Improved Verification for Aerospace Systems

Mark A. Powell
Attwater Consulting
P.O. Box 57702
Webster, TX 77598-7702
208-521-2941
attwater@aol.com

Abstract—Aerospace systems are subject to many stringent performance requirements that must be verified with low residual risk to the customer. This report investigates the improvements in verification planning and requirements validation made possible by using conditional approaches vice classical statistical procedures. The example used in this report to illustrate the results of these investigations is a proposed mission assurance requirement with its concomitant maximum acceptable verification risk for the NASA Constellation Program Orion Launch Abort System (LAS).

This report demonstrates the following improvements possible through use of conditional approaches: 1) verification plans are more achievable and feasible than those developed using classical statistical methods; 2) historical surrogate data can be used to validate proposed performance requirements; and, 3) historical surrogate data may be incorporated in verification planning to produce even less costly and more reasonable verification plans. The procedures presented in this report may produce similar improvements and cost savings in validation and verification for any performance requirement for any aerospace system. [1] [2]

TABLE OF CONTENTS

1. INTRODUCTION

Aerospace systems are subject to many stringent performance requirements. Verification assures that the risk that the as-built system does not satisfy these requirements is below some level acceptable to the customer. The customer upon accepting the verified aerospace system shoulders this residual risk. Verification planning and execution for such requirements, to achieve low levels of residual risk, is often difficult and costly. For many performance requirements, verification planning that uses classical statistical methods becomes nearly impossible without many difficult to defend assumptions. The customer may be subsequently forced to accept a much higher level of residual risk than preferred, perhaps unknowingly.

These concerns were realized for a proposed mission assurance requirement for the NASA Constellation Program, specifically for the Orion Launch Abort System (LAS).

1

[1] 978-1-4244-2622-5/09/$25.00 ©2009 IEEE
[2] IEEEAC paper #1445, Version 4, Updated January 8, 2009

This report presents the following investigations using as an example a proposed Orion LAS mission assurance requirement with a required maximum acceptable verification risk: 1) the comparison of verification planning based on classical statistical methods with difficult to defend assumptions, and verification planning using conditional approaches avoiding indefensible assumptions; 2) use of historical test and flight operations data from surrogate systems for validation of a mission assurance requirement; and, 3) use of historical test and flight operations data from surrogate systems in verification planning for a mission assurance requirement.

2. METHODS

Verification planning is essentially an inverse hypothesis testing process. Hypothesis testing consists of three steps. The first is to design and perform an experiment to obtain data about a hypothesis. The second is to process the observed data using a statistical procedure. The third is to decide, based on the results of statistical processing of the observed data from the experiment, whether to reject the hypothesis or retain it for further consideration.

In verification planning and execution, these steps are performed in a rather convoluted order. The first is to design an experiment to obtain data about a hypothesis for the verification plan (to be performed later in actual verification execution). The second is to identify the desired outcome, rejection of the hypothesis or its retention. The third is then to determine which data sets observable from this planned experiment, that when subjected to a statistical procedure, would clearly lead to the decision that the desired outcome was achieved.

The hypotheses considered for verification planning always relate to whether or not an as-built system satisfies a requirement. The decision about the hypothesis becomes whether to accept the as-built system or not. In verification planning, the objective is to find the data sets possible from the experiment, that when performed, will make this decision clear. These defined data sets establish the success criterion for the verification plan. In verification execution, the experiment is performed. The data sets are then compared to those identified in the verification plan, and the decision is made based on the data actually observed.

Both for hypothesis testing and verification planning, the decision on the hypothesis is based on some specified acceptable level of risk of making an incorrect decision. For verification planning, this is the risk of accepting the as-built system when the requirement is not truly satisfied, even though the experiment produced the requisite data indicating successful verification. This is the residual risk that the customer shoulders with successful verification.

Another important factor in verification planning is the risk of the experiment failing to produce a data set defined for successful verification when the requirement is truly satisfied in the as-built system. Clearly, it is very desirable that the experiment be designed in verification planning such that this risk is very low. This is especially important for the contractor or supplier if verification results are used for contract fulfillment. Ideally, the probability of verification plan success when the as-built system truly satisfies the requirement (the complement of this risk) should exceed 90%.

The experiments or tests used in verification of a mission assurance requirement for an aerospace system exercise the system and observe whether the system accomplished its mission or not. The data then are a collection of Bernoulli trial results [1], i.e., numbers of observed successes and failures of the system to accomplish its mission in a testing scenario. The verification plan for a mission assurance requirement thus defines the numbers of observed successes and failures of the mission that, when subjected to a statistical procedure, enables the decision whether the mission assurance requirement is satisfied or not with less than some maximum acceptable residual risk.

In practice, a verification requirement for an aerospace system mission requirement may state the maximum acceptable verification risk in terms of a required *confidence*. The complement of the confidence will be the maximum acceptable verification risk. The use of this term often leads to a misinterpretation that a *confidence interval* procedure from classical statistics is required to be used in verification. As will be seen later in this report, this misinterpretation can be very costly.

Comparison: Classical versus Conditional Approaches

Classical Statistical Procedures—Classical statistical procedures operate on the principle of developing an experiment to produce evidence that will support the *falsity* of a hypothesis [1]. The concept is simple: if the hypothesis is truly false, then the data obtained from an experiment about the hypothesis, when processed by a statistical procedure or recipe, should indicate that the hypothesis is false. Falsification of a hypothesis is always possible if it is truly false, where proof that it is true when truly true may not be. If the results of the statistical processing of the data observed from the experiment do not support the falsity of the hypothesis, then the hypothesis remains under consideration as *potentially* being true. Classical statistical procedures never infer the truth of the hypothesis, only the falsity of the hypothesis, and the selection of the hypothesis for the procedure can be important in the design of the experiment. As a result, classical statistical procedures inherently incorporate conservatism. Inability to falsify a hypothesis from a single experiment suggests nothing about whether it is true. Both classical hypothesis testing and the

equivalent confidence interval testing recipes adhere to these fundamental precepts.

A common misconception is that the confidence interval of classical statistics contains the probability that the hypothesis is true given the observed data [3]. Classical statistical recipes that produce confidence intervals are based on a *significance* value, the complement of which is informally referred to as the confidence, most likely due to the name Karl Pearson gave to the interval. The probability contained in the confidence interval is always greater than the complement of the significance value that is used to develop the confidence interval [4]. The actual probability contained in the classical confidence interval may be calculated from the observed data. In [4], a verification plan for a reliability requirement was developed based on a requirement that the verification provide 90% confidence that the requirement was satisfied. The 90% confidence interval (based on a significance of 0.1) for the classical statistical test used to verify this requirement actually contained a 99.92% probability that the requirement was satisfied when the requisite data were observed. This resulted in an actual residual verification risk of 0.08%, far below the required 10% (complement of the 90% confidence required). This unrequired and essentially invisible risk reduction more than doubled the cost of the verification. The inherent conservatism in using classical statistical procedures in verification planning always comes with a cost.

