

# Reliability as an Independent Variable Applied to Liquid Rocket Engine Test Plans

Richard Strunz\*

EADS Astrium, GmbH, 81663 Munich, Germany

and

Jeffrey W. Herrmann†

University of Maryland, College Park, Maryland 20742

DOI: 10.2514/1.B34279

**Manufacturers lack an adequate method to balance performance, reliability, and affordability. The reliability-as-an-independent-variable methodology 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. The overall solution strategy of reliability-as-an-independent-variable 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 reliability-as-an-independent-variable methodology is exemplarily applied to the actual hot-fire test history of the F-1, the space shuttle main engine, and the RS-68 liquid rocket engine, showing adequate agreement between computed results and actual flight engine reliability.**

## Nomenclature

$\text{Beta}(\cdot)$	= Beta probability density function
$\text{Bi}(\cdot)$	= Binomial probability density function
$C$	= level of confidence (credibility bound)
$c$	= number of cycles
CTD	= cumulated test duration, s
$F_x$	= cumulative density function
HW	= number of hardware
$L(\cdot)$	= likelihood function
$n$	= number of trials, i.e., hot-fire tests
$p$	= probability of success
$q$	= failure fraction
$q_{\text{ind}}^*$	= candidate probability density function
$R$	= reliability
$r$	= number of failures
$t$	= hardware life, s
tp	= hot-fire test proportion
$T_{50}$	= median lifetime, s
$w$	= weighting factor
$x$	= number of successes, i.e., hot-fire tests
$\alpha$	= shape parameter
$\alpha_1, \alpha_2$	= weighting factors
$\beta$	= shape parameter
$\beta_{(\cdot)}$	= median regression coefficients
$\Theta$	= integration domain
$\theta$	= parameters to be estimated
$\gamma$	= acceptance rate
$\Gamma(\cdot)$	= Gamma function

$\lambda$	= parameter of Poisson probability mass function
$\pi$	= probability of success of a functional node
$\pi(\cdot)$	= posterior probability
$\pi_0$	= prior probability
$\sigma$	= standard deviation

## Subscripts

$D$	= integration domain
$i$	= functional node
$j$	= hot-fire test group
med	= median
MD	= mission duration
MIQ	= mission ignition quantity
pr	= prior used in empirical Bayes estimation
rem	= remaining
tot	= total
$u$	= upper percentile
0	= used in empirical Bayes estimation

## I. Introduction

**L** IQUID rocket engines have always been one of the major affordability drivers of launch vehicles, but in the past, national prestige or military requirements were dominating the decisions for the development of a new launch vehicle. This paradigm has changed. Performance, reliability, and affordability 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.<sup>‡</sup> 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 competitive launcher<sup>§</sup>

Received 7 March 2011; revision received 25 May 2011; accepted for publication 9 June 2011. Copyright © 2011 by Astrium, GmbH. Published by the American Institute of Aeronautics and Astronautics, Inc., with permission. Copies of this paper may be made for personal or internal use, on condition that the copier pay the \$10.00 per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923; include the code 0748-4658/11 and \$10.00 in correspondence with the CCC.

\*Program Manager, Advanced Programs and Systems Technology; Ph.D. Student, Department of Mechanical Engineering, University of Maryland, College Park, Maryland 20742 (Corresponding Author).

†Associate Professor, Department of Mechanical Engineering and Institute for Systems Research, Martin Hall, Room 2181.

<sup>‡</sup>Data available at [http://www.esa.int/esaMI/Launchers\\_Home/SEMNCI1PGQD\\_0.html](http://www.esa.int/esaMI/Launchers_Home/SEMNCI1PGQD_0.html) [retrieved 30 October 2010].

<sup>§</sup>Data available at <http://www.satnews.com/cgi-bin/story.cgi?number=2083189758> [retrieved 30 October 2010].

**Table 1 Qualification/certification hot-fire test attributes**

Test attributes	F-1	J-2	RL10	LR87	LR91	SSME	RS-68 <sup>a</sup>
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
Fleet leader concept used	No	No	No	No	No	Yes	No
Overstress testing	No	No	Yes	No	No	Yes	Yes

<sup>a</sup>Values are based on data given in [6]

**Table 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 nominal time (max), s	165	500	700	165	225	520	250
Multiple of mission nominal time, s	13.6	3.8	2.1	12.1	11.3	4.8	3.6

but at no expense of the mission success reliability and other launch service competitive factors compared with the current European launch vehicle Ariane 5, as well as other operational launch vehicles.

The affordability of expendable launch vehicles is largely determined by the development and production costs of liquid rocket engines [1,2]. The major part of the development cost is spent on development test hardware subjected to hot-fire tests in order to sufficiently demonstrate design maturity/robustness and to qualify/certify the liquid rocket engines for a successful flight operation [3]. The reliability-as-an-independent-variable (RAIV) methodology 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 mainly driven 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 levers on the development cost is the chosen verification strategy with the objective to minimize the number of hardware sets by testing the mission requirements on a single hardware set multiple times but at the expense of a higher 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, and range cost, as well as insurance fees [4].

Therefore, the selection of the optimal choice for the settings of 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; this paper proposes the RAIV methodology to address this shortcoming.

## II. Background

Liquid rocket engine qualification or, synonymously, flight certification has always been a concern of space industry and agency alike because no industry or governmentwide 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 [5].

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 [2]. 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 (HFTD) [6].

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 1 [7]. 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 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.

The surveys performed by Emdee [8,9] and Pempie and Vernin [10] 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 s accumulated test duration spread over 15 hardware sets to 150 hot-fire tests with at least 50,000 s of accumulated test duration but without a statement about a required number of hardware sets.

Despite these two recommendations, Wood [6] reports that the RS-68 engine was subjected to 183 hot-fire tests with an accumulated test duration of only 18,945 s spread over eight new and four

refurbished hardware sets before the maiden flight on Delta IV. Greene [11] 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 s accumulated hot-fire duration spread over more than 90 hardware sets [12].

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 [13]. 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) [14]. The methodology 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 [6], most likely due to budget constraints. Consequently, the lack of an industry or government standard or guideline remains evident.

### III. Reliability-as-an-Independent-Variable Methodology

The RAIV methodology is the proposed 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 methodology is to generate the quantitative figures of the reliability trade space.

The inputs to the RAIV methodology 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 methodology 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.

