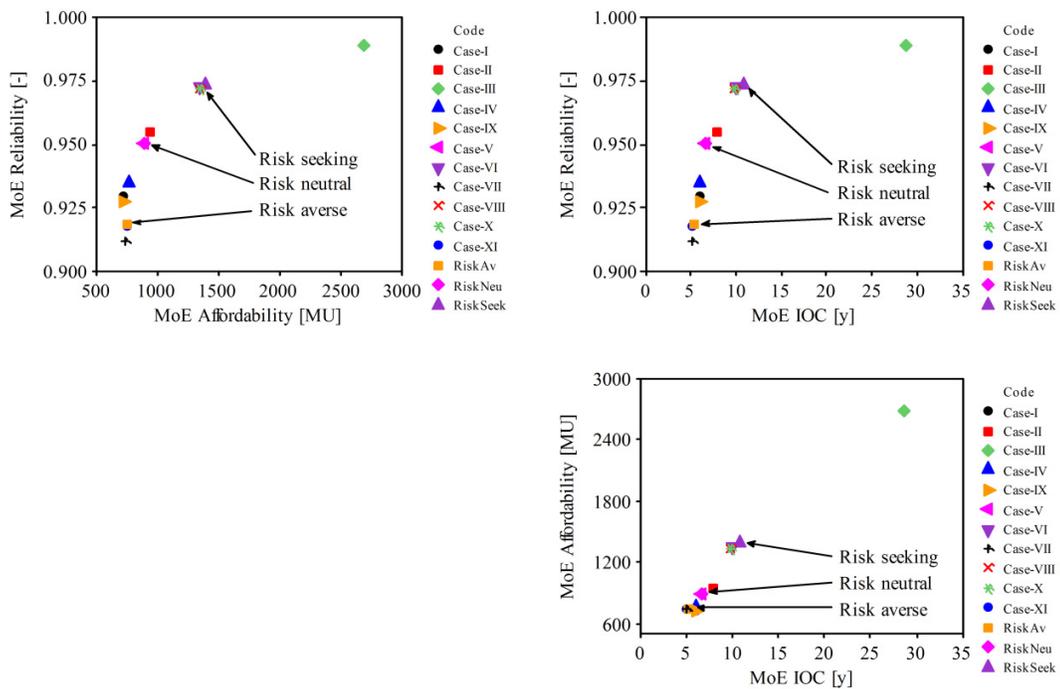
Figure 4-17:   Satisficed Solutions of the SOGA using different Risk Attitudes for Reliability

By comparing the satisficed solutions of the SOGA that were obtained with different weights with the solutions obtained with different risk attitudes for reliability, it is apparent that both decision-maker inputs influence the fitness evaluation of the SOGA. The impact of the risk attitude is even more influential; therefore, the determination of the weights is less critical in the satisficing approach than usually

considered. As a consequence, fixed values can be used in the overall decision-making. Instead, more focus should be given to the decision-maker's uncertainty about the objective spaces and the definition of credible objective targets.

Technical, Financial, and Schedule Program Risks

A Monte Carlo simulation was run to explore impact of this uncertainty on the performance of specific solutions. This simulation sampled from the distributions of these uncertain parameters, which are listed in Table 4-30. All were modeled as triangular distributions with the parameters given in Table 4-30. (Note that no test conductance variables or epistemic uncertainty is given to respect the confidentiality of the data such as the number of test facility, test cadence, and direct cost.) The simulation created 500 samples and, for each sample, calculated the performance measures for the satisficed solution that takes into account the decision-maker inputs (see Table 4-25). Figure 4-18 depicts the simulation results and the results generated in previous section for ease of reference.

Table 4-30:   Uncertainty Bounds of Decision-maker Inputs and Epistemic Uncertainty

	Min	Most likely	Max
Knowledge transfer factor, -	0.609	0.676	0.744
Failure occurrence allocation 1, -	0.54	0.60	0.66
Failure occurrence allocation 2, -	0.18	0.20	0.22
Failure mechanisms weight, -	0.4	0.5	0.6
Low cycle fatigue Weibull shape parameter	2	3	4
High cycle fatigue Weibull shape parameter	5	6	7
Life capacity weight, -	0.4	0.5	0.6
Design safety factor, -	2	4	6
Producibility, -	0.30	0.35	0.40
Overhead for development hardware, -	1.300	1.325	1.350
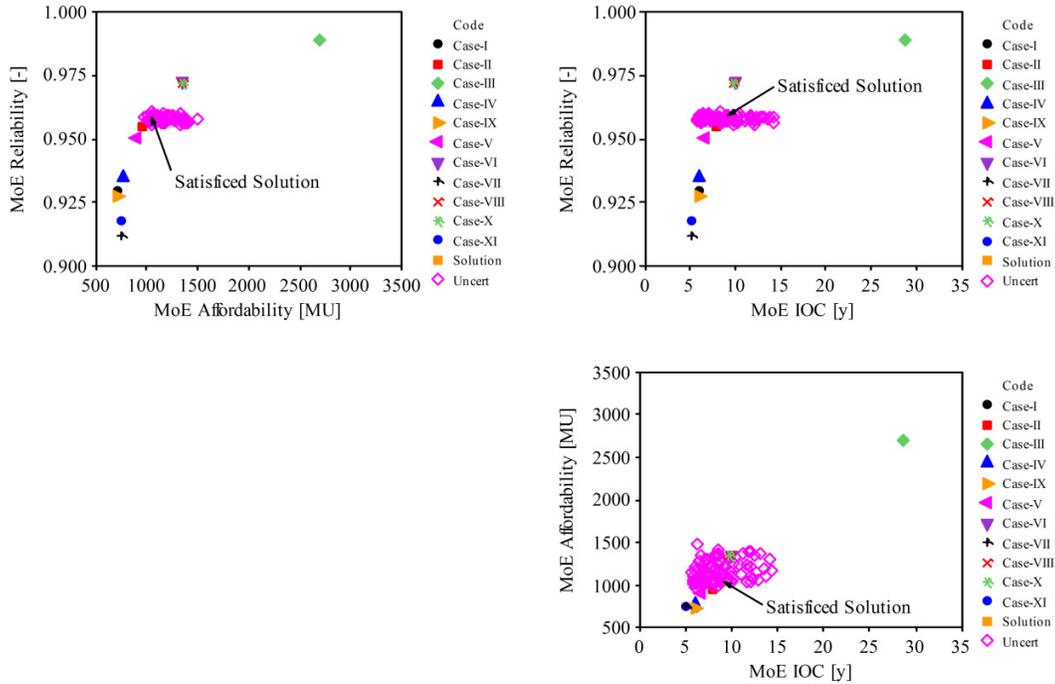
Figure 4-18: Risks and Epistemic Uncertainty Impact on the Satisfced Solution