Conditional Approaches—Conditional approaches operate on the principle of developing an experiment to obtain data to support the truth of the hypothesis. This is a much stronger proposition than the retention of a hypothesis for consideration as true solely because data was not observed to indicate that it was false. Conditional statistical procedures infer from the data observed in the experiment the probability that the hypothesis is true. The residual risk of the hypothesis not being true for the data observed is the complement of this inferred probability.

Where classical statistical procedures include a host of recipes for different problems, conditional approaches are all based on one simple recipe, Bayes' Law [5]. Equation 1 provides this simple recipe showing the uncertainty relationship between the observed data and a hypothesis h_0.

$$P(h_0 \mid data) \propto L(data \mid h_0) P(h_0) \tag{1}$$

The term to the left of the proportion in equation 1 is called the posterior probability and is the probability that the hypothesis h_0 is true given the data. The hypothesis h_0 usually concerns the values of the parameters of a probability distribution model, from which the residual risk may be calculated. The first term to the right of the proportion is called the likelihood, the probability of obtaining the data or information actually observed given hypothetical values of the parameters of the probability distribution model. For data consisting solely of observed events, it is the same likelihood function used in calculating the maximum likelihood estimates. The likelihood may also include terms for probabilities of observing censored data, which are not handled well if at all by classical statistical recipes. Events and observations about any phenomenon are generally independent. As such, the likelihood is usually a product of the probabilities of each event or observation given hypothetical values of the parameters. For data consisting of only actual observed events (no censored data) where the probability distribution model is a continuous model, the likelihood is a product of evaluations of the probability density function for the model given hypothetical values of the parameters.

The second term to the right of the proportion in equation 1 is called the prior probability, the joint probability for the probability distribution model parameters before processing the data. In general, selecting a model for the joint uncertainty for the parameters of any model for any phenomenon is problematic. For most probability distribution model parameters, there is no real physical meaning for these parameters, and hence no physics to govern any uncertainty about the joint relationship. The solution in this case is to use a joint model that actually reflects this lack of information about these parameters. These models for the prior are called variously non-informative priors [5], ignorance priors [3], Jeffrey's priors [5], maximum entropy priors [6], and reference priors [7]. As discussed in each reference, these uncertainty models provide the least possible information about the parameters in the inference, and thus provide maximum objectivity to the statistical inference of the posterior model. [7] provides a list of these models for most of the commonly used probability distribution models. Using such a prior provides maximum objectivity in the inference, and indefensible assumptions are avoided.

Because of the linearity of the derivative operator for real world verification problems, equation 1 can be formulated with the derivatives of the posterior and prior probabilities [5], yielding Bayes' Law in terms of probability density functions as in equation 2.

$$pd(h_0 \mid data) \propto L(data \mid h_0)\, pd(h_0) \tag{2}$$

The posterior probability density function can always be integrated to produce probability values.

Verification is fundamentally used to make a decision whether or not to believe that the requirement is truly satisfied in the as-built system. Subsequently, for many aerospace systems, the decision for contract fulfillment is based on successful verification. Decision theory [8] exclusively uses conditional approaches. Verification

planning seems then to be a natural application on which to apply conditional approaches.

A Useful Model for Mission Assurance Requirement Verification Planning—Classical statistical procedures and conditional approaches both start by selecting a useful probability distribution model that could generate the observed data. The selection of this model should where possible be based on the physics or mathematics of the phenomenon that will generate the data. For verification of Orion LAS mission assurance, the numbers of observed mission successes and failures form a set of Bernoulli trial results, the uncertainty for obtaining this set is naturally modeled using the binomial probability distribution model. Equation 3 provides this binomial model where $\tilde{n}_s$ is the number of successes observed, $\tilde{n}_f$ is the number of failures observed, and θ is the mission assurance.

$$P\left(\tilde{n}_s, \tilde{n}_f \mid \theta\right) = \frac{\Gamma\left(\tilde{n}_s + \tilde{n}_f + 1\right)}{\Gamma\left(\tilde{n}_s + 1\right)\Gamma\left(\tilde{n}_f + 1\right)} \theta^{\tilde{n}_s} \left(1-\theta\right)^{\tilde{n}_f} \qquad (3)$$

Mission Assurance Requirement Verification Planning using Classical Statistical Procedures—Verification planning that is based on classical statistical procedures uses the null hypothesis that *the requirement is **not** satisfied in the as-built system*. Thus, the verification is considered successful if the experiment produces data that supports rejection of this null hypothesis; i.e., evidence is observed to infer that it is ***not*** true that the requirement is ***not*** satisfied. Despite the double negative in the preceding statement, this is not equivalent to an experiment producing evidence that the requirement is in truth satisfied. If the alternate hypothesis were selected as the null hypothesis in verification planning using classical statistical procedures, i.e. that *the requirement is satisfied in the as-built system*, then the verification would be considered successful if evidence were ***not*** observed to reject the hypothesis. In this case, failure to observe evidence to reject the hypothesis that the requirement is satisfied in the as-built system is an even weaker proposition. These subtleties escape many systems engineers in the aerospace industry responsible for verification.

The classical statistical procedure suitable for use in verifying an aerospace system mission assurance requirement, using as data the numbers of observed successes and failures of the mission, is the binomial test [2]. One way to implement the binomial test for verification planning is to construct a one-sided confidence interval that ranges from the required mission assurance level to unity for a given significance level (usually derived as the complement of a specified confidence). The significance level is traditionally set to the value of the maximum acceptable verification risk to be shouldered by the customer. An estimator for mission assurance is calculated from the data observed from the experiment, in this case the numbers of mission successes and failures. If this estimate of mission assurance produced by the binomial test recipe falls inside this confidence interval, the hypothesis that the required mission assurance has ***not*** been satisfied by the as-built system is rejected, and the verification is considered successful. The objective of verification planning using the binomial test is to determine the numbers of successes and failures that if observed would produce a mission assurance estimate that will indeed fall into this confidence interval.

