IDA

INSTITUTE FOR DEFENSE ANALYSES

Improving Reliability Estimates with Bayesian Statistics

Laura J. Freeman, *Project Leader*
Kassandra Fronczyk

June 2015

IDA

The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-5452

Improving Reliability Estimates with Bayesian Statistics

Laura J. Freeman, *Project Leader*
Kassandra Fronczyk

IMPROVING RELIABILITY ESTIMATES WITH BAYESIAN STATISTICS

Dr. Kassandra Fronczyk, IDA

Dr. Laura Freeman, IDA

Abstract

There are many challenges with assessing the reliability of a complex system. One of the more difficult aspects of system reliability assessment is integrating multiple sources of information, including component, subsystem, and full system data, as well as possible previous test data or subject matter expert (SME) opinion. The Bayesian paradigm is tailor made for these types of situations, allowing for allowing for the combination of multiple sources of data and variability to obtain more robust reliability estimates and uncertainty quantification. The Bayesian approach to combining reliabilities is discussed through a recent multi-mission ship example. The Bayesian approach to combining information from various subsystems/components and other sources to estimate full system reliability has many advantages. Most notably, multiple sources of prior information can be incorporated, complex systems (and their structures) can be analyzed without seriously increasing the computations, and uncertainty intervals are straight-forward to calculate and interpret.

Introduction

Assessing the operational reliability of complex systems can often require the incorporation of multiple sources of data in order build a credible statistical analysis of system reliability. Ideally, adequate data on the system reliability would be collected during operational testing, using representative users under a range of operationally realistic conditions. Unfortunately, it is often not possible or cost effective to collect all of the data on system reliability in operational testing. Additionally, it is common to have operational tests that result in zero failures. In these cases, using a range of additional sources of information may provide a better assessment of the system reliability and provide a methodology for estimating reliability where there is no solution in a traditional frequentist paradigm. Examples of reliability information that may be desirable to include in an operational reliability assessment include data from systems tested in stages (e.g., captive-carry and live fire), or that have multiple subsystems, with limited full scale testing. Resulting data from testing these systems may include pass-fail, lifetime, and/or degradation information, all of which needs to be coherently combined into the reliability analysis. In some cases, we may have information from prior testing that we want to incorporate into the analysis.

Within the reliability literature, there are many instances of Bayesian approaches for complex system assessments. Anderson-Cook et al. (2007) assess the reliability of the stockpile of nuclear weapons, where it is not possible to conduct many full system tests. Hamada et al. (2004), Graves et al. (2010), Wilson et al. (2011), and Guo et al. (2013) all

consider methods that allow for the combination of component-level data or data from other variants of the system that can be incorporated into the system-level analysis. Johnson et al. (2003) and Reese et al. (2011) show how one can use Bayesian hierarchical models to integrate component, subsystem and system data, along with prior expert opinion, to assess the reliability of a complex system. For examples of advances, opportunities, and issues in data combination, see Anderson-Cook (2009) and Wilson et al. (2006).

Bayesian methods are valuable for their natural integration of prior information and their practical convenience for modeling and estimation. The Bayesian approach allows for simultaneous assessment of the reliability of each subsystem or stage and the overall system reliability. To illustrate the Bayesian method, we will look at an example based on the reliability of a recent multi-mission ship. This example illustrates a proof of concept for combining information in operational reliability assessments and can be extended to analyze a myriad of operational test data.

Bayesian Reliability

Models used to estimate reliability are characterized by a number of parameters. For example, the exponential or Weibull distributions are commonly used to estimate the mean time between failures (MTBF) for continuous data, while the binomial distribution is commonly used to estimate probability of failure for pass/fail systems. The classical statistical approach considers these parameters (e.g., mean time to failure) as fixed, unknown constants to be estimated using data. A confidence interval for an unknown parameter is a frequency statement about the likelihood that numbers calculated from a sample capture the true parameter. Strictly speaking, we cannot make probability statements about the true parameter.

The Bayesian approach, however, treats the model parameters as random quantities. There is no assumption that the model parameters (e.g., MTBF) are actually random; but instead, we use the tools of probability to express our uncertainty about the values of the parameters. Before looking at the test data, we construct a prior distribution, or starting assessment, for the model parameters. We use the data to revise our starting assessment, and derive the updated assessment (i.e. the posterior distribution) for the model parameters. Parameter estimates, along with confidence intervals (known as credible intervals), are calculated directly from the posterior distribution. Credible intervals are probability statements about the unknown parameters.

The use of Bayesian methods is becoming increasingly popular because it makes practical sense to leverage all the information available when making decisions under uncertainty, and when interpretation of the results is intuitive. Moreover, these methods often have good properties for small sample sizes, and as sample sizes get larger, the inferences match those of traditional frequentist methods. Finally, these methods provide flexible analyses that can account for complex system configurations with relatively simple numerical methods when compared to traditional methods.

Bayesian reliability is explained by the simple formula:

$$f_{posterior}(R|data) \propto L(data| R) f_{prior}(R).$$

In words, the posterior distribution, $f_{posterior}$, for reliability R given the data equals the (normalized) product of the prior information (stored in f_{prior}) and the data, formulated as a likelihood function (L). This is Bayes' theorem. A traditional frequentist analysis considers only the likelihood function of the data, so the fundamental mathematical shift in a Bayesian analysis is the incorporation of the prior information.