Based on the simulation results for the financial and schedule risks (which are correlated as shown in Figure 4-15), the probability that the program will meet both the affordability and IOC targets (1.200 and 10.25 years) is approximately 0.65 based on the bivariate normal probability density function. Therefore, a 35 percent risk is given to accrue a cost overrun and a schedule slippage. In reality, the RS-68 liquid rocket engine development cost overrun was 40 percent, and its development schedule slippage was 12 months [118].

4.4.4 Conclusion

This section presented a preference-based risk-informed satisfced decision-making method that uses a SOGA and includes utility functions that reflect the decision-maker's risk attitude (expressed as targets). The SOGA implementation is shown to be computationally efficient and effective in finding the Pareto-optimal solutions by

comparing the results with the NSGA-II. It was also shown that the set of non-dominated solutions is convex, which allows the application of a SOGA.

The inclusion of the decision-maker's risk attitude into the fitness function by means of truncated exponential utility functions with associated efficient risk coefficient is shown to be more important than the weights in SOGAs or preference incorporation in the multi- or even many-objective EMO algorithms. The utility-probability duality is an adequate model that is easily implemented in a SOGA because it affects only the fitness evaluation.

The preference-based risk-informed satisficed decision-making method equips program managers and systems engineers with a simulation framework that is capable of treating program risks efficiently and adequately. The technical risk is measured by the number of TAAF cycle failures, and the financial and schedule risks are determined by the model variables and the epistemic uncertainty. Joint confidence intervals for the objectives affordability and IOC can be estimated to support this new trend in program risk management.

Chapter 5:   Conclusion

This dissertation described a risk-informed decision-making methodology to improve liquid rocket engine program tradeoffs with conflicting areas of concern, which includes non-technical and technical parameters. The solution strategy is based on a multiobjective satisficing problem formulation using the weighted sum of normalized objective functions. The objectives correspond to three areas of concern: affordability, reliability, and IOC, which are modeled with classical CERs, the RAIV strategy (introduced here), and classical SERs.

This dissertation also described the RAIV strategy, which is an important component of the methodology. The RAIV strategy was developed to estimate the demonstrated reliability of complex systems by aggregating multilevel hot-fire test data with different failure mechanisms and the characteristics that the testing profiles differ from the mission profile.

The problems that were discussed in Chapter 4 addressed: (1) the validation of the RAIV strategy using the U.S. liquid rocket engines F-1, SSME, and RS-68 and the European liquid rocket engine Vulcain 1 and Vulcain 2 (Section 4.1), (2) the application of the methodology in a multiattribute decision-making to select the best liquid rocket engine design alternative (Section 4.2), (3) the application of the methodology in a multiobjective satisficed decision-making to define the optimum hot-fire test plan (Sections 4.1, 4.2, and 4.4), (4) and the application of the RAIV strategy as a reliability growth model (Section 4.3).

All logical model constructions include evolutions of the tools that are used to find the solution to a formulated problem statement as it was the case also for the risk-informed decision-making methodology. The definition of the EQM and the Bayesian estimation of the functional node reliabilities were fundamental to the generation and validation of the logical model. The first major progress was the Bayesian estimation of the parameters that are used to calculate the weighting of the testing profiles instead of a simple quantile regression of historical data. The second major progress was the implementation of the coding trick [see Eq. (3.33)] that improved significantly the numerical stability and relaxed several impediments with regard to the use of a mixture prior distribution. The third main progress was the inclusion of the knowledge transfer factor as the mix parameters of the mixture prior distribution and the prediction of the TAAF cycle failures, which take into account the novelty of the new system. Consequently, the final model requires only minimum user inputs such as the targets for the objectives and the performance requirements of the liquid rocket engine alternatives to generate Pareto-optimal fronts or the satisficed solution for each of the liquid rocket engine design alternative.

5.1 Summary of Results

5.1.1 Reliability-as-an-independent-variable Strategy

The RAIV strategy provides a mathematical framework for planning and tracking the demonstrated reliability of complex systems by aggregating multilevel hot-fire test data with different failure mechanisms and the characteristics that the testing profiles differ from the mission profile. The planning of hot-fire test data includes the

prediction of the number of the typical TAAF cycle failures, which is based on the technology maturity of the competing risks system components.

The RAIV strategy is validated with the U.S. liquid rocket engines F-1, SSME, and RS-68 as well as the European liquid rocket engine Vulcain 1 and Vulcain 2 that were based on the different hot-fire test strategies ranging from a formal reliability demonstration, the DVS, and the objective-based variable test/time philosophy. It is shown that the three hot-fire test strategies are not different from a reliability engineering point of view. The differences are with regard to a stringent cost reduction approach by cutting the scope of the hot-fire test plan with the consequence of a reduced demonstrated reliability prior to the first flight.

5.1.2 Test plan optimization

Hot-fire test plan optimization, which maximizes the demonstrated reliability while optimizing the affordability and test schedule, is an important use of the risk-informed decision-making methodology. Therefore, the quantitative link between affordability and reliability is provided to the decision-maker.

As discussed in Section 4.2.3, the test plan optimization approach was used to quantitatively substantiate the claim that a prior test bed or demonstrator reduces the development cost of the actual flight engine in case there is a substantial technology maturation need.

5.1.3 Reliability Growth

The application of the RAIV strategy as a reliability growth planning, tracking, and projection model (discussed in Section 4.3) provides a true alternative to the

empirical Duane and analytical Crow/AMSAA models. In particular, the inclusion of testing profiles that are different to the mission profile and aggregated over several system integration levels offers specifically advantages over the classical reliability growth models.