Behind the binomial test recipe is an assumption that the central limit theorem is valid for the data and experiment. This is apparent in the recipe through the use of critical points selected from the standard normal distribution. The central limit theorem is based on some key assumptions as well. These assumptions are well documented [1], [3], [5], [7], [8], and are: 1) the data are all independent and identically distributed; 2) these probability distributions have finite means and variances; and, 3) the data numbers approach infinity. For most verification planning, this compounding of assumptions upon assumptions may be difficult to defend, especially if the verification is very expensive. It may also be easily demonstrable that these assumptions are false. Seldom can experiments used for verification of aerospace systems producing very large numbers of data be cost feasible. Experiments are difficult to develop that produce data that are truly independent and identically distributed. And, it is not unusual for the physics of the phenomena producing the data to indicate that the probability distribution for the data has a mean or variance that is undefined, i.e., the data are Cauchy or Lorentz distributed. The inherent conservatism in classical statistical procedures often compensates for the limited validity of these assumptions, but with a cost. Where verification cost is a programmatic issue for an aerospace system development, these assumptions may be difficult to defend.

Mission Assurance Requirement Verification Planning using Conditional Approaches—The operant hypothesis for verification planning using conditional approaches is that *the requirement is satisfied in the as-built system*. Using conditional approaches, if a verification plan, developed for a maximum acceptable verification risk of 5%, is successful, then there is a 95% probability that the requirement is truly satisfied in the as-built design, based on the test data. This information resonates particularly well with managers making decisions to accept an aerospace system and authorize contractor payment for it.

To develop a verification plan using conditional approaches, the posterior density model must be formed for the uncertainty about the hypothesis that the requirement is indeed satisfied in the as-built system. Equation 3, the binomial probability, directly provides the likelihood for verification of a mission assurance requirement. The binomial model provides the probability of obtaining the observed data $\tilde{n}_s$ and $\tilde{n}_f$ as a function of the mission assurance θ, the parameter of the binomial model. The kernel for the uncertainty model that provides the ultimate

objectivity for the posterior based on using the binomial model [7] is presented in equation 4.

$$pd(\theta) \propto \theta^{-\left(\frac{1}{2}\right)}(1-\theta)^{-\left(\frac{1}{2}\right)} \tag{4}$$

Equation 5 presents the kernel of the posterior density model obtained via the product of equation 3 (the likelihood) and the maximum objectivity prior in equation 4.

$$pd(\theta \mid \tilde{n}_s, \tilde{n}_f) \propto \theta^{\left(\tilde{n}_s - \frac{1}{2}\right)}(1-\theta)^{\left(\tilde{n}_f - \frac{1}{2}\right)} \tag{5}$$

This posterior density model is immediately recognizable as the kernel of a beta probability density model, and is presented in its full form in equation 6 where **B(·,·)** is the beta function.

$$pd(\theta \mid \tilde{n}_s, \tilde{n}_f) = \frac{\theta^{\left(\tilde{n}_s - \frac{1}{2}\right)}(1-\theta)^{\left(\tilde{n}_f - \frac{1}{2}\right)}}{\mathbf{B}\left(\tilde{n}_s + \frac{1}{2}, \tilde{n}_f + \frac{1}{2}\right)} \tag{6}$$

The probability that the mission assurance requirement is satisfied, and hence the verification risk, given that $\tilde{n}_s$ and $\tilde{n}_f$ are observed, are obtained from an integral of the posterior density model in equation 6. Equation 7 presents this integral to calculate the probability that the true mission assurance the mission assurance θ exceed the required mission assurance θ_R with a maximum acceptable verification risk V_R.

$$P(\theta \geq \theta_R \mid \tilde{n}_s, \tilde{n}_f) = \int_{\theta_R}^{1} \left(\frac{\theta^{\left(\tilde{n}_s - \frac{1}{2}\right)}(1-\theta)^{\left(\tilde{n}_f - \frac{1}{2}\right)}}{\mathbf{B}\left(\tilde{n}_s + \frac{1}{2}, \tilde{n}_f + \frac{1}{2}\right)} \right) d\theta \geq 1 - V_R \tag{7}$$

Given the mission assurance requirement θ_R, and the maximum acceptable verification risk V_R, the verification plan developed using conditional approaches finds the limiting values of $\tilde{n}_s$ and $\tilde{n}_f$ that satisfy the inequality on the right of equation 7.

Validation Considering Historical Data

Quite often, several systems very similar to the aerospace system of interest have been designed, tested, and fielded. These surrogate systems, if fielded and flown, were verified to satisfy their performance requirements within the maximum acceptable verification risk at acceptable costs for design, implementation, and verification. The mission assurance for these surrogate systems, as verified and as achieved in operations, establishes the bounds of validity for the mission assurance requirement with the specified maximum acceptable verification risk for the aerospace system of interest.

It is not necessary to know what the mission assurance requirements were for these surrogate systems, or what the maximum acceptable verification risks were. From the surrogate systems' test data, the numbers of test successes and failures, statistical procedures may be used to establish the actual verified mission assurance parameterized as a function of maximum acceptable verification risk. This parametric relationship was achievable and cost feasible, else these surrogate systems would not have been fielded, whether verified or not.

The mission assurance actually achieved in the as-built surrogate system in operations may be different from the level that was verified or verifiable, usually higher. The same statistical procedures used to establish the verified mission assurance parameterized as a function of maximum acceptable verification risk may be used with the surrogate systems actual operations or flight data, the numbers of successful and failed flights. This establishes the achieved operational mission assurance parameterized as a function of bounding uncertainty levels. These levels of mission assurance performance were apparently achievable and cost feasible in design and manufacture, else these systems would have never been fielded and operated.

These parametric relationships for surrogate systems can be used to validate the achievability and feasibility of a proposed mission assurance requirement for an aerospace system of interest, both for verification planning and operations performance. Both classical statistical procedures and conditional approaches may be used to analyze the surrogate test and flight data to establish these parametrics.

Validation based on Classical Statistical Procedures—The classical statistical procedure used to develop this parameterization again relies upon the binomial test, with its difficult to defend assumptions. For a given set of data, the achieved mission assurance for a given significance level using the binomial test recipe is just the lower limit of the confidence interval computed from the data. Typically in practice, this significance level is set to the specified maximum acceptable verification risk, or complement of the confidence if so specified. For a range of hypothetical maximum acceptable verification risk levels between 0.001 and 0.5, Figure 1 shows the verified or verifiable mission assurance parametric resulting from using the binomial test for a hypothetical aerospace system with test data consisting of eight mission successes and one mission failure (significance levels set to the risk levels).