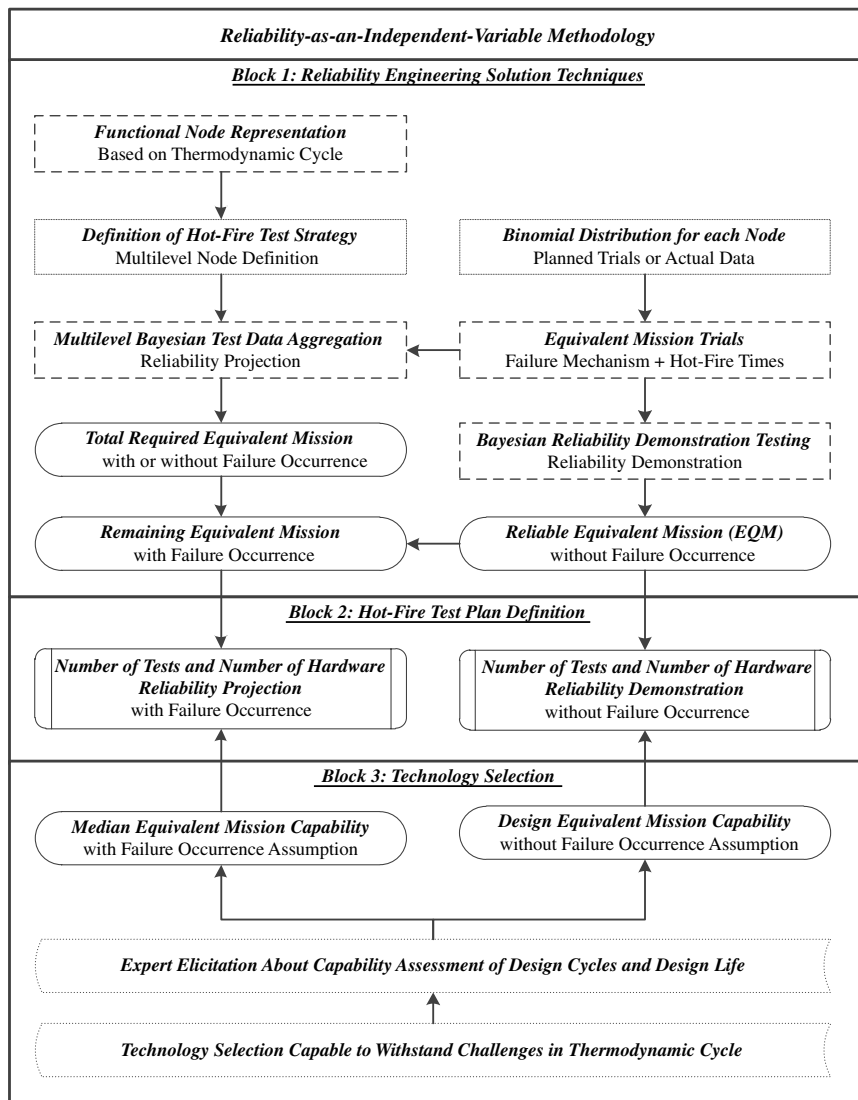


Fig. 1 RAIV methodology.

The overall RAIV methodology is depicted in Fig. 1. The main steps of the methodology are listed below. Details related to each step are given in Secs. III.A–III.G.

1) To define the hot-fire test strategy, the functional architecture of liquid rocket engines 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.

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 are 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.

#### A. 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 [14–17]. 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 methodology, 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 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 [6]. 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 in Sec. IV.A.

#### B. 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 HFTD group is a set of tests of the components associated with a functional node, where all the tests in this group

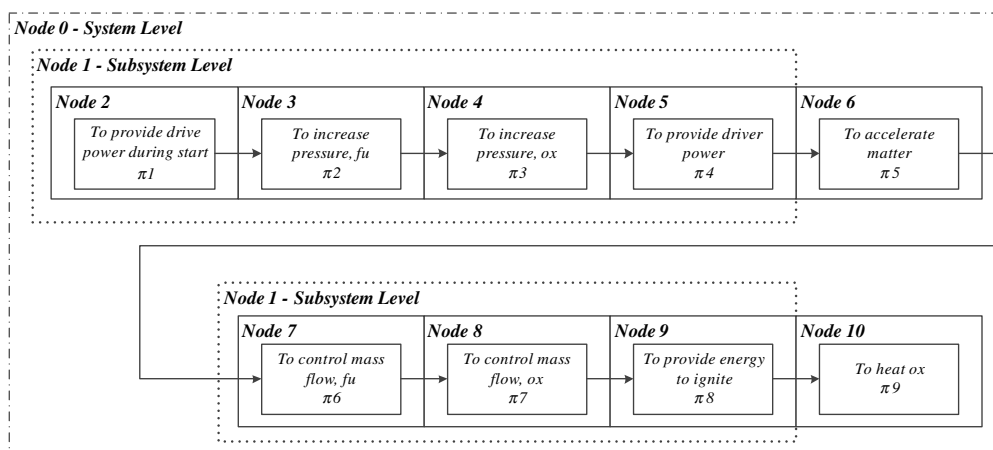


Fig. 2 Functional node representation of the RS-68 liquid rocket engine.

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  $CT_{i,j}$ ; the number of failures  $r_{i,j}$ ; the test duration  $HFTD_{i,j}$ ; the weighting factors for the two failure mechanisms,  $\alpha_{1,i,j}$  and  $\alpha_{2,i,j}$ ; a weighting factor to account for hot-fire tests shorter than full mission duration  $w_{i,j}$ ; and an acceleration factor (AF) to account for different operational load points  $AF_{i,j}$ . As described in the following paragraphs, these data are used to determine  $EQM_{i,j}$ , the EQMs of these tests,  $EQM_i$ , and the EQMs of all of the tests for a functional node.

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 until Secs. III.D–III.F, because the different hot-fire events must be normalized first with respect to the mission and different HFTDs 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 [18,19].

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. (1). The first term reflects the stress-increased failure mechanism, and the second term reflects the strength-reduced failure mechanism, respectively:

$$EQM = \alpha_1 \cdot \frac{CT}{MLC_{MIQ}} + \alpha_2 \cdot \frac{CTD}{MLC_{MD}} \quad (1)$$

where

$$\sum_k \alpha_k = 1$$

$\alpha_1$  and  $\alpha_2$  are weighting factors, CT is the number of cycles tested, CTD is the accumulated test duration,  $MLC_{MIQ}$  is the MLC ignition quantity without overhaul of the system in between the missions, and  $MLC_{MD}$  is the mission duration accumulated in various hot-fire operations in the MLC.