5.1.4 Satisficing

The satisficing operation within the risk-informed decision-making methodology can be performed with a computationally efficient and effective SOGA because the set of non-dominated solutions is convex. This was shown by comparing the results of a SOGA with the well-known and frequently used NSGA-II. The SOGA approach combines the dimensionality reduction, preference incorporation, and different fitness evaluation schemes in order to handle the multiobjective problem in a single-objective problem formulation.

The dimensionality reduction is based on a weighted normalized fitness function that includes the decision-maker's risk attitude by means of truncated exponential utility functions (preference incorporation) with associated efficient risk coefficient using the utility-probability duality. The fitness function evaluation is then performed on the transformed objective space, i.e., the score values of the utility functions.

5.2 Contributions

The risk-informed decision-making methodology and the RAIV strategy contribute to improving decision-making in the liquid rocket engine industry by providing decision-makers with an integrated way to consider tradeoffs between demonstrated

reliability, affordability, and schedule (IOC). These tools can be used by customers (agency), program managers, systems engineers, and reliability engineers throughout the entire product life cycle.

The risk-informed decision-making methodology and the RAIV strategy improve the multiattribute decision-making in the NASA "pre-Phase A" or ECSS "Phase 0 and A" with regard to the selection of the best liquid rocket engine alternative by providing a quantitative link between the three areas of concern affordability, reliability, and IOC.

In addition, the risk-informed decision-making methodology and the RAIV strategy improve the multiobjective decision-making in the NASA "Phase A and B" or ECSS "Phase B" with regard to the definition of an optimized multilevel hot-fire test allocation that defines the overall test plan in order to achieve the liquid rocket engine flight certification with a stated reliability-by-credibility requirement.

Finally, the RAIV strategy is used for reliability growth modeling in all remaining product life cycle phases, i.e., NASA "Phase C, D, and E" or ECSS "Phase C, D, and E". Flight missions and production assurance tests are used as evidence.

5.3 Future Work

The risk-informed decision-making methodology is applied to liquid rocket engine systems that can be categorized as competing risks systems. The methodology is, however, generally formulated so that any other complex hardware system may be used that is subject to testing profiles that are different to the final mission profile. Therefore, future work could focus on the application of the risk-informed decision-

making methodology to other complex hardware system or non-competing risks systems.

Only expendable liquid rocket engines are currently considered. The inclusion of renewal theory and the application to maintenance models not only for the development but also for the operation and support is of particular interest for reusable liquid rocket engine applications. This research suggestion may seem odd with regard to the current launch vehicle development directions but the future will reintroduce reusable launch vehicles [219].

The TAAF cycle failure prediction is based on a system level approach and retrospective failure fraction allocation. The development of more sophisticated physics-of-failure component models may allow the definition of a component level knowledge transfer factor. By that means the TAAF cycle failure prediction may be improved.

The RAIV strategy application focuses on reliability growth planning and tracking. Future work could focus on the implementation of reliability growth projection models that incorporate the general framework of delayed and non-delayed fixes.

Software applications are not at all addressed in this research but the RAIV strategy is principally also applicable to software reliability verifications and validations by treating the multilevel as functions, modules, and fully integrated software instead of component, subsystem, and hardware product.

Bibliography

[1]   Hammond, W. E., *Space Transportation: A Systems Approach to Analysis and Design*, American Institute of Aeronautics and Astronautics, 1999.

[2]   Chang, I.-S., Investigation of Space Launch Vehicle Catastrophic Failures, *Journal of Spacecraft and Rockets*, Vol. 33, No. 2, 1996, pp. 198–205.

[3]   Huzel, D. K. and Huang, D. H., *Modern Engineering for Design of Liquid-Propellant Rocket Engines*, American Institute of Aeronautics and Astronautics, 1992.

[4]   Emdee, J. L., A Survey of Development Test Programs for Hydrogen Oxygen Rocket Engines, No. 2001-0749, AIAA paper, 2001.

[5]   Emdee, J. L., A Survey of Development Test Programs for LOx/Kerosene Rocket Engines, No. 2001-3985, AIAA paper, 2001.

[6]   Joyner, C. R., Laurie, J. R., Levack, D. J., and Zapata, E., Pros, Cons, and Alternatives to Weight Based Cost Estimating, No. 2011-5502, AIAA paper, 2011.

[7]   Hunt, C. D., Personal Communication, May 2010.

[8]   Hamaker, J. W., *Improving the predictive Capability of Spacecraft Cost Models by Considering Non-technical Variables*, Ph.D. thesis, University of Alabama, Huntsville, AL, USA, 2006, Department of Industrial and Systems Engineering.

[9]   Richards, S., Liquid Rocket Engine Flight Certification, *Space Transportation Propulsion Technology Symposium*, Vol. 3, 1990, pp. 1031–1038.

[10]   Meisl, C. J., Life Cycle Cost Considerations for Launch Vehicle Liquid Propellant Rocket Engines, No. 86-1408, AIAA paper, 1986.

[11]   Meisl, C. J., Advanced Life Cycle Cost Modeling, *Journal of Parametrics*, Vol. 4, No. 1, 1989, pp. 74–102.

[12]   Meisl, C. J., Life Cycle Cost Methodology for Space Rocket Engines, *International Society of Parametric Analysts*, 1991, 13th Annual Conference.

[13]   SAIC, NAFCOM Government Version 2008, 2008.

[14]   Yang, G., *Life Cycle Reliability Engineering*, John Wiley & Sons, 2007.

[15]   Kumamoto, H., *Satisfying Safety Goals by Probabilistic Risk Assessment*, Springer Science+Business Media, 2007.

[16]   Sackheim, R., Overview of United States Space Propulsion Technology and Associated Space Transportation Systems, *Journal of Propulsion and Power*, Vol. 22, No. 6, 2006, pp. 1310–1333.

[17] Hamada, M. S., Martz, H. F., Reese, C. S., Graves, T., Johnson, J., and Wilson, A. G., A fully Bayesian Approach for combining multilevel failure information in Fault Tree Quantification and optimal follow-on Resource Allocation, *Reliability Engineering and System Safety*, 2004, pp. 297–305.

[18] Hamada, M. S., Wilson, A. G., Reese, C. S., and Martz, H. F., *Bayesian Reliability*, Springer Science+Business Media, 2008.

[19] Graves, T. L. and Hamada, M. S., A Demonstration of Modern Bayesian Methods for Assessing System Reliability with Multilevel Data and for Allocating Resources, *International Journal of Quality, Statistics, and Reliability*, 2009, pp. 1–10.