For illustrative purposes throughout the remainder of the method explanation, we begin with a small example: a kill chain with two components. For each of the two components, 10 pass/fail tests are administered and results are recorded. Component 1 fails twice and Component 2 fails zero times. We can calculate the reliability of each component and find confidence intervals of the reliability of each system, R_1 and R_2. We also want an assessment of the system reliability, assuming the components work in series (this will be discussed further in a subsequent section). Under the frequentist approach the reliability is calculated as the product of the reliability of each component, or

$$R_{system} = R_1 * R_2 = \left(1 - \frac{2}{10}\right) * \left(1 - \frac{0}{10}\right) = 0.8 * 1 = 0.8.$$

The estimated system reliability is 0.8. The confidence interval associated with the reliability would reflect our uncertainty about this estimate. How do we calculate a confidence interval on R_{system}? There is no straightforward answer to this, especially given a subsystem with zero failures. Additionally, if Component 1 also had zero failures, do we believe the system has a reliability of 1? That is debatable. There may be previous test data that support the perfect reliability claim, so a reliability estimate of 1 might be close to the truth. By adopting a Bayesian framework, we can quantify our belief in such a claim.

We'll focus on a single component, and then describe the process to combine it with the second component for the system reliability. The mechanics of applying Bayesian probability to reliability distributions can be summarized in three steps:
- Construct prior from prior information (e.g., previous testing)
- Construct likelihood from test data
- Estimate posterior distribution using Bayes' theorem

In the first step, the prior distribution of the reliability, $f_{prior}(R)$, is constructed from previous data or expert knowledge. The prior reliabilities are assigned in the form of a distribution to be determined before the data are obtained. Careful thought should always be put into the prior distribution, as naively specified priors for Bayesian system reliability can lead to misleading results.

Turning to our two-component example, say Component 1 was previously tested and failed 3 out of 40 tests. Depending on how operationally realistic the previous testing was, we may choose to include none or all of the prior information into our prior assessment of reliability, $f_{prior}(R_1)$. One approach to including this information is through a Beta distribution

$$f_{prior}(R_1) \propto R_1{}^{n_p p}(1 - R_1)^{n_p(1-p)},$$

with p as the reliability estimate and $n_p \geq 0$ as the confidence factor of that prior estimate. When n_p is set to 0, we do not believe that the prior data are relevant to the current test data and the prior distribution gives equal probability to all values between 0 and 1 (see the middle panel of Figure 1). As n_p increases, our confidence in the prior reliability estimate increases and the distribution peaks around said estimate (see the left and right panels of Figure 1).

Second, tests are performed and the resulting test data are used in the likelihood function, L(data|R). This likelihood function is the starting point for classical reliability analysis. As in the frequentist framework, the binary test data of Component 1 follow a Binomial distribution with probability of a pass of R_1. That is,

$$L(data|R_1) \propto R_1^{s_1}(1 - R_1)^{f_1},$$

where s_1 is the number of successes and f_1 is the number of failures from Component 1.

Finally, Bayes' theorem is used to find the posterior reliability distribution, $f_{posterior}$(R|data). The posterior distribution is proportional to the product of the prior distribution and the likelihood function for all subsystems in the unit. For our small example, choosing the Beta distribution as a prior is ideal for a few reasons: it ensures that R is between (0, 1) and it is the "conjugate" prior for the Binomial distribution. Conjugate priors are such that the form of the prior distribution, when combined with the likelihood, is the same as the posterior distribution. That is, combining the likelihood with the prior and rearranging terms, we have

$$f_{posterior}(R_1|data) \propto R_1^{s_1}(1 - R_1)^{f_1}R_1^{n_p p}(1 - R_1)^{n_p(1-p)}$$
$$\propto R_1^{s_1+n_p p}(1 - R_1)^{f_1+n_p(1-p)}$$

which is a Beta distribution, with parameters $s_1 + n_p p + 1$ and $f_1 + n_p(1 - p) + 1$. Turning to our Component 1 reliability, Figure 1 shows three different priors (dashed lines), the Binomial likelihood of 8 successes out of 10 tests (solid lines), and the posterior distributions (dotted lines) based on the respective priors. The left panel has a prior assessment that is based on the previous testing of Component 1, which failed only 3 of 40 tests and included a weight of 10; that is, the prior on R_1 is a Beta distribution with parameters a=10*37/40 and b=10*3/40. While we have 40 tests, we do not believe they are fully relevant to the full system test. However, the prior distribution has most of the probable reliability estimates between 0.6 and 1. After the 10 current tests, the posterior combines both distributions into a Beta distribution with parameters a+8 and b+2; the mean estimate is about 0.86 with an 80 percent credible interval of (0.76,0.95). The middle panel shows the non-informative prior analysis (i.e., $n_p = 0$ or Beta(1,1)). The posterior distribution follows a Beta(1+8,1+2) distribution giving a mean estimate of 0.75, but our uncertainty is wider because we did not include any previous information (0.58,0.89). Finally, the right panel gives a prior reliability assessment of just under 0.4, which contradicts the data assessment, and a prior weight of 3 or Beta(3*15/40,3*25/40). The posterior distribution is Beta(3*15/40+8, 3*25/40 +2), which is shifted closer to the data estimate, providing a posterior mean of 0.70 and an 80 percent interval of (0.54,0.85). Which of these three prior distributions are sensible? They all are. However, a Bayesian analysis needs to justify how they set their prior. This will change depending

on the specific situation and the analyst. One possible method is to choose the one that reflects the information at hand prior to the analysis, and report how much the results change under more or less informative prior distributions.