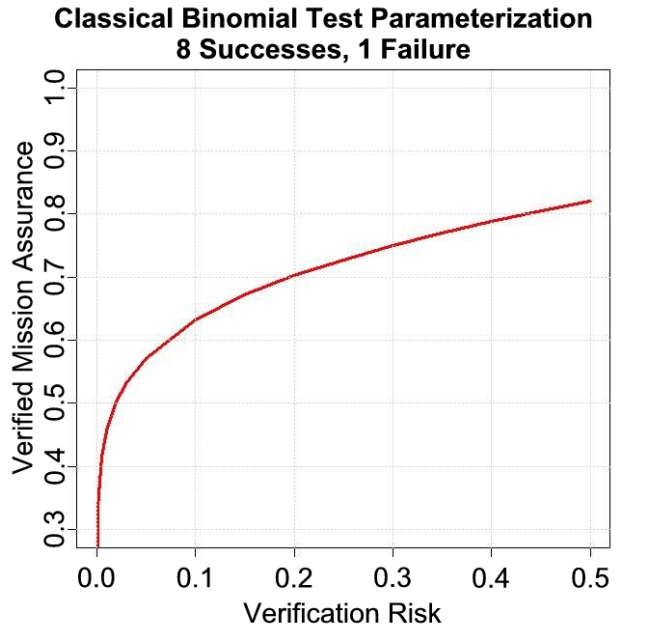

Figure 1 – The classical binomial test can be used to parameterize verified mission assurance as a function of maximum acceptable verification risk level for a surrogate system.

The solid red line in Figure 1 represents the mission assurance requirement verified to be satisfied by the as-built surrogate system, at the verification risk level on the abscissa, when the test data were eight mission successes and one mission failure processed, using the classical binomial test. A point selected on the red solid line in Figure 1 at the abscissa value of 0.1 has the following properly stated interpretation: *the test data indicates that the as-built surrogate system did* ***not*** *satisfy a mission assurance requirement* ***below 0.64*** *at a 0.1 significance*. This properly stated interpretation might make a customer's decision to accept the as-built system and pay the contractor somewhat difficult. Nota bene: the double negative in the previous statement does NOT equate to the interpretation that the *as-built surrogate system does satisfy a mission assurance requirement of* ***at least 0.64*** *with no more than* ***10% residual risk***.

Validation based on Conditional Approaches—The conditional approach presented earlier in this report for verification planning without using any indefensible assumptions may also be used to parameterize achieved mission assurance performance at verification risk levels. To do this, equation 7 is evaluated for a given mission assurance level θ_R with surrogate system mission successes $\tilde{n}_s$ with the allowed number of failures $\tilde{n}_f$, providing the maximum acceptable level of verification risk V_R. For a range of hypothetical maximum acceptable verification risk levels between 0.001 and 0.5, Figure 2 shows the verified mission assurance thus obtained, compared with that obtained using the binomial test, for a hypothetical aerospace system with test data of eight mission successes and one mission failure.

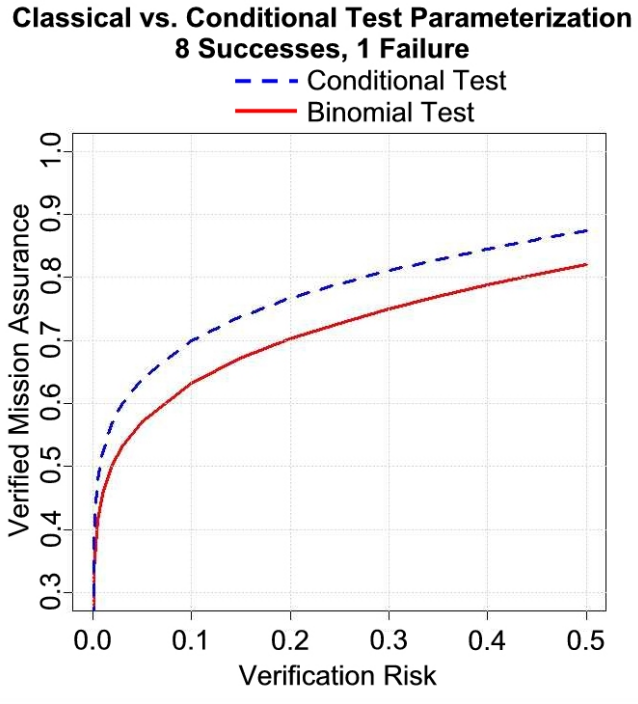

Figure 2 – Conditional approaches, without using indefensible assumptions, indicate levels of verified mission assurance superior to those obtained using the classical binomial test.

The surrogate system given its test data was verified to achieve at least the mission assurance level as indicated by the dashed blue line in Figure 2, at the verification risk on the abscissa when using conditional approaches. A point selected on the dashed blue line in Figure 2 at the abscissa value of 0.1 has the following properly stated interpretation: *there is a 90% certainty that the as-built surrogate system satisfied a mission assurance requirement of* ***at least 0.7*** *based on the test data*. An equivalent interpretation at the same point is: *there is no more than a 10% risk that the as-built surrogate system did not satisfy a mission assurance requirement at least 0.7 based on the test data*. These statements do make the customer's decision to accept the as-built system and pay the contractor rather straightforward.

Note the conservatism inherent using the classical statistical procedures; everywhere in Figure 2, the verified mission assurance for the dashed blue line is higher than for the solid red line for all verification risk levels. This conservatism using the binomial test means that more than eight test successes would be needed using the binomial test to achieve the mission assurance level produced using conditional approaches for every verification risk level. These additional tests mean additional costs.

Such investigations considering historical data for surrogate aerospace systems allow validation of the reasonableness, feasibility, and achievability of a proposed mission assurance requirement at the maximum acceptable verification risk for the system of interest. The mission assurance parametrics for the fielded surrogate system in test and operations, calculable from the historical data, were achieved and verifiable; else it would not have been fielded. The cost of the verification of the parametric performance must have been acceptable; else the verification would not have been performed. This information on surrogate systems yields invaluable insight into the validity of the performance requirements and verification risk, as well as into the design and verification for the system of interest.

As will be seen later in this report, flight data quantities are usually much larger than test data quantities, and better mission assurance performance is usually achieved in actual mission flights than in testing at the same level of risk. Actual flight mission assurance may be much more stringent than can be cost feasible to verify.

Use of Historical Data in Verification Planning

Quite often, the similarities of surrogate aerospace systems to the system of interest extend to commonalities in design and manufacturing standards, and even to the manufacturers themselves. In these cases, it is reasonable to expect that the system of interest will exhibit similar properties and performance as the historical systems. While classical statistical procedures offer no means to take advantage of surrogate test data in verification planning, conditional approaches do.