[20] Practices, J. L. P. S. T. and Panel, S., DRAFT: Test and Evaluation Guideline for Liquid Rocket Engines, Tech. Rep. 50260651, 2010.

[21] Pempie, P. and Vernin, H., Liquid Rocket Engine Test Plan Comparison, No. 2001-3256, AIAA paper, 2001.

[22] SAE, Liquid Rocket Engine Reliability Certification, Tech. Rep. ARP 4900, The Engineering Society for Advancing Mobility Land Sea Air and Space, 1996.

[23] ESA, ESA - Launchers Home - Next Generation Launchers, 2008, retrieved on January 05, 2010.

[24] Simon, H. A., A Behavioral Model of Rational Choice, *The Quarterly Journal of Economics*, Vol. 69, No. 1, 1955, pp. 99–118.

[25] Simon, H. A., Rational Choice and the Structure of the Environment, *Psychological Review*, 1956, pp. 129–138.

[26] Simon, H. A., Theories of Decision-making in Economics and Behavioral Science, *The American Economic Review*, 1959, pp. 253–283.

[27] Strunz, R. and Herrmann, J. W., Reliability As an Independent Variable Applied to Liquid Rocket Hot Fire Test Plans, *AIAA Journal of Propulsion and Power*, Vol. 27, No. 5, 2011, pp. 1032–1044.

[28] Strunz, R. and Herrmann, J. W., Planning, Tracking, and Projecting Reliability Growth A Bayesian Approach, *IEEE Conference Publications*, 2012.

[29] Strunz, R. and Herrmann, J. W., A reliability as an independent variable (RAIV) methodology for optimizing test planning for liquid rocket engines, *CEAS Space Journal*, 2011.

[30] Williams, W. C., Report of the SSME Assessment Report, Tech. rep., Washington D.C.: National Aeronautics and Space Administration, 1993.

[31] Krasich, K., Reliability Growth Test Design – Connecting Math to Physics, The Annual Reliability and Maintainability Symposium, 2011.

[32] Hill, P. G. and Peterson, C. R., *Mechanics and Thermodynamics of Propulsion*, Addison-Wesley, 1992.

[33] Humble, R. W., Henry, G. N., and Larson, W. J., *Space Propulsion Analysis and Design*, The McGraw-Hill Companies, 1995.

[34] Sutton, G. P. and & Biblarz, O., *Rocket Propulsion Elements*, John Wiley & Sons, 2001.

[35] Howard, R. A., Decision Analysis: Applied Decision Theory, *Proceedings of the Fourth International Conference on Operational Research*, 1966, pp. 55–77.

[36] Daft, R. L., *Organization Theory and Design*, South-Western College Pub, 2009.

[37] Cyert, R. M. and March, J. G., *A behavioral Theory of the Firm*, Prentice-Hall Inc, 1963.

[38] Manktelow, K., *Reasoning and Thinking*, Psychology Press, 2000.

[39] Gilboa, I., Decision Theory, Entry in the Encyclopedia of Political Science.

[40] Mintzberg, H., Raisinghani, D., and Théorêt, A., The Structure of "Unstructured" Decision Processes, *Administrative Science Quarterly*, Vol. 21, No. 2, 1976, pp. 246–275.

[41] Soelberg, P. O., Unprogrammed Decision Making, *Industrial Management Review*, 1967, pp. 19–29.

[42] Meisl, C. J., Space Transportation Booster Engine Selection Methodology, No. 87-1852, AIAA paper, 1987.

[43] Krevor, Z. C., *A Methodology to Link Cost and Reliability for Launch Vehicle Design*, Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2007, School of Aerospace Engineering.

[44] Bush, G. W., President Bush announces new Vision for Space Exploration Program, 2004, Retrieved on July 26, 2006.

[45] Bernoulli, D., Exposition of a New Theory on the Measurement of Risk, *Econometrica*, 1738, pp. 22–36.

[46] von Neumann, J. and Morgenstern, O., *Theory of games and economic behavior*, Princeton University Press, 1947.

[47] Savage, L. J., *The foundations of statistics*, Wiley, 1954.

[48] Kahneman, D. and Tversky, A., Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 1979, pp. 263–291.

[49] Durlauf, S. and Blume, L., *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, 2008.

[50] Tversky, A. and Kahneman, D., Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, Vol. 5, 1992, pp. 297–323.

[51] Birnbaum, M. H., New Paradoxes of Risky Decision Making, *Psychological Review*, Vol. 115, No. 2, 2008, pp. 463–501.

[52]   Sewell, M., *Decision Making Under Risk: A Prescriptive Approach*, University of Cambridge, 2009.

[53]   Harrison, G. W. and Rutström, E. E., Expected Utility Theory and Prospect Theory: One Wedding and a decent Funeral, *Experimental Economics*, Vol. 12, 2009, pp. 133–158.

[54]   Bordley, R. and LiCalzi, M., Decision Analysis using Targets instead of Utility Functions, *Decisions in Economics and Finance*, Vol. 23, 2000, pp. 53–74.

[55]   Abbas, A. E. and Matheson, J., Normative target-based decision making, *Managerial and Decision Economics*, Vol. 26, 2005, pp. 373–385.

[56]   Abbas, A. E. and Matheson, J., Utility - Probability Duality, Stanford University, Department of Management Science and Engineering.

[57]   Wilson, P., *The Utility of Reliability and Survival*, Ph.D. thesis, The George Washington University, 2007.

[58]   Odhnoff, J., On the Techniques of Optimizing and Satisficing, *The Swedish Journal of Economics*, 1965, pp. 24–39.

[59]   Eilon, S., Goals and Constraints in Decision-making, *Operational Research Quarterly*, 1972, pp. 3–15.

[60]   Wierzbicki, A. P., A Mathematical Basis for Satisficing Decision Making, *Mathematical Modelling*, Vol. 3, No. 5, 1982, pp. 391–405.

[61]   Rao, S. S., *Engineering Optimization: Theory and Practice*, John Wiley & Sons, 2009.

[62]   Weise, T., Global Optimization Algorithms - Theory and Application, self-published as e-book, 2009.