The weighting factors  $\alpha_1$  and  $\alpha_2$  are assumed to be equal (both are 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 [19].

Startup and shutdown modes are more detrimental than the steady-state operational mode of liquid rocket engines [20,21]; therefore, some weighting inside the CTD is used to account for these different effects. Worlund et al. [22] 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 [20]. These data were partially used to describe the relationship as given in Eq. (2):

$$w_j = \frac{Ln(tp_j) - \beta_{med,0}}{\beta_{med,1}} \quad (2)$$

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 [20] with a Dirichlet prior distribution.

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 methodology can be seen in Sec. IV.

Introducing all extensions, the final EQM equation for a hot-fire test group  $j$  within a functional node  $i$  is given in Eq. (3):

$$EQM_{i,j} = \alpha_{1,i,j} \cdot \frac{CT_{i,j}}{MLC_{MIQ}} + \alpha_{2,i,j} \cdot \frac{CT_{i,j} \cdot (AF_{i,j} \cdot w_{i,j} \cdot HFTD_{i,j})}{MLC_{MD}} \quad (3)$$

where  $\alpha_{1,i,j}$  and  $\alpha_{2,i,j}$  are the weighting factors for the two failure mechanisms,  $CT_{i,j}$  is the number of cycles tested (one cycle consists of the startup and shutdown),  $MLC_{MIQ}$  is the MLC ignition quantity without overhaul of the system in between the missions,  $(AF_{i,j} \cdot w_{i,j} \cdot HFTD_{i,j})$  is the rated and weighted test duration times the number of cycles tested  $CT_{i,j}$ , and  $MLC_{MD}$  is the HFTD 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  $EQM_i$  and equivalent number of successful trials  $x_i$ . The number of total trials  $EQM_i$  is given in Eq. (4):

$$EQM_i = \sum_{j=1}^k EQM_{i,j} \quad (4)$$

where  $EQM_{i,j}$  is the EQM as defined in Eq. (3). The number of equivalent successful trials  $x_i$  is given in Eq. (5):

$$x_i = EQM_i - \sum_{j=1}^k \left[ \alpha_{1,i,j} \cdot \frac{CT_{i,j}}{MLC_{MIQ}} + \alpha_{2,i,j} \cdot \frac{CT_{i,j} \cdot (AF_{i,j} \cdot w_{i,j} \cdot HFTD_{i,j})}{MLC_{MD}} \right] \quad (5)$$

where  $EQM_i$  is defined in Eq. (4), and the second term is equivalent to Eq. (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) and (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.

### C. 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 in Eq. (6). The solution of Eq. (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.,  $\pi_i = f_i(\pi_i)$ :

$$\pi(\underline{\pi}_i|\text{Data}) = \frac{\prod_i \pi_i^{x_i} \cdot (1 - \pi_i)^{\text{EQM}_i - x_i} \cdot \prod_i \pi_{0,i}}{\int \dots \int_{\Theta} (\prod_i \pi_i^{x_i} \cdot (1 - \pi_i)^{\text{EQM}_i - x_i} \cdot \prod_i \pi_{0,i}) \cdot d\pi_{\Theta}} \quad (6)$$

where  $\pi(\underline{\pi}_i|\text{Data})$  is the posterior of the parameter vector  $\underline{\pi}_i$  given the Data,  $\pi_i$  is the individual lowest level functional node success probability,  $\pi_{0,i}$  is the prior distribution of the individual lowest level functional node success probability, and Data is the multilevel data in terms of EQM  $\text{EQM}_i$  as defined in Eq. (4) and equivalent successes  $x_i$  as defined in Eq. (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  $\Theta$ , even with modern general-purpose multidimensional integration algorithms [23]. 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 [24–30]. The 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  $q_{\text{ind}}^*$  as given in Eq. (7):

$$q_{\text{ind}}^* = \frac{1}{1 + e^{-F_X^{-1}[F_X(N(\logit \pi_i, \sigma_i))]}]} \quad (7)$$

where  $F_X^{-1}$  is the inverse of the cumulative distribution function for the candidate distribution as a function of the quantile of the distribution function, i.e.,  $N(\logit \pi_i, \sigma_i)$ . The standard deviation  $\sigma_i$  of the distribution function is a tunable constant that influences the one-variable-at-a-time acceptance rate of the acceptance probability  $\gamma_i$  for new candidate values for each functional node probability  $\pi_i$ . The acceptance probability is given in Eq. (8):

$$\gamma_i = \text{Min} \left\{ 1, \frac{\pi^*(\pi_i|\text{Data})}{\pi(\pi_i|\text{Data})} \cdot \frac{q_{\text{ind},i}^* \cdot (1 - q_{\text{ind},i}^*)}{q_{\text{ind},i} \cdot (1 - q_{\text{ind},i})} \right\} \quad (8)$$

where  $\pi^*(\pi_i|\text{Data})$  is the unscaled target density calculated with a new candidate value,  $\pi(\pi_i|\text{Data})$  is the unscaled target density calculated with the previous accepted value,  $q_{\text{ind},i}^*$  is the new candidate value, and  $q_{\text{ind},i}$  is the previous value of the Markov chain.

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  $\sigma_i$  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  $\sigma_i$  in such a way that the acceptance rates of each individual parameter are close to 0.35 [15,29,30].

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

The selection of the prior distributions for the functional node parameters  $\pi_i$  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.

#### D. Bayesian Reliability Demonstration Testing

The main advantage of the BRDT technique is the reduction of test sample size [31]. 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. (9):

$$L(\text{Data}|q) = \binom{\text{EQM}_{\text{bogeey}}}{r} \cdot q^r \cdot (1 - q)^{\text{EQM}_{\text{bogeey}} - r} \quad (9)$$

where  $q$  is the failure fraction,  $\text{EQM}_{\text{bogeey}}$  is the number of EQMs without failures performed at system level, and  $r$  is the number of observed equivalent failures during testing (usually assumed to be zero in the BRDT). 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. (10):

$$\pi_0(q) = \frac{q^{\alpha-1} \cdot (1 - q)^{\beta-1}}{\text{Beta}(\alpha, \beta)} \quad (10)$$

where  $q$  is the failure fraction (the parameter to be estimated in the Bayes theorem),  $\alpha$  and  $\beta$  are the shape parameters of the Beta distribution, and  $\text{Beta}(\alpha, \beta)$  is the solution of the Eulerian integral of the first kind:

$$\int_0^1 x^{\alpha-1} (1 - x)^{\beta-1} dx$$

An empirical Bayes approach was used to estimate the parameter settings for the shape parameters  $\alpha$  and  $\beta$  using the data given in McFadden and Shen [32]. The procedures described by Martz and Waller [33,34] or by Modarres et al. [18] were applied that lead to the same parameter estimates, but the latter one is mathematically more appealing and is given in Eq. (11):