If in equation 6, $\tilde{n}_s$ and $\tilde{n}_f$ are the surrogate test data, then the left hand side represents a reasonable model of uncertainty for mission assurance for the aerospace system of interest, *prior* to its design and manufacture. Thus, rather than use a prior of maximum objectivity as in equation 4 to form the posterior density model for verification planning, equation 6 using the surrogate test data is used as the prior density model. Multiplying equation 6 with the likelihood of equation 3 with number of test successes n_s with the allowed number of failures n_f provides a posterior density model for a verification plan that takes advantage of the surrogate test data. The plan to verify a mission assurance θ_R requirement with a maximum acceptable verification risk V_R that takes advantage of the surrogate test data $\tilde{n}_s$ and $\tilde{n}_f$ is obtained by solving equation 8 for n_s and n_f.

$$
\begin{aligned}
&P\left(\theta \geq \theta_R \mid n_s, n_f, \tilde{n}_s, \tilde{n}_f\right) \\
&\quad = \int_{\theta_R}^{1} \left(\frac{\theta^{n_s+\tilde{n}_s-\frac{1}{2}}\left(1-\theta\right)^{n_f+\tilde{n}_f-\frac{1}{2}}}{\mathrm{B}\left(n_s+\tilde{n}_s+\frac{1}{2}, n_f+\tilde{n}_f+\frac{1}{2}\right)} \right) d\theta \qquad (8) \\
&\quad \geq 1-V_R
\end{aligned}
$$

The posterior density model thus obtained using equation 6 as a prior for verification planning, taking advantage of the surrogate system test data $\tilde{n}_s$ and $\tilde{n}_f$, is the integrand in equation 8, another beta probability density model. As will be seen later in this report, the numbers of tests successes and failures allowed obtained from equation 8 can be considerably smaller than if the surrogate system test data are ignored.

A caveat for this process is appropriate at this point for using surrogate data for both validation and verification. The determination of suitability of surrogate system test data, as representative of the system of interest, requires serious and thorough engineering analysis and judgment. Inappropriate application of data always invalidates a statistical process.

3. EXAMPLE

Description of the Orion Launch Abort System

The Orion Launch Abort System (LAS) is a rocket-based system mounted atop the Orion Crew Exploration Vehicle [9]. The Orion LAS will be the first such system to be employed by NASA since the Apollo program.

The primary function of the Orion LAS is to provide crew survival should the Ares I launch vehicle explosively malfunction on the pad or during launch. The Orion LAS accomplishes this function by firing solid rockets that safely take the Orion capsule containing the astronauts away from the malfunctioning Ares I launch vehicle. The Orion crew is not expected to survive should the Orion LAS not function when needed. Successful operation of this system will be required on every Orion flight, regardless of whether the system is merely jettisoned, or used for pad or launch abort. Therefore, the Orion LAS must be designed to achieve a high level of mission assurance performance, which must be verified with a low level of verification risk. As such, the Orion LAS will have a very stringent mission assurance requirement to assure crew survival should an Ares I launch vehicle malfunction.

Figure 3 shows the Orion LAS attached to the Orion capsule at the very top.

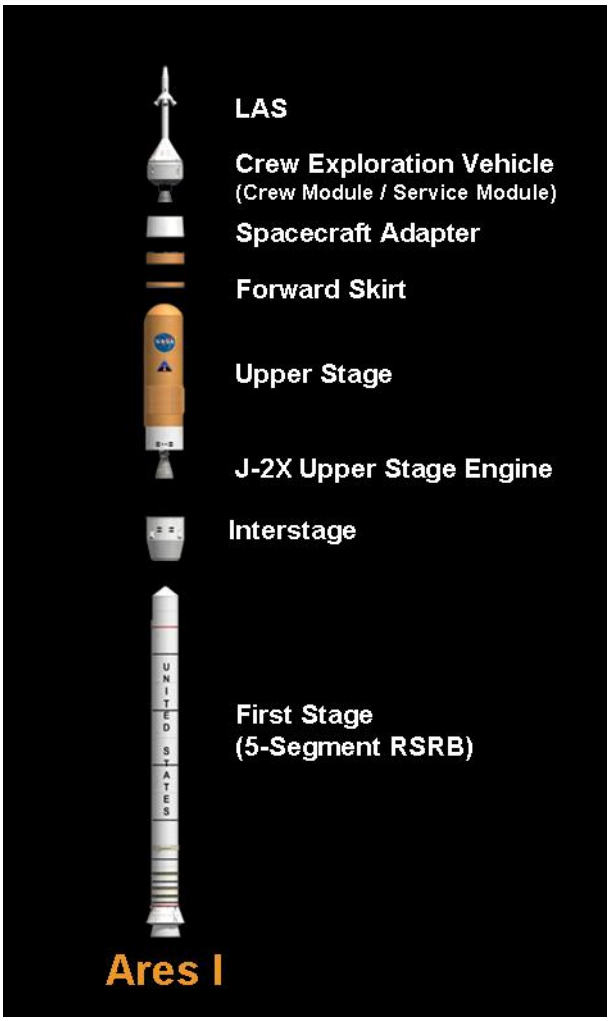

Figure 3 – The Constellation Ares I stack, showing the Orion Launch Abort System atop the Orion Crew Exploration Vehicle

Verification of this mission assurance requirement at a reasonable level of risk to be shouldered by NASA can be very expensive. A proposed mission assurance requirement considered initially for the Orion LAS was 0.9973 with a maximum acceptable verification risk of 10%, stated as a *90% confidence*.[3] These proposed requirements indicate that the maximum acceptable risk of Orion LAS mission failure is 0.0027, or 0.27%, and that the maximum acceptable risk of exceeding this value after successful testing is 10%. Acceptance of the Orion LAS with a successful verification means that NASA accepts no more than a 10% risk that the risk of loss of mission and crew due to failure of the Orion LAS exceeds 0.27%.

Historical Ground Test and Flight Data for Surrogate Systems Similar to the Orion LAS

Forty-one different systems [11] were identified as being sufficiently similar to the proposed design for the Orion LAS to merit consideration for validation and verification planning for the Orion LAS. Appendix A of [11] lists these surrogates with relevant descriptive data as well as ground test and flight mission assurance data. Appendix B of [11] describes in detail the in-depth engineering analysis involved in selecting these systems as suitable surrogates for the Orion LAS. All 41 of these surrogate systems were fielded. Numbers of successful and failed ground tests as well as numbers of successful and failed flights were provided. Table 1 lists the aggregation of the ground test and flight results data from these 41 surrogate systems.

Table 1: Aggregate Surrogate Data Considered for Validation and Verification Planning for the Orion LAS

Data Type	$\tilde{n}_s$	$\tilde{n}_f$
Ground Tests	263	5
Flights	4381	8
Totals	5094	13

The numbers of successes in column two of Table 1 reflect the numbers of times that the surrogate systems successfully accomplished their intended function in the launch mission in ground test and/or flight. The numbers of failures in column three of Table 1 reflect the numbers of times the surrogate systems were fired and did not accomplish their intended function in the launch mission in ground test and/or flight. The 41 surrogate systems averaged fewer than seven tests, with only nine using more than 10 tests, and with one using a maximum of 39 tests. Only eight of the surrogate systems had more than 100 flights, with one having 1,888 flights.