[63]   Gandomkar, M. and Vakilian, M., A Combination of Genetic Algorithm and Simulated Annealing for optimal Distributed Generation Allocation in Distribution Networks, *18th Annual Canadian Conference on Electrical and Computer Engineering*, 2005, pp. 645–648.

[64]   Kuo, W. and Wan, R., *Optimal Reliability Design – Algorithms and Comparisons*, John Wiley & Sons, 2007.

[65]   Nahas, N. and Nourelfath, M., Ant System for Reliability Optimization of a series System with Multiple-choice and Budget Constraints, *Reliability Engineering and System Safety*, 2005, pp. 1–12.

[66]   Tao, Y., Huang, H.-Z., and Yang, B., An Interactive Preference-Weight Genetic Algorithm for Multi-criterion Satisficing Optimization, *International Conference on Natural Computation*, 2006, pp. 643–652.

[67]   Tamura, H., Shibata, T., Tomiyama, S., and Hatono, I., A Meta-Heuristic Satisficing Trade-off Method for Solving Multi-objective Combinatorial Optimization Problems, *International Conference on Systems, Man and Cybernetics*, 1999, pp. 539–544.

[68]   Bayes, T. and Price, R., An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions of the Royal Society of London*, Vol. 53, 1763, pp. 370–418.

[69]   Miller, I. and Miller, M., *John E. Freund's Mathematical Statistics*, Pearson Education, 2004.

[70]   Martz, H. F. and Waller, R. A., The Basics of Bayesian Reliability Estimation from Attribute Test Data, Tech. rep., Los Alamos Scientific Laboratory of the University of California, 1976.

[71]   Modarres, M., Kaminskiy, M., and Krivtsov, V., *Reliability Engineering and Risk Analysis: A Practical Guide*, CRC Press, 2010.

[72]   Casella, G. and Berger, R. L., *Statistical Inference*, Duxbury, 2002.

[73]   Kass, R. E. and Wasserman, L., The Selection of Prior Distributions by Formal Rules, *Journal of the American Statistical Association*, Vol. 91, No. 435, 1996, pp. 1343–1370.

[74]   Robert, C. P., *The Bayesian Choice*, Springer Science+Business Media, 2007.

[75]   Weiler, H., The Use of Incomplete Beta Functions for Prior Distributions in Binomial Sampling, *Technometrics*, 1965, pp. 335–347.

[76]   Pham, T. G., A Bayes Sensitivity Analysis when using the Beta Distribution as a Prior, *IEEE Transactions on Reliability*, Vol. 39, No. 1, 1990, pp. 6–7.

[77]   Duran, B. S. and Booker, J. M., A Bayes Sensitivity Analysis when Using the Beta Distribution as a Prior, *IEEE Transactions on Reliability*, 1988, pp. 239–247.

[78]   Siu, N. O. and Kelly, D. L., Bayesian Parameter Estimation in Probabilistic Risk Assessment, *Reliability Engineering and System Safety*, 1998, pp. 89–116.

[79]   Waterman, M. S., Martz, H. F., and Waller, R. A., Fitting Beta Prior Distributions in Bayesian Reliability Analysis, Tech. rep., Los Alamos Scientific Laboratory of the University of California, 1976.

[80]   McFadden, R. H. and Shen, Y., An Analysis of the historical Reliability of US liquid-fuel Propulsion Systems, No. 90-2713, AIAA paper, 1990.

[81]   Krolo, A., *Planung von Zuverlässigkeitstests mit weitreichender Berücksichtigung von Vorkenntnissen*, Ph.D. thesis, University of Stuttgart, 2004, Mechanical Engineering Department.

[82]   Kleyner, A. V., *Determining optimal Reliability Targets through Analysis of Product Validation Cost and Field Warranty Data*, Ph.D. thesis, University of Maryland, College Park, MD, USA, 2005, Mechanical Engineering Department.

[83]   Hitziger, T., *Übertragbarkeit von Vorkenntnissen bei der Zuverlässigkeitstestplanung*, Ph.D. thesis, University of Stuttgart, 2007, Mechanical Engineering Department.

[84]   Giffin, W. C., *Transform Techniques for Probability Modeling*, Academic Press, 1975.

[85]   Mastran, D. V., *On Bayesian Assessments of Coherent Systems*, Ph.D. thesis, George Washington University, 1973, School of Engineering and Applied Science.

[86]   Springer, M. D. and Thompson, W. E., Bayesian Confidence Limits for the Reliability of Cascade Exponential Subsystems, *IEEE Transaction on Reliability*, 1967, pp. 86–89.

[87]   Springer, M. D. and Byers, J. K., Bayesian Confidence Limits for the Reliability of Mixed Exponential and Distribution-Free Cascade Subsystems, *IEEE Transactions on Reliability*, 1971, pp. 24–28.

[88]   Martz, H. F. and Duran, B., A Comparison of three Methods for Calculating Lower Confidence Limits on System Reliability using Binomial Component Data, *IEEE Transactions on Reliability*, 1985, pp. 113–120.

[89]   Martz, H. F. and Waller, R. A., Bayesian Reliability Analysis of complex series/parallel Systems of Binomial Subsystems and Components, *Technometrics*, 1990, pp. 407–416.

[90]   Martz, H. F., Waller, R. A., and Fickas, E. T., Bayesian Reliability Analysis of Series Systems of Binomial Subsystems and Components, *Technometrics*, 1988, pp. 143–154.

[91]   Bier, V. M., On the Concept of Perfect Aggregation in Bayesian Estimation, *Reliability Engineering and Safety*, 1994, pp. 271–281.

[92]   Azaiez, M. N. and Bier, V. M., Perfect Aggregation for a Class of General Reliability Models with Bayesian Updating, *Applied Mathematics and Computation*, 1995, pp. 281–302.

[93]   Johnson, V. E., Graves, T. L., Hamada, M., and Reese, C. S., A Hierarchical Model for Estimating the Reliability of Complex Systems, *Bayesian Statistics 7*, 2003, pp. 199–213.

[94]   Thompson, W. E. and Chang, E., Bayes Confidence Limits for Reliability of Redundant Systems, *Technometrics*, 1975, pp. 127–146.

[95]   Chang, E. Y. and Thompson, W. E., Bayes Analysis of Reliability of Complex Systems, *Operation Research*, 1976, pp. 156–168.