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

$$\alpha = x_0 \quad \text{and} \quad \beta = n_0 - x_0 \quad (11)$$

where  $p_{\text{pr}}$  is the prior mean,  $x_0$  are the successes,  $n_0$  are the trials,  $k$  is the coefficient of variation, and  $\alpha$  and  $\beta$  are the shape parameters of the Beta distribution. The estimated shape parameters  $\alpha$  and  $\beta$  that correspond to the mean as well as the 0.05 and 0.95 percentiles of liquid rocket engine reliability are listed in Table 3.

The posterior distribution percentiles of the failure fraction  $q$  are related to the binomial distribution as given in Eq. (12):

**Table 3 Shape parameter for the Beta prior distribution in the BRDT plan**

	0.05 percentile	Average	0.95 percentile
Shape parameter $a$	21	39	42.0
Shape parameter $b$	1.2	0.5	0.6

$$P(q \leq q_u) = \int_0^{q_u} \pi(q|\text{Data}) \cdot dq = C \quad (12)$$

where  $q_u$  is the upper percentile of the posterior distribution,  $\pi(q|\text{Data})$  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. (13):

$$\frac{\text{Beta}_{q_u}(q_u, \alpha + r, \beta + \text{EQM}_{\text{bogy}} - r) \cdot \Gamma(\alpha + \beta + \text{EQM}_{\text{bogy}})}{\Gamma(\beta + \text{EQM}_{\text{bogy}} - r) \cdot \Gamma(\alpha + r)} = C \quad (13)$$

where  $\text{EQM}_{\text{bogy}}$  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,  $\alpha$  and  $\beta$  are the Beta distribution shape parameters,  $C$  is the credibility bound,  $\text{Beta}_{q_u}(\cdot)$  is the incomplete beta function, and

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

is the Gamma function.

The equivalency of Eq. (13) with the well-known frequentist binomial model  $(1 - C) = R^n$  may not be obvious, but if Eq. (13) is rewritten using a vague prior (parameters  $\alpha$  and  $\beta$  are set to 1), and assuming a zero failure success testing, the Bayesian-like binomial model can be stated as given in Eq. (14):

$$\text{Bi}(0, q, n + 1) = (1 - q)^{n+1} = (1 - C) \Leftrightarrow R^{n+1} = (1 - C) \quad (14)$$

where  $\text{Bi}(r, q, n + 1)$  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. Life Capability of Hardware Sets

The RAIV methodology 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, but one of the main deficits of the reported values is the lack of an associated reliability statement [7].

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 methodology. It is important to note that the bogey EQL ( $\text{EQL}_{\text{bogy}}$ ) 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  $\text{EQL}_{\text{bogy}}$  is also listed in Table 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.

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. (1) but with relevant modifications linked to the bogey number of cycles, the bogey life, and the 0.5 percentiles. The bogey EQL ( $\text{EQL}_{\text{bogy}}$ ) is given in Eq. (15), and the median EQL ( $\text{EQL}_{\text{med}}$ ) is given in Eq. (16):

$$\text{EQL}_{\text{bogy}} = \alpha_1 \cdot \frac{c_{\text{bogy}}}{\text{MLC}_{\text{MIQ}}} + \alpha_2 \cdot \frac{t_{\text{bogy}}}{\text{MLC}_{\text{MD}}} \quad (15)$$

where

$$\sum_k \alpha_k = 1$$

$\alpha_1$  and  $\alpha_2$  are weighting factors both set to 0.5,  $c_{\text{bogy}}$  is the bogey number of cycles,  $t_{\text{bogy}}$  is the bogey life,  $\text{MLC}_{\text{MIQ}}$  is the MLC ignition quantity without overhaul of the system in between the missions, and  $\text{MLC}_{\text{MD}}$  is the MLC mission duration:

$$\text{EQL}_{\text{med}} = \alpha_1 \cdot \frac{\lambda}{\text{MLC}_{\text{MIQ}}} + \alpha_2 \cdot \frac{T_{50}}{\text{MLC}_{\text{MD}}} \quad (16)$$

where

$$\sum_k \alpha_k = 1$$

$\alpha_1$  and  $\alpha_2$  are weighting factors set to 0.5,  $\lambda$  is the median number of cycles to failure,  $T_{50}$  is the median life,  $\text{MLC}_{\text{MIQ}}$  is the MLC ignition quantity without overhaul of the system in between the missions, and  $\text{MLC}_{\text{MD}}$  is the HFTD accumulated during the MLC. The median number of cycles to failure  $\lambda$  is given by Eq. (17), and the median life  $T_{50}$  is given by Eq. (18):

**Table 4 Engine design and mission requirements**

Liquid rocket engine	Design cycles	Design life, s	Missions	Mission starts	Mission nominal time, s	MLC cycles	MLC HFTD, s	$\text{EQL}_{\text{bogy}}$
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
					150 <sup>a</sup>	4	600	6.8
					350 <sup>a</sup>		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

<sup>a</sup>First hot fire and restart.

$$P(\text{CT} \leq c_{\text{bogy}}) = R(c_{\text{bogy}}) = 1 - \frac{\Gamma(1 + \lfloor c_{\text{bogy}} \rfloor, \lambda)}{\Gamma(1 + \lfloor c_{\text{bogy}} \rfloor)} \quad (17)$$

where  $P(\text{CT} \leq c_{\text{bogy}})$  is the probability of the discrete random variable cycle tested CT evaluated at the bogey number of cycles  $c_{\text{bogy}}$  that equals  $R(c_{\text{bogy}})$  as the number of cycles for which the reliability will be  $R$ ,  $\lambda$  is the mean of the Poisson distribution,

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

is the Gamma function, and  $\lfloor \cdot \rfloor$  is the floor function. The search parameter is the mean  $\lambda$  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  $\lambda$  is above nine, for which the mean and the median will be indistinguishable:

$$T_{50} = t_{\text{bogy}} \left[ \frac{\ln(2)}{-\ln(R)} \right]^{1/\beta} \quad (18)$$

where  $T_{50}$  is the median life time,  $t_{\text{bogy}}$  is the bogey life,  $R$  is the reliability level of interest, and  $\beta$  is the shape parameter of a two-parameter 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  $\beta$  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 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 Sec. IV for the initial SSME design cycle and design life assumptions).