It is important to note that none of these surrogate systems had construction and performance characteristics identical to the proposed Orion LAS design. There indeed is a wide variety of surrogate systems' designs and performances listed in Appendix A of [11] that broadly bracket the current design concepts for the Orion LAS. The engineering analysis concluded that design and manufacturing for solid rockets is a relatively mature art, conducted by a very small number of firms in the industry, and that the Orion LAS would be designed, manufactured, and tested, using

8

[3] The term "confidence" is subject to misinterpretation between managers and statisticians, and thus between managers and verification planners. For the Constellation program, NASA Associate Administrator Scott "Doc" Horowitz defined the term as "… a calculation of the probability of performing a certain task over a given time within a specific budget (percent chance that a given project will cost an indicated amount or less)" [10]. The precedent this definition establishes for the Constellation Program relevant to verification of performance requirements is that a "90% confidence" is defined as a 90% probability that the performance requirement is satisfied with a successful verification.

sufficiently similar standards and processes as used for the surrogate systems selected.

4. RESULTS

The methods discussed in section 2 of this report are employed for the proposed Orion LAS mission assurance requirement with the stated maximum acceptable verification risk. Verification plans developed with both classical and conditional approaches are compared. Validation of the Orion LAS mission assurance requirement is performed using historical surrogate data. And, a cost-feasible verification plan is developed for a valid Orion LAS mission assurance requirement by incorporating historical surrogate ground test data.

Comparison: Orion LAS Mission Assurance Verification Plans using Classical and Conditional Approaches

Table 2 provides the minimum numbers of successes for the specified number of allowed failures that will define successful verification of the proposed Orion LAS mission assurance requirement as developed using classical statistical procedures. These data numbers will produce an estimate of mission assurance that will fall in the rejection region for a 90% *confidence interval* for the binomial test. Table 2 also provides in column three the associated probability of verification plan success given that the mission assurance requirement is truly satisfied in the as-built design.

Table 2: Data Required for Successful Verification Based on the Binomial Test

n_s	n_f	$P(V_{success})$
852	0	9.88%
1439	1	14.90%
1968	2	15.49%
2470	3	15.63%

To verify this proposed Orion LAS mission assurance requirement, with a 90% *confidence interval*, the numbers of tests required in Table 1 would be prohibitively expensive, even allowing no failures. Recall that the largest number of tests among the surrogate systems was only 39, and the average number of surrogate tests was less than 7. Column three in Table 1 indicates that even if the proposed Orion LAS mission assurance requirement of 0.9973 were satisfied in the as-built design, the probability of obtaining the specified numbers of successes with no more than the allowed number of failures is terribly low. If contract fulfillment depends on successful verification of this requirement, then column three provides the probability that the contractor would get paid if they Orion LAS they designed and built satisfied the proposed mission assurance requirement. There is always a trade between verification cost and the values in column 3. Contractors supplying the Orion LAS to NASA will generally want much higher values than presented in column 3, especially if contract payment depends on successful verification.

Table 3 illustrates that improvements in test numbers result from verification planning using conditional approaches over using classical statistical procedures in numbers of tests required and probability of verification success.

Table 3: Comparison of Data Required for Successful Verification Obtained using Classical and Conditional Procedures in Verification Planning

Classical (from Table 2)		**Conditional**		n_f
n_s	$P(V_{success})$	n_s	$P(V_{success})$	
852	9.88%	501	25.81%	0
1439	14.90%	1156	31.75%	1
1968	15.49%	1707	36.57%	2
2470	15.63%	2221	40.58%	3

The improvements in verification planning using conditional methods as displayed in Table 3 are rather dramatic. However, too many tests are still required for the verification plan to be cost feasible, and the probabilities of verification plan success do not approach the desirable range above 90%. The expense of testing 501 Orion LAS is just not consistent with a system that may be used in fewer than 100 missions. Even if the expense were reasonable, there is still an almost 75% probability that the verification plan would fail if the stringent 0.9973 mission assurance requirement were truly satisfied in every as-built Orion LAS (from column four in Table 3).

The conservatism of using classical statistical procedures for verification plan development (Table 3, comparing columns 1 and 3) is readily apparent, and significantly increases the expense of verification and reduces the probability of verification success.

Orion LAS Mission Assurance Requirement Validation based on Historical Data

The historical ground test and flight data for the 41 surrogate systems discussed in section 3 of this report were processed with the conditional procedure in section 2 to parameterize achieved mission assurance as a function of residual risk level (verification risk in the case of the ground test data). Figure 4 illustrates for each surrogate system the achieved risk levels at the proposed Orion LAS mission assurance requirement of 0.9973, calculated from both ground test data and flight data.

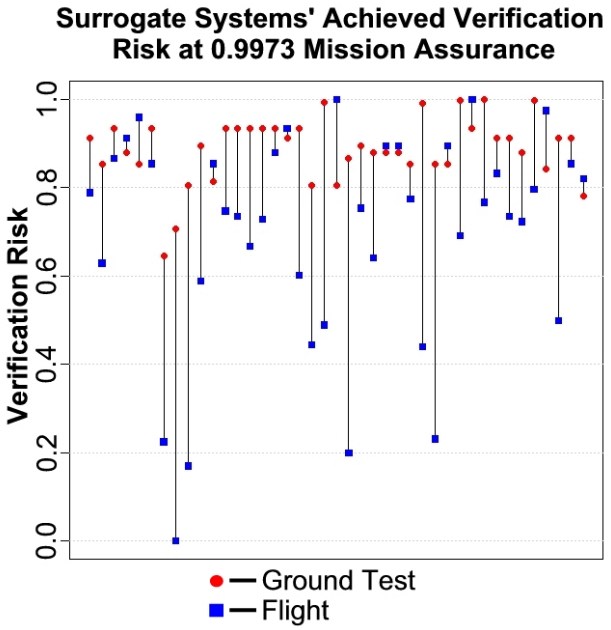

Figure 4 – Achieved risk levels at a mission assurance level of 0.9973 for the 41 surrogate systems only fell below the 10% maximum acceptable level in one case, and that from actual flight data.