[96]   Martz, H. F. and Baggerly, K. A., Bayesian Reliability Analysis of High-Reliability Systems of Binomial and Poisson Subsystems, *International Journal of Reliability, Quality, and Safety Engineering*, 1997, pp. 283–307.

[97]   Bolstad, W. M., *Understanding Computational Bayesian Statistics*, John Wiley & Sons, 2010.

[98]   Robert, C. P. and Casella, G., *Monte Carlo Statistical Methods*, Springer Science-Business Media, 2004.

[99]   Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., Equation of State Calculations by Fast Computing Machines, *The Journal of Chemical Physics*, 1953, pp. 1087–1092.

[100]   Hastings, W. K., Monte Carlo Sampling Methods using Markov chains and their Applications, *Biometrika*, 1970, pp. 97–109.

[101]   Held, L., *Methoden der statistischen Inferenz: Likelihood und Bayes*, Spektrum Akademischer Verlag Heidelberg, 2008.

[102]   Albert, J., *Bayesian Computation with R*, Springer Science+Business Media., 2007.

[103]   Yu, B. and Mykland, P., Looking at Markov samplers through cusum path plots: A simple diagnostic idea, *Statistics and Computing*, 1998, pp. 275–286.

[104]   Gelman, A. and Rubin, D., Inference from iterative simulation using multiple sequences (with discussion), *Statistical Science*, 1992, pp. 457–511.

[105]   Cowles, M. K. and Carlin, B. P., Markov chain Monte Carlo Convergence Diagnostics: A Comparative Review, *Journal of the American Statistical Association*, 1996, pp. 883–904.

[106]   Brooks, S. P. and Roberts, G. O., Convergence assessment techniques for Markov chain Monte Carlo, *Statistics and Computing*, 1998, pp. 319–335.

[107]   Gregory, P., *Bayesian Logical Data Analysis for the Physical Sciences*, Cambridge University Press, 2005.

[108]   Liu, J. S., *Monte Carlo Strategies in Scientific Computing*, Springer Science+Business Media, 2008.

[109]   Graves, T. L., YADAS Yet Another Data Analysis Software, 2003, retrieved on August 30, 2011.

[110]   Gelman, A. and Hill, J., *Data Analysis using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, 2008.

[111]   NASA, NASA Systems Engineering Handbook, Tech. Rep. NASA/SP-2007-6105 Rev. 1, 2007.

[112]   ECSS, Space Project Management: Project planning and Implementation, 2009.

[113]   Hammond, W. E., *Design Methodologies for Space Transportation Systems*, American Institute of Aeronautics and Astronautics, 2001.

[114]   Kraemer, R. S., *Rocketdyne: Powering Humans into Space*, American Institute of Aeronautics and Astronautics, 2006.

[115]   Wasserman, G. S., *Reliability Verification, Testing, and Analysis in Engineering Design*, Marcel Dekker, Inc, 2003.

[116]   Schwarz, W., Schwub, S., Quering, K., Wiedmann, D., Höppel, H. W., and Göken, M., Life prediction of thermally highly loaded components: modelling the damage process of a rocket combustion chamber hot wall, *CEAS Space Journal*, 2010, pp. 1–15.

[117]   Koelle, D. E., *Handbook of Cost Engineering for Space Transportation Systems with TRANSCOST 7.3 Rev 3*, TCS-TransCostSystems, 2010.

[118]   Wood, B. K., Propulsion for the 21st Century – RS-68, No. 2002-4324, AIAA paper, 2002.

[119]   Greene, W., Development of the J-2X Engine for the ARES I Crew Launch Vehicle and the ARES V Cargo Launch Vehicle: Building on the APOLLO Program for Lunar Return Missions, 2008, retrieved on February 07, 2010.

[120]   Biggs, R. E., *Space Shuttle Main Engine: The first twenty years and beyond*, American Astronautical Society, 2008.

[121]   Rocketdyne, Background, F-1 Rocket Engine, 1965, retrieved on May 15, 2011.

[122]   O'Hara, K. J., Liquid Propulsion System Reliability "Design for Reliability", No. 89-2628, AIAA paper, 1989.

[123]   Pugh, R. L., Space Transportation Main Engine Reliability Demonstration Technique, *Proceedings Annual Reliability and Maintainability Symposium*, 1993, pp. 173–180.

[124]   Lloyd, D. K. and Lipow, M., *Reliability: Management, Methods, and Mathematics*, The American Society for Quality Control, 1977.

[125]   Worlund, A. L., Monk, J. C., and Bachtel, F. D., NLS Propulsion Design Considerations, No. 92-1181, AIAA paper, 1992.

[126]   Butts, G., Mega Projects Estimates - A History of Denial, NASA, 7th Annual NASA Project Management Challenge, 2010.

[127]   NASA, NASA Cost Estimating Handbook, Tech. rep., NASA, 2008.

[128]   Kumar, D. U., Crocker, J., Chitra, T., and Saranga, H., *Reliability and Six Sigma*, Springer Science+Business Media, 2006.

[129]   Dhillon, B. S., *Life Cycle Costing*, Gordon and Breach Science Publishers, 1989.

[130]   Fabrycky, W. J. and Blanchard, B. S., *Life-Cycle Cost and Economic Analysis*, Prentice-Hall, 1991.

[131]   Blanchard, B. S. and Fabrycky, W. J., *Systems Engineering and Analysis*, Pearson Prentice Hall, 2006.

[132]   Gupta, Y. and Chow, W. S., Twenty-five years of life cycle costing - theory and applications: a survey, *The International Journal of Quality and Reliability*, 1985, pp. 51–76.

[133]   Asiedu, Y. and Gu, P., Product Life Cycle Cost Analysis: State-of-the-art Review, *International Journal of Production Research*, Vol. 36, No. 4, 1998, pp. 883–908.

[134]   Christensen, P. N., Sparks, G. A., and Kostuk, K. J., A method-based Survey of Life Cycle Costing Literature pertinent to Infrastructure Design and Renewal, *Canadian Journal of Civil Engineering*, Vol. 32, 2005, pp. 250–259.

[135]   Strunz, R., *An integrated View of Project Management and Systems Engineering Principles*, Master's thesis, Munich University of Applied Sciences, 2002.

[136]   Stewart, R. D., *Cost Estimating*, John Wiley & Sons, 1991.