#### F. Number of Hardware Sets

The number of hardware sets needed to complete the RAIV methodology hot-fire test scope is calculated using the hardware reliability necessary to support the total required EQM ( $\text{EQM}_{\text{tot}}$ ) based on the multilevel BTDA technique as given in Eq. (19):

$$\text{EQM}_{\text{tot}} = \text{EQM}_{\text{bogy}} + \text{EQM}_{\text{rem}} \Leftrightarrow \text{HW}_{\text{tot}} = \text{HW}_{\text{bogy}} + \text{HW}_{\text{rem}} \quad (19)$$

where  $\text{EQM}_{\text{bogy}}$  is the required EQM to support the BRDT, and  $\text{EQM}_{\text{rem}}$  is the remaining EQM needed to complete the overall RAIV methodology hot-fire test scope defined by Eq. (6).

Equation (19) can be modified with the corresponding life capability in order to define the required number of hardware sets as given in Eq. (20):

$$\text{HW}_{\text{tot}} = \frac{\text{EQM}_{\text{bogy}}}{\text{EQL}_{\text{bogy}}} + \frac{\text{EQM}_{\text{rem}}}{\text{EQL}_{\text{med}}} \quad (20)$$

where  $\text{EQM}_{\text{bogy}}$  is the required EQM for the BRDT with corresponding bogey EQL ( $\text{EQL}_{\text{bogy}}$ ), and  $\text{EQM}_{\text{rem}}$  is the remaining EQM to complete the overall RAIV methodology hot-fire test scope with corresponding median EQL ( $\text{EQL}_{\text{med}}$ ) based on the 0.5 percentile.

#### G. 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 HFTDs, the development duration, and the development cost. The RAIV methodology 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 [35,36].

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

### IV. Numerical Examples

The application and demonstration of the RAIV methodology with artificial hot-fire test data would lack credibility in the space industry. Therefore, the numerical examples used for the validation of the RAIV methodology are based on the hot-fire test histories of the F-1 liquid rocket engine, 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 Sec. III.

#### A. Define Hot-Fire Test Strategy

The RAIV methodology 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<sup>1</sup> for the number of hot-fire tests of 1081, for which 278 tests were for 150 s or longer. No information is given on the accumulated HFTDs. The SSME hot-fire test history featured 726 hot-fire tests with 110,253 s of accumulated HFTD [37]. The RS-68 was qualified with 183 hot-fire tests and 18,945 s [6]. 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 5 for the F-1, as given in Table 6 for the SSME, and as given in Table 7 for the RS-68.

#### B. 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 [37]. 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 tests that are shorter than full mission duration that were given in Worland et al. [22] and a Bayesian solution for the parameters of the likelihood function of the model introduced by Lloyd and Lipow [20]. The resulting accumulated hot-fire test time is about 111,000 s, with the average HFTD of around 100 s that can be compared with the data given in Emdee [9], which result in the average HFTD of roughly 90 s. Likewise, the RS-68 hot-fire testing groups are based on a test allocation that resulted in the accumulated HFTD that is given by Wood [6]. 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 [20].

<sup>1</sup>Data available at [http://history.msfc.nasa.gov/saturn\\_apollo/documents/Background\\_F-1\\_Rocket\\_Engine.pdf](http://history.msfc.nasa.gov/saturn_apollo/documents/Background_F-1_Rocket_Engine.pdf) [retrieved 15 May 11].



**Table 5 Functional nodes of the F-1 mapped to physical components**

Functional node	Physical component
To provide ignition power $\pi_1$	Ignition system components
To increase pressure $\pi_2$	Single shaft turbopump arrangement (including gear)
To provide drive power $\pi_3$	GG
To accelerate matter $\pi_4$	Thrust chamber assembly (TCA)
To control mass flow, fuel side $\pi_5$	Valves on fuel side
To control mass flow, oxidizer side $\pi_6$	Valves on oxidizer side

**Table 6 Functional nodes of the SSME mapped to physical components**

Functional node	Physical component
To increase pressure, fuel side $\pi_1$	Boost and turbopump, fuel side
To increase pressure, oxidizer side $\pi_2$	Boost and turbopump, oxidizer side
To provide drive power, fuel side $\pi_3$	Preburner to drive turbine, fuel side
To provide drive power, oxidizer side $\pi_4$	Preburner to drive turbine, oxidizer side
To accelerate matter $\pi_5$	TCA (flight nozzle extension)
To control mass flow, fuel side $\pi_6$	Main fuel valve
To control mass flow, oxidizer side $\pi_7$	Main oxidizer valve, preburner oxidizer valves
To provide energy to ignite $\pi_8$	Igniters for preburners and TCA
To heat oxidizer $\pi_9$	Heat exchanger to pressurize tank

**Table 7 Functional nodes of the RS-68 mapped to physical components**

Functional Node	Physical component
To provide drive power during start, $\pi_1$	Starter
To increase pressure, fuel side $\pi_2$	Turbopump, fuel side
To increase pressure, oxidizer side $\pi_3$	Turbopump, oxygen side
To provide drive power, $\pi_4$	GG to drive the fuel and oxygen pumps
To accelerate matter $\pi_5$	TCA
To control mass flow, fuel side $\pi_6$	GG and TCA valves, fuel
To control mass flow, oxidizer side $\pi_7$	GG and TCA valves, ox
To provide energy to ignite $\pi_8$	Igniters for GG and TCA
To heat oxidizer $\pi_9$	Heat exchanger to pressurize tank

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 testbed (ISTB) testing, and the RS-68 that includes GG component-level and PP subsystem-level testing are given in Tables 8–10 (RS-68 with AF of one) and Table 11

(RS-68 with AF of five), respectively. The assumption of an AF of five is given to investigate the impact on the number of hardware sets and the resulting reliability projection level. Using Eqs. (3–5), the EQMs and the number of successful trials were determined as required inputs for Eq. (6).

### C. Estimate the Reliability Projection Metric

The EQMs and the number of successful trials given in Tables 8–11 are used in the multilevel BTDA using Eq. (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 12.

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% [2] and the reported reliability level of 0.984 [3], respectively. In addition, Koelle [3] 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 s 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 up to now.

Table 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.