In Figure 4, the vertical lines merely connect the achieved verification risk based on the ground test data with that based on the flight data for each surrogate system. None of these surrogate systems achieved mission assurance performance of 0.9973 verified within a 10% residual risk. Only one of the surrogate systems achieved less than 10% risk for a mission assurance of 0.9973 with actual flights. The lowest achieved verification risk for ground tests for the mission assurance of 0.9973 was more than 60%, hardly a level acceptable by most aerospace system customers. For most of these surrogate systems, there was more flight data than ground test data. There are 11 surrogate systems in Figure 4 for which the achieved risk using flight data was higher than that using ground test data. In five of these cases, there were fewer flights than ground tests, and higher achieved risk should be expected with fewer data. For the remaining six surrogate systems with higher flight than ground test achieved risk, all of the ground tests attempted were successful and there were one or more failures observed in actual flights, albeit with larger numbers of flights than tests. A single failure in a set of successful Bernoulli trials can dramatically increase the achieved risk, and again this should be expected. Tables 2 and 3 demonstrate this effect by virtue of the dramatically increased numbers of successes needed to assure the same level of verification risk when failures occur. Based on figure 4, the proposed Orion LAS mission assurance requirement of 0.9973 may not have been achieved widely in industry, and rarely verified at 10% residual risk.

Figure 5 illustrates for each surrogate system the achieved mission assurance at the proposed Orion LAS verification risk of 10%, calculated from both ground test data and flight data.

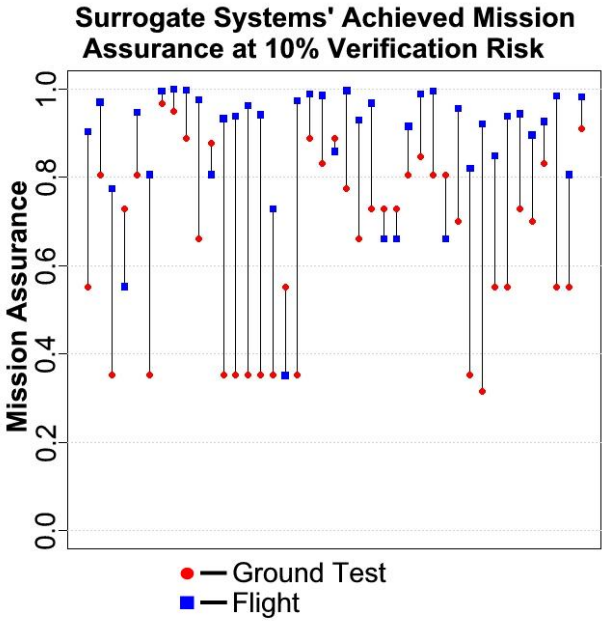

Figure 5 – Achieved mission assurance for the 41 surrogate systems only exceeded 0.9973 at a 10% verification risk in one case, and that from actual flight data.

In Figure 5, the vertical lines merely connect the achieved mission assurance based on the ground test data with that based on the flight data for each surrogate system. While the scale of Figure 5 is too coarse to demonstrate this, none of these surrogate systems achieved a mission assurance of 0.9973 verified at a 10% risk level based on the ground test data. Only two surrogate systems achieved verification (based on ground test data) of a mission assurance level above 0.9 at 10% risk. While many of the surrogate systems achieved mission assurance levels above 0.9 at a 10% risk considering actual flight data, only one achieved a mission assurance above 0.9973 at a 10% risk. This surrogate system had 26 test successes out of 26 tests, and 1880 flight successes of 1888 flight attempts.

Figures 4 and 5 overall may provide some significant insight into the state of the art for rocket design and manufacture. While the 10% maximum acceptable verification risk level is not an unusual verification risk value, a mission assurance requirement of 0.9973 may be rare. Better than half of the 41 surrogate systems in figure 5 achieved mission assurance in ground tests better than 0.7 with no more than 10% verification risk, and only 15 exceeded an achieved mission assurance of 0.8.

Considering the discussion in section 3 of this report concerning how these surrogate systems broadly bracketed the performance and design characteristics of the Orion LAS, it is illuminating to investigate the aggregate mission assurance performance. Table 4 provides the achieved verification risk for a mission assurance level of 0.9973 and the achieved mission assurance level for a 10% verification risk for the aggregated data in Table 1.

Table 4: Achieved Mission Assurance and Verification Risk based on Aggregate Surrogate Systems Data

Data Type	$\tilde{n}_s$	$\tilde{n}_f$	Achieved Risk for 0.9973 Mission Assurance	Achieved Mission Assurance at 10% Risk
Ground Tests	263	5	99.97%	0.9680
Flights	4381	8	7.19%	0.9974
Totals	5094	13	43.28%	0.9964

The parameters in the rightmost two columns of Table 4, as calculated from the flight data, suggest that systems similar to the proposed Orion LAS generally perform in flight with mission assurance slightly better than the proposed required level of 0.9973, with a no more than a 10% risk. However, the calculations using the aggregate ground test data suggest that surrogate systems were verified at lower mission assurance levels at 10% verification risk. Considering the aggregate data from these surrogate systems, it appears that if a mission assurance of 0.97 is verified at 10% risk, then better mission assurance performance, perhaps close to the proposed Orion LAS 0.9973 level, can be achieved in actual flight at less than a 10% risk. Thus, an ambitious yet valid mission assurance requirement to be verified at 10% risk to be considered for the Orion LAS then would be at 0.97. The remaining validation question is whether or not verification of a 0.97 mission assurance requirement is achievable and cost feasible.

Conditional Orion LAS Verification Plan Using Historical Ground Test Data

As discussed in section 2 of this report, a verification plan can be developed taking advantage of historical surrogate data. The historical surrogate data that is appropriate for verification of a mission assurance requirement for the Orion LAS is the aggregate ground test data in Table 1. Table 5 presents the plans developed using classical statistical procedures and using conditional approaches to verify a 0.97 mission assurance requirement for the Orion LAS with a 10% risk.

Table 5: Data Required for Successful Verification of 0.97 Mission Assurance with 10% Risk Using Classical Methods and Conditional Approaches taking Advantage of Surrogate Systems' Ground Test Data

Classical		Conditional		n_f
n_s	$P(V_{success})$	n_s	$P(V_{success})$	
76	9.88%	18	57.80%	0
128	14.90%	59	67.00%	1
174	15.49%	100	74.80%	2
218	15.63%	140	81.5%	3

The verification plans in the two leftmost columns of Table 5 are again developed using the binomial test, and cannot take advantage of the surrogate test data. The reduction of required test successes needed in column one of Table 5 vice in Table 2 is solely due to the relaxed mission assurance requirement to be verified.

As indicated in [11], one of the initiators of this investigation was a discussion about the numbers of tests needed to verify the proposed 0.9973 mission assurance requirement with no more than a 10% risk (stated in [11] as a 90% confidence). The total number of tests being considered as cost feasible based on heuristics was 15, very close to the 18 required in Table 5 if no failures occur when conditional approaches are used in verification planning for the mission assurance requirement of 0.97. The probability of verification success given that the 0.97 mission assurance requirement is satisfied in the as-built Orion LAS does not reach the preferred level above 90%. However, there is a better than two-fold improvement over the original conditional verification plan and five-fold improvement over the classical verification plan.

5. CONCLUSIONS

The investigations in this report yield three important conclusions for improving validation and verification of aerospace systems' performance requirements, and suggest a number of investigatory extensions. The improvements should be extensible to validation and verification of any performance requirement for any aerospace system.

Verification Planning Using Conditional Approaches Reduces Numbers of Tests Required

Conditional approaches are used as the basis for applying decision theory. Verification for aerospace systems always supports the decision making process for acceptance of the system. Therefore, verification planning using conditional approaches is suitable and appropriate for aerospace systems.

As demonstrated in section 2 of this report, conditional approaches can be used with models of maximum objectivity to eliminate most, if not all, questionable assumptions. This is not possible when developing a verification plan using classical statistical methods. As observed with the Orion LAS mission assurance requirement example, the use of maximum objectivity models in verification planning did not introduce any unnecessary inherent conservatism, producing verification plans using fewer numbers of tests, with higher probabilities of verification success than those developed using classical statistical methods. This results in more cost feasible verification plans due to the reduced numbers of tests that need to be performed.

Requirements Validation is Improved by Using Historical Data

When considering verification planning for a stringent aerospace system requirement, historical surrogate systems data may be used with conditional methods to investigate actual achieved requirement performance levels in both verification and operations as a function of residual risk. Surrogate systems that were fielded, were verified to satisfy their performance requirements at stated verification risk levels, and were cost feasible in design, manufacturing, and verification. Conditional approaches allow this validation process to be performed without any knowledge of the actual requirements levied on the surrogate systems, and without using any indefensible assumptions.

As was observed in this report for the proposed mission assurance requirement for the Orion LAS, 41 surrogate systems actually performed close to the Orion LAS requirement in operations considering the aggregated data. However, none of these surrogate systems were verified to the level of the proposed Orion LAS mission assurance requirement with the stated residual verification risk. This suggests that the industry that produced the surrogates, and that will produce the Orion LAS, may indeed design and built such systems to perform at levels that cannot be verified at those levels at a feasible cost. Relaxing the stringency of the Orion LAS mission assurance requirement to the 0.97 level for verification purposes should be valid based on results from processing the aggregate surrogate systems data.

Further, consider that the Ares I rocket will be designed, manufactured, and verified by the same industry that does so for the Orion LAS. A 0.97 mission assurance for the Ares I means that there is a 3% risk of failure. The Orion LAS with a mission assurance requirement of 0.97 has a 3% risk of failing, given that the Ares I fails. With these values for mission assurance requirements for both vehicles, the actual risk of loss of crew due to failure of the Orion LAS when the Ares I fails is the product of these two risks, a 0.09% risk.

When ground test and flight data from surrogate systems is available, conditional approaches may be used to validate both the level of required performance and required maximum acceptable verification risk for the system of interest.

Greater Verification Plan Achievability and Feasibility Results from Combining Conditional Approaches with Historical Data

Beyond the improvements offered to verification planning when using conditional approaches, the cost feasibility can be further improved if historical surrogate system data are available. As seen for verification of the Orion LAS mission assurance requirement in Table 5, once relaxed to a validated level, a verification plan taking advantage of the historical surrogate systems' ground test data can be developed that can be cost feasible, with a greatly improved probability of verification success.

Extensions to the Investigations

Conditional approaches may be used to predict future system performance based on observed data, i.e., test results [7]. Considering that anticipated use of the Orion LAS may be limited to fewer than 100 flights, it is possible to predict the probabilities for number of failures of the Orion LAS among a fixed number of flights, given numbers of surrogate ground test and flight successes and failures and actual Orion LAS verification results. Factoring these results into verification planning, coupled with requirements validation based on processing surrogate data using conditional approaches, may yield further improvements in cost feasibility for verification of stringent requirements for aerospace systems.

REFERENCES

[1] Anthony J. Hayter, Probability and Statistics for Engineers and Scientists, Third Edition, Belmont, CA, Duxbury, 2007.

[2] C. J. Clopper & E. S. Pearson, "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial," Biometrika, 26, 404-413, 1934.

[3] James O. Berger, Statistical Decision Theory and Bayesian Analysis, Second Edition. Springer-Verlag, New York, 1980.

[4] Mark A. Powell, "Optimal and Adaptable Reliability Test Planning Using Conditional Methods," Proceedings from the 14th Annual International Symposium, International Council on Systems Engineering, Toulouse, FR, June 20–24, 2004.

[5] Harold Jeffreys, Theory of Probability. Oxford University Press, Oxford, 1939.

[6] Edward Jaynes, "Prior Probabilities". IEEE Transactions Systems, Science and Cybernetics, 4, 227-291, 1968.

[7] Jose Bernardo and Adrian Smith, Bayesian Theory. John Wiley & Sons, LTD, New York, 1994.

[8] Howard Raifa and Robert Schlaifer, Applied Statistical Decision Theory. John Wiley & Sons, Inc. New York, 1960.

[9] Brian Muirhead, "Constellation Architecture and System Margins Strategy," Proceedings from the International Astronautical Congress, Glasgow, Scotland, September 29 – October 3, 2008.

[10] Scott "Doc" Horowitz, "Current Constellation Planning," presented to the House Science and Technology Committee Staff, March 13, 2007.

[11] Mark A. Powell, Safety and Mission Assurance Special Assessments, CxP Orion LAS Verification Planning Evaluation Report, Repository Number: JS-2008-002, November 12, 2007.

BIOGRAPHY

Mark Powell has practiced Systems Engineering for over 35 years in a wide range of technical environments including DoD, NASA, DOE, and commercial. More than 25 of those years have been in the aerospace arena. His roles in these environments have included project manager, engineering manager, chief systems engineer, and research scientist. He is currently an adjunct member of the Stevens Institute of Technology Systems Engineering Faculty, and of the University of Houston, Clear Lake Systems Engineering Faculty. Mr. Powell maintains an active engineering and management consulting practice (currently in affiliation with SAIC) in North America, Europe, and Asia. Beyond consulting, he is sought frequently as a symposium and conference speaker and for training, workshops, and tutorials on various topics in Systems Engineering, Project Management, and Risk Management. Mr. Powell is an active member of AIAA, Sigma Xi, the International Society for Bayesian Analysis, and the International Council on Systems Engineering, where he serves as Assistant Director for Systems Engineering Processes.