[137]   Ventura, M., The Lowest Cost Rocket Propulsion System, No. 2006-4782, AIAA paper, 2006.

[138]   Emblemsvåg, J., *Life-Cycle Costing: Using Activity Based Costing and Monte Carlo Methods to manage future Costs and Risks*, John Wiley & Sons, 2003.

[139]   Strunz, R., *An Extensive Case Study of Design of Experiments applied to the Characterization of a Parametric Cost Estimation Model*, Master's thesis, Munich University of Applied Sciences, 2005.

[140]   Meisl, C. J., The Future of Design Integrated Cost Modeling, No. 92-1056, AIAA paper, 1992.

[141]   Lee, P., A Process Oriented Parametric Cost Model, No. 93-1029, AIAA paper, 1993.

[142]   PRICESystems, *Your Guide to PRICE H*, PRICESystems, 1998.

[143]   Galorath, SEER for Manufacturing: Project Management Software, Project Tracking Software, Manufacturing Estimation Software - Galorath, 2008, retrieved on January 24, 2010.

[144]   SAIC, *Cost and Schedule Analytical Techniques Development*, SAIC, 1998.

[145]   Harwick, W. T., Space Economics Measurement: A Survey of Space Cost Models, No. 97-3945, AIAA paper, 1997.

[146]   Broemm, W. J., Ellner, P. M., and Woodworth, W. J., AMSAA Reliability Growth Guide, Tech. rep., United States Army Material Systems Analysis Activity, Aberdeen Proving Ground, 2000.

[147]   Codier, E. O., Reliability Growth in Real Life, Vol. In Proceedings of the 1968 Annual Symposium on Reliability, 1968, pp. 458–469.

[148]   Hall, J. B., *Methodology for Evaluating Reliability Growth Programs of discrete Systems*, Ph.D. thesis, University of Maryland, College Park, MD, USA, 2008, Department of Mechanical Engineering.

[149]   Crow, L. H., A Methodology for Managing Reliability Growth during Operational Mission Profile Testing, Annual Reliability and Maintainability Symposium, IEEE, 2008.

[150]   Safie, F. and Fuller, R. P., NASA Applications and Lessons Learned in Reliability Engineering, The Annual Reliability and Maintainability Symposium, 2012.

[151]   McDonnell, B. J., The Application of a Design Verification System and Accelerated Mission Testing to Gas Turbine Engine Development, Tech. rep., Society of Automotive Engineers, 1978.

[152]   Hajela, P. and Lin, C., Genetic search strategies in multi-criterion optimal design, *Structural Optimization*, Vol. 4, No. 2, 1993, pp. 99–107.

[153]   Deb, K., *Multi-objective Optimization using Evolutionary Algorithms*, John Wiley & Sons, 2001.

[154]   Miettinen, K., *Nonlinear Multiobjective Optimization*, Kluwer, 1999.

[155]   Zhang, Q., MOEA/D: A Multiobjective Evolutionary Algorithm based on Decomposition, *IEEE Transaction on Evolutionary Computation*, Vol. 11, No. 6, 2007, pp. 712–731.

[156]   Palisade, DecisionTools Suite (5.7) [Software], 2011.

[157]   LLP, E. A. A., SolveXL (Version 1.0.5.2) [Software], 2011.

[158]   Saaty, T. L., *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*, RWS Publications, 2000.

[159]   Mustajoki, J., Hämäläinen, R., and Salo, A., Decision Support by Interval SMART/SWING - Incorporating Imprecision in the SMART and SWING Methods, *Decision Sciences*, Vol. 36, No. 2, 2005, pp. 317–339.

[160]   Cook, J. D., Determining distribution parameters from quantiles, Working paper 55, Anderson Cancer Center Department of Biostatistics Working Paper Series, 2010.

[161]   NATO, Methods and Models for Life Cycle Costing, Tech. Rep. TR-SAS-054, NATO, 2007.

[162]   Krugman, P. R. and Obstfeld, *International Economics*, Addison Wesley, 2002.

[163]   Schankman, M. S. and Reynolds, J. W., Advancing the Art of Technology Cost Estimating- a Collaboration between NASA and Boeing, *ISPA/SCEA International Conference*, 2010.

[164]   OSD, *Manufacturing Readiness Level Deskbook*, DoD, version 2 ed., 2011.

[165]   Macret, J. L., Maitrise des Risques des Développements Technologiques: Méthodologie TRL et IRL en Ingénierie Système, 2008.

[166]   Mankins, J. C., Technology Readiness Levels, Tech. rep., NASA, 1995.

[167]   Sauser, B. J. and Forbes, E., Defining an Integration Readiness Level for Defense Acquisition, *International Symposium of the International Council on Systems Engineering (INCOSE)*, 2009.

[168]   Long, J. M., Integration readiness levels, *Aerospace Conference, 2011 IEEE*, 2011, pp. 1–9.

[169]   Nelson, W., *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*, John Wiley & Sons, 1990.

[170]   Rausand, M. and Hoyland, A., *System Reliability Theory*, John Wiley & Sons, 2004.

[171]   Martz, H. F. and Waller, R. A., *Bayesian Reliability Analysis*, John Wiley & Sons, 1982.

[172]   Feynman, R., *What Do You Care What Other People Think?*, W. W. Norton & Company, 2001.

[173]   Leemis, L. M., *Reliability Probabilistic Models and Statistical Methods*, Lawrence Leemis, 2009.

[174]   Rocketdyne, Rocket Engine Life Analysis, Task Final Report, DR-4, 1996.

[175]   McCulloch, C., Searle, S., and Neuhaus, J., *Generalized, Linear, and Mixed Models*, John Wiley & Sons, 2008.

[176]   Givens, G. and Hoeting, J., *Computational Statistics*, John Wiley & Sons, 2012.

[177]   Ebeling, C. E., *An Introduction to Reliability and Maintainability Engineering*, Waveland Press, 2005.

[178]   Young, A., *The Saturn V F-1 Engine*, Springer, 2009.

[179]   Anderson, M. J. and Whitcomb, P. J., *RSM Simplified: Optimizing Processes using Response Surface Methods for Design of Experiments*, Productivity Press, Ltd., 2005.

[180]   Kleijnen, J. P. C., *Design and Analysis of Simulation Experiments*, Springer, 2008.