### D. Estimate the Reliability-by-Confidence Metric

The *R-by-C* metric is used as input for the hardware reliability requirements, and consequently for the number of hardware sets required for the overall hot-fire test strategy. Equation (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, a confidence level of 50% was used for the F-1 engine [2], confidence level of 60% was used for the SSME, and a confidence level of 90% was used for the RS-68 engine [14].

### E. 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 13 lists the bogey or design cycles and bogey or design life for the F-1 [9] and SSME [7]. The bogey cycles and bogey life for the

**Table 8 Multilevel BTDA scope: F-1<sup>a</sup>**

	$r_{i,j}$	EQM <sub><i>i,j</i></sub>	Hot-fire tests	Acc. hot-fire time, s	$w_{i,j}$	$\alpha_{1,i,j}$	$\alpha_{2,i,j}$	$x_i$	EQM <sub><i>i</i></sub>	$n_i$	Total acc. hot-fire time, s
<i>Node 0: system</i>											
Group 1 [15 s]	1	5.4	30	450	0.44	0.50	0.50				
Group 2 [50 s]	1	11.3	50	2,500	0.66	0.50	0.50				
Group 3 [80 s]	1	90.3	323	25,840	0.78	0.50	0.50				
Group 4 [100 s]	1	127.6	400	40,000	0.84	0.50	0.50				
Group 5 [150 s]	2	107.4	250	37,500	0.96	0.50	0.50				
Group 6 [165 s]	0	13.1	28	4,620	1.00	0.50	0.50				
Node 0: engine								353.5	355.1	1,081	110,910

<sup>a</sup>Acc. denotes accumulated.

**Table 9 Multilevel BTDA scope: SSME<sup>a</sup>**

	$r_{i,j}$	$EQM_{i,j}$	Hot-fire tests	Acc. hot-fire time, s	$w_{i,j}$	$\alpha_{1,i,j}$	$\alpha_{2,i,j}$	$x_i$	$EQM_i$	$n_i$	Total acc. hot-fire time, s
<i>Node 1: ISTB</i>											
Group 1 [0–100 s]	0		1000	100,000	0.75	0.50	0.50				
<i>Node 0: system</i>											
Group 1 [mean 2 s]	0	3.4	27	54	0.15	0.50	0.50				
Group 2 [mean 21 s]	0	14.1	107	2,247	0.51	0.50	0.50				
Group 3 [mean 97 s]	3	31.1	184	17,848	0.74	0.50	0.50				
Group 4 [mean 158 s]	4	26.9	132	20,856	0.82	0.50	0.50				
Group 5 [mean 183 s]	4	26.5	121	22,143	0.84	0.50	0.50				
Group 6 [mean 283 s]	3	36.0	128	36,224	0.91	0.50	0.50				
Group 7 [mean 400 s]	0	7.5	21	8,400	0.96	0.50	0.50				
Group 8 [520 s]	0	2.6	6	3,120	1.00	0.50	0.50				
Node 1: ISTB								197.1	197.1	1,000	100,000
Node 0: engine								147.5	148.1	726	110,892

<sup>a</sup>Acc. denotes accumulated.

**Table 10 Multilevel BTDA Scope (AF = 1): RS-68<sup>a</sup>**

	$r_{i,j}$	$EQM_{i,j}$	Hot-fire tests	Acc. hot-fire Time, s	$w_{i,j}$	$\alpha_{1,i,j}$	$\alpha_{2,i,j}$	$x_i$	$EQM_i$	$n_i$	Total acc. hot-fire time, s
<i>Node 5: GG</i>											
Group 1 [0–50 s]	2	12.5	62	3,100	0.04	0.50	0.50				
<i>Node 1: PP</i>											
Group 1 [0–100 s]	1	1.8	6	600	0.42	0.50	0.50				
<i>Node 0: system</i>											
Group 1 [Mean 28 s]	3	13.1	78	2,195	0.04	0.50	0.50				
Group 2 [mean 136 s]	3	4.4	18	2,450	0.53	0.50	0.50				
Group 3 [mean 139 s]	3	6.9	28	3,900	0.53	0.50	0.50				
Group 4 [mean 163 s]	3	6.5	24	3,900	0.59	0.50	0.50				
Group 5 [mean 173 s]	3	4.2	15	2,600	0.59	0.50	0.50				
Group 6 [mean 195 s]	3	5.8	20	3,900	0.59	0.50	0.50				
Node 5: GG								12.2	12.5	62	3,100
Node 1: PP								1.8	1.8	6	600
Node 0: engine								37.0	40.9	183	18,945

<sup>a</sup>Acc. denotes accumulated.

**Table 11 Multilevel BTDA scope (AF = 5): RS-68<sup>a</sup>**

	$r_{i,j}$	$EQM_{i,j}$	Hot-fire tests	Acc. hot-fire time, s	$w_{i,j}$	$\alpha_{1,i,j}$	$\alpha_{2,i,j}$	$EQM_i$	$n_i$	Total acc. hot-fire time, s
<i>Node 5: GG</i>										
Group 1 [0–50 s]	2	16.7	62	3,100	0.04	0.50	0.50			
<i>Node 1: PP</i>										
Group 1 [0–100 s]	1	1.9	6	600	0.42	0.50	0.50			
<i>Node 0: system</i>										
Group 1 [mean 28 s]	3	20.3	78	2,195	0.04	0.50	0.50			
Group 2 [mean 136 s]	3	5.7	18	2,450	0.53	0.50	0.50			
Group 3 [mean 139 s]	3	8.9	28	3,900	0.53	0.50	0.50			
Group 4 [mean 163 s]	3	7.9	24	3,900	0.59	0.50	0.50			
Group 5 [mean 173 s]	3	5.0	15	2,600	0.59	0.50	0.50			
Group 6 [mean 195 s]	3	7.3	20	3,900	0.59	0.50	0.50			
Node 5: GG								16.7	62	3,100
Node 1: PP								1.9	6	600
Node 0: engine								55.1	183	18,945

<sup>a</sup>Acc. denotes accumulated.

RS-68 were defined through the RAIV methodology. 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 methodology.

**F. 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 12, the confidence levels of 50, 60, and 90%, the life capability, and the number of hardware sets are determined with Eq. (20). The results in terms of average, minimum, and maximum numbers of hardware sets

**Table 12 Reliability Projection Levels using the RAIV Methodology**

Liquid rocket engine	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

**Table 13 Total number of hardware sets**

Liquid rocket engine	<i>R</i> -by- <i>C</i>	Reported HW	Design life	Design cycles	$\beta = 3 \pm 0.5$	$\beta = 4 \pm 0.5$
F-1 (1,081 hot-fire tests)	0.9894 at 0.5	56 <sup>a</sup>	2,250	20	$20 \leq 22 \leq 24$	$24 \leq 26 \leq 28$
F-1 (2,740 hot-fire tests)	0.9952 at 0.5	56 <sup>a</sup>	2,250	20	$39 \leq 46 \leq 51$	$51 \leq 55 \leq 59$
SSME, specified life capability <sup>b</sup>	0.9825 at 0.6	20	27,000 <sup>c</sup>	55 <sup>c</sup>	$4 \leq 5 \leq 6$	$6 \leq 6 \leq 7$
SSME, realistic test bogey <sup>b</sup>	0.9825 at 0.6	20	5,000 <sup>d</sup>	20 <sup>d</sup>	$17 \leq 21 \leq 23$	$23 \leq 26 \leq 27$
RS-68	0.9454 at 0.9	8 + 4 <sup>e</sup>	4,000 <sup>f</sup>	15 <sup>f</sup>	$9 \leq 10 \leq 10$	$10 \leq 10 \leq 10$

<sup>a</sup>Reported in [2,9] but spread over roughly 2,740 hot-fire tests.

<sup>b</sup>Same hot-fire test plan assumed.

<sup>c</sup>Original design life and cycles requirement.

<sup>d</sup>Realistic lifetime and cycle numbers derived from [38].

<sup>e</sup>Four engine hardware sets were refurbished [6].

<sup>f</sup>Estimated test bogey life and test bogey cycles based on the RAIV methodology.

are given in Table 13 assuming Weibull shape parameters of  $3 \pm 0.5$  and  $4 \pm 0.5$  for the median lifetime estimation. The estimation of the  $EQM_{\text{bogey}}$  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.

The findings of Table 13 may suggest the use of Weibull shape parameters of  $3 \pm 0.5$  for LOX/liquid hydrogen and  $4 \pm 0.5$  for LOX/kerosene liquid rocket engines based on the estimated number of hardware sets using the RAIV methodology and the corresponding reported values. Certainly, further investigations are needed to make conclusions. The more important aspect of the results listed in Table 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 s), only five to six hardware sets would have been required for the complete development program using the RAIV methodology 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 methodology using the more realistic test bogey capability of 20 cycles and 5000 s. 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 [38]. Therefore, the SSME example demonstrates that any unrealistic test bogey capability assumption, which is used in tradeoff studies, may result in infeasible hot-fire test plan definitions and may cause strong program cost overruns and schedule slippage.

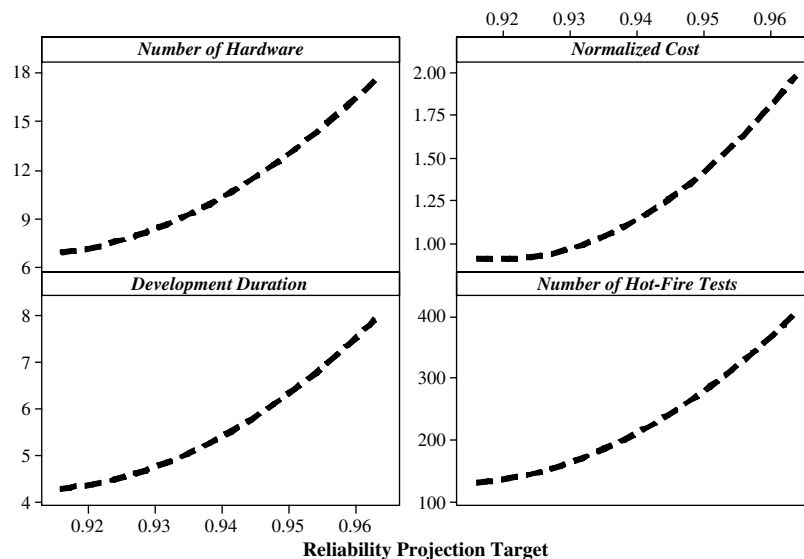
### G. Test Plan Optimization

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 (Note that 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/U.S Air Force Costing Model (NAFCOM) in combination with a specific hot-fire test cost model using the results of the RAIV methodology. 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 methodology determined the optimal hot-fire test plan. The results (presented in Fig. 3) highlight quantitatively the expected tendencies of the claims presented in Sec. II. 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



**Fig. 3 Key test plan metrics for various reliability projection targets: RS-68 test case.**

reliability projection level is raised from the initial level of 93.3 to 95% (Note that the reliability level indicated is based on the case where no AF is used in the RAIV methodology).

## V. Conclusions

The presented RAIV methodology 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 methodology can be used to define test boogies that are associated with a reliability requirement that may also be used as a design requirement. One additional strength of the RAIV methodology 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 methodology with thermodynamic cycle models and parametric cost models. Although the RAIV methodology was exemplarily 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.

## Acknowledgments

The Corresponding Author would like to acknowledge Astrium, GmbH, for financial support and Günter Langel and Dieter Preclik for their vision and trust in the endeavor of pursuing the graduate program at the University of Maryland.

## References

- [1] Hammond, W. E., *Space Transportation: A Systems Approach to Analysis and Design*, AIAA Education Series, AIAA, Reston, VA, 1999, pp. 407–413.
- [2] Meisl, C. J., “Life-Cycle-Cost Considerations for Launch Vehicle Liquid Propellant Rocket Engine,” 22nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Huntsville, AL, AIAA Paper 1986-1408, 16–18 June 1986.
- [3] Koelle, D. E., “Handbook of Cost Engineering for Space Transportation Systems,” TransCostSystems Rept. TCS TR-190, Ottobrunn, Germany, 2010, p. 36.
- [4] Parkinson, R. C., “The Hidden Costs of Reliability and Failure in Launch Systems,” *Acta Astronautica*, Vol. 44, Nos. 7–12, April–June 1999, pp. 419–424.  
doi:10.1016/S0094-5765(99)00093-4
- [5] Sackheim, R., “Overview of United States Space Propulsion Technology and Associated Space Transportation Systems,” *Journal of Propulsion and Power*, Vol. 22, No. 6, Nov.–Dec. 2006, pp. 1310–1332.  
doi:10.2514/1.23257
- [6] Wood, B. K., “Propulsion for the 21st Century—RS-68,” 38th Joint Liquid Propulsion Conference, Indianapolis, IN, AIAA Paper 2002-4324 July 2002.
- [7] Richards, S., “Liquid Rocket Engine Flight Certification,” *Space Transportation Propulsion Technology Symposium*, NASA Conference Publ. 3112, Vol. 3, 1990.
- [8] Emdee, J. L., “A Survey of Development Test Programs for Hydrogen Oxygen Rocket Engines,” 39th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, AIAA Paper 2001-0749, Jan. 2001.
- [9] Emdee, J. L., “A Survey of Development Test Programs for LOX/Kerosene Rocket Engines,” 37th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Salt Lake City, UT, AIAA Paper 2001-3985, July 2001.
- [10] Pempie, P., and Vernin, H., “Liquid Rocket Engine Test Plan Comparison,” 37th AIAA/ASME/SAE/ASEE Joint Propulsion, Salt Lake City, UT, AIAA Paper 2001-3256, July 2001.
- [11] 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,” NASA Marshall Space Flight Center, 2008.
- [12] Rachuk, V. S., “Design, Development, and History of the Oxygen/Hydrogen Engine RD-0120,” 31st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, San Diego, CA, AIAA Paper 1995-2540, July 1995.
- [13] O’Hara, K. J., “Liquid Propulsion System Reliability: ‘Design for Reliability,’” 25th ASME/SAE/ASEE Joint Propulsion Conference, Monterey, CA, AIAA Paper 1989-2628, July 1989.
- [14] Pugh, R. L., “Space Transportation Main Engine Reliability Demonstration Technique,” *Proceedings Annual Reliability and Maintainability Symposium*, IEEE Publ., Piscataway, NJ, 1993, pp. 173–180.
- [15] 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*, Vol. 2009, 2009, Paper 754896.  
doi:10.1155/2009/754896
- [16] Hamada, M. S., Martz, H. F., Reese, C. S., Graves, T., Johnson, V., 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*, Elsevier, New York, 2004, pp. 297–305.
- [17] 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*, Vol. 21, No. 4, 2006, pp. 514–531.  
doi:10.1214/088342306000000439
- [18] Modarres, M., Kaminskiy, M., and Krivtsov, V., *Reliability Engineering and Risk Analysis: A Practical Guide*, Taylor and Francis, Philadelphia, 1999, pp. 327–335.
- [19] Schwarz, W., Schwub, S., Quring, K., Wiedmann, D., Höppel, H. W., and Göken, M., “Life Prediction of Thermally Highly Loaded Components: Modeling the Damage Process of a Rocket Combustion Chamber Hot Wall,” Space Propulsion Conference, San Sebastian, Spain, Assoc. Aéronautique et Astronautique de France Paper 1841633, May 2010.
- [20] Lloyd, D. K., and Lipow, M., *Reliability: Management, Methods, and Mathematics*, Prentice-Hall, Upper Saddle River, NJ, 1962, Chap. 16.
- [21] Meisl, C. J., “Rocket Engine versus Jet Engine Comparison,” AIAA/SAE/ASME/ASEE 28th Joint Propulsion Conference and Exhibit, Nashville, TN, AIAA Paper 1992-3686, July 1992.
- [22] Worlund, A. L., Monk, J. C., and Bachtel, F. D., “NLS Propulsion Design Considerations,” AIAA Aerospace Design Conference, Irvine, CA, AIAA Paper 1992-1181, Feb. 3–6, 1992.
- [23] Hahn, T., “CUBA: A Library for Multidimensional Numerical Integration,” *Computer Physics Communications*, Vol. 168, No. 2, June 2005, pp. 78–95.  
doi:10.1016/j.cpc.2005.01.010
- [24] Bolstad, W. M., *Understanding Computational Bayesian Statistics*, Wiley, Hoboken, NJ, 2010, Chaps. 5, 6, 7, 8.
- [25] Hastings, W. K., “Monte Carlo Sampling Methods Using Markov Chains and Their Applications,” *Biometrika*, Vol. 57, No. 1, 1970, pp. 97–109.  
doi:10.1093/biomet/57.1.97
- [26] Chib, S., and Greenberg, E., “Understanding the Metropolis–Hastings Algorithm,” *American Statistician*, Vol. 49, No. 4, Nov. 1995, pp. 327–335.  
doi:10.2307/2684568
- [27] Neath, R. C., and Jones, G. L., “Variable-at-a-Time Implementations of Metropolis–Hastings” [online], March 2009, <http://arxiv.org/abs/0903.0664v1> [retrieved July 2011].
- [28] Fishman, G. S., *A First Course in Monte Carlo*, Duxbury, Belmont, MA, 2006, Chaps. 3, 6.
- [29] Gregory, P., *Bayesian Logical Data Analysis for the Physical Sciences*, Cambridge Univ. Press, Cambridge, England, U. K., 2005, Chap. 12.
- [30] Liu, J. S., *Monte Carlo Strategies in Scientific Computing*, Springer Science and Business Media, New York, 2008, Chaps. 3, 4, 5, 12.
- [31] Wasserman, G. S., *Reliability Verification, Testing, and Analysis in Engineering Design*, Marcel Dekker, New York, 2003.
- [32] McFadden, R. H., and Shen, Y., “An Analysis of the Historical Reliability of US Liquid-Fuel Propulsion Systems,” 26th AIAA/ASME/SAE/ASEE Joint Propulsion, Orlando, FL, AIAA Paper 1990-2713, July 1990.
- [33] Martz, H., and Waller, R., *Bayesian Reliability Analysis*, Wiley, New York, 1982.

- [34] Martz, H., and Waller, R., "The Basics of Bayesian Reliability Estimation from Attribute Test Data," Los Alamos Scientific Lab. Report UC-79P, Los Alamos, NM, 1976, pp. 1–17.
- [35] Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, New York, 1989, pp. 85–86.
- [36] Deb, K., *Multi-Objective Optimization using Evolutionary Algorithms*, Wiley, New York, 2008, pp. 289–313.
- [37] Biggs, R. E., *Space Shuttle Main Engine: The First Twenty Years and Beyond*, AAS History Series, Vol. 29, American Astronomical Soc., San Diego, CA, 2008, Chaps. 2, 4.
- [38] Williams, W. C., Himmel, S. C., Mobley, T. B., Reese, B. A., Suarez-Alfonso, E., and Weiss, R., "Report of the SSME Assessment Team," NASA, Washington, D.C., 1993.

E. Kim  
Associate Editor