[181]   Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to linear regression analysis*, Wiley & Sons, 2006.

[182]   Rahman, S. A. and Hebert, B. J., Large Liquid Rocket Testing – Strategies and Challenges, No. 2005-3564, AIAA paper, 2005.

[183]   Kelly, D. and Smith, C., *Bayesian Inference for Probabilistic Risk Assessment*, Springer Series in Reliability Engineering, Springer Science+Business Media, 2011.

[184]   Fabrycky, W. J., Ghare, P. M., and Torgersen, P. E., *Industrial Operations Research*, Prentice-Hall, 1972.

[185]   Ramsey, F. L. and Schafer, D. W., *The Statistical Sleuth*, Duxbury, 2002.

[186] Strunz, R. and Herrmann, J. W., RAIV - The Sibling of CAIV, European Aerospace Cost Engineering Working Group (EACE), 2011.

[187] SatNews, 2010, retrieved on October 30, 2010.

[188] Parkinson, R. C., The Hidden Costs of Reliability and Failure in Launch Systems, *Acta Astronautica*, Vol. 44, No. 7-12, 1999, pp. 419–424.

[189] Rachuk, V. S., Design, Development, and History of the Oxygen/Hydrogen Engine RD-0120, No. 1995-2540, AIAA paper, 1995.

[190] Wilson, A. G., Graves, T. L., Hamada, M. S., and Reese, C. S., Advances in Data Combination, Analysis and Collection for System Reliability Assessment, *Statistical Science*, 2006, pp. 514–534.

[191] Hahn, T., CUBA: A Library for Multidimensional Numerical Integration, *Computer Physics Communications*, Vol. 168, No. 2, 2005, pp. 78–95.

[192] Chib, S. and Greenberg, E., Understanding the Metropolis-Hastings Algorithm, *American Statistician*, Vol. 49, No. 4, 1995, pp. 327–335.

[193] Neath, R. and Jones, G., Variable-at-a-Time Implementations of Metropolis-Hastings, 2009, retrieved on July 15, 2011.

[194] Fishman, G. S., *A First Course in Monte Carlo*, Duxbury, 2006.

[195] Goldberg, D., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.

[196] Gunter, D., Gunter's Space Page, 2011, Retrieved on August 26, 2011.

[197] ESA, Next Generation Launchers, 2010, retrieved on October 30, 2010.

[198] Fukushima, Y., Development status of LE-7A and LE-5B engines for H-IIA Family, *Acta Astronautica*, Vol. 50, No. 5, 2002, pp. 275–284.

[199] Shermon, D., *Systems Cost Engineering*, Gower Publishing, 2009.

[200] Kirkwood, C. W., *Strategic Decision Making*, Duxbury Press, 1995.

[201] McPherson, J. W., *Reliability Physics and Engineering*, Springer Science+Business Media, 2010.

[202] Saunders, S. C., *Reliability, Life Testing, and Prediction of Service Lifes*, Springer Science+Business Media, 2007.

[203] Branke, J., Kaußler, T., and Schmeck, H., Guidance in Evolutionary Multi-Objective Optimization, *Advances in Engineering Software*, Vol. 32, 2001, pp. 499–507.

[204] Branke, J. and Deb, K., *Knowledge Incorporation in Evolutionary Computation*, Springer, 2005.

[205] Deb, K., Mohan, M., and S., M., Evaluating the eps-Domination Based Multi-Objective Evolutionary Algorithm for a Quick Computation of Pareto-Optimal Solutions, *Evolutionary Computation*, Vol. 13, No. 4, 2005, pp. 501–525.

[206]   Ishibuchi, H., Hitotsuyanagi, Y., Ohyanagi, H., and Nojima, Y., Effect of the Existence of Highly Correlated Objectives on the Behavior of MOEA/D, *Lecture Notes in Computer Science 6576: Evolutionary Multi-Criterion Optimization*, Vol. EMO 2011, 2011, pp. 166–181.

[207]   Ishibuchi, H., Hitotsuyanagi, Y., Ohyanagi, H., and Nojima, Y., Behavior of EMO algorithms on many-objective optimization problems with correlated objectives, *Proc. of 2011 IEEE Congress on Evolutionary Computation*, 2011, pp. 1465–1472.

[208]   Jin, Y. and Branke, J., Evolutionary Optimization in Uncertain Environment - A Survey, *IEEE Transactions on Evolutionary Computation*, Vol. 9, No. 3, 2005, pp. 303–317.

[209]   Trautmann, H. and Mehnen, J., Preference-based Pareto optimization in certain and noisy environments, *Engineering Optimization*, Vol. 41, No. 1, 2009, pp. 23–38.

[210]   Goh, C. K. and Tan, K. C., Noise Handling in Evolutionary Multi-Objective Optimization, *IEEE Congress on Evolutionary Computation*, 2006.

[211]   Hughes, E. J., Evolutionary multi-objective ranking with uncertainty and noise, *Evolutionary multi-criterion optimization*, Vol. 1993, 1999, pp. 329–343.

[212]   Teich, J., Pareto-front exploration with uncertain objectives, *Evolutionary multi-criterion optimization*, Vol. 1993, 2001, pp. 314–328.

[213]   Salo, A. A. and Hämäläinen, R. P., Preference programming through approximate ratio comparisons, *European Journal of Operational Research*, Vol. 82, 1995, pp. 458–475.

[214]   See, T.-K., Gurnani, A., and Lewis, K., Multi-Attribute Decision Making using Hypothetical Equivalents and Inequivalents, *Transactions of the ASME*, Vol. 126, 2004, pp. 950–958.

[215]   Keeney, R. L. and Raiffa, H., *Decisions with Multiple Objectives*, Cambridge University Press, 1993.

[216]   NASA, NASA Schedule Management Handbook, Tech. Rep. NASA/SP-2010-3403, 2010.

[217]   Book, S. A., Why correlation matters in cost estimating, 32nd Annual DoD Cost Analysis Symposium, Williamsburg, VA, February, 1999.

[218]   Harmon, B., Formulation of default correlation values for cost risk analysis, Presented at SSCAG/SCAF/EACEWG Joint Meeting, Berlin, Germany, May 11-12, 2010.

[219]   Coppinger, R., Russia Wants Reusable Rockets By 2020, 2012, Retrieved on February 13, 2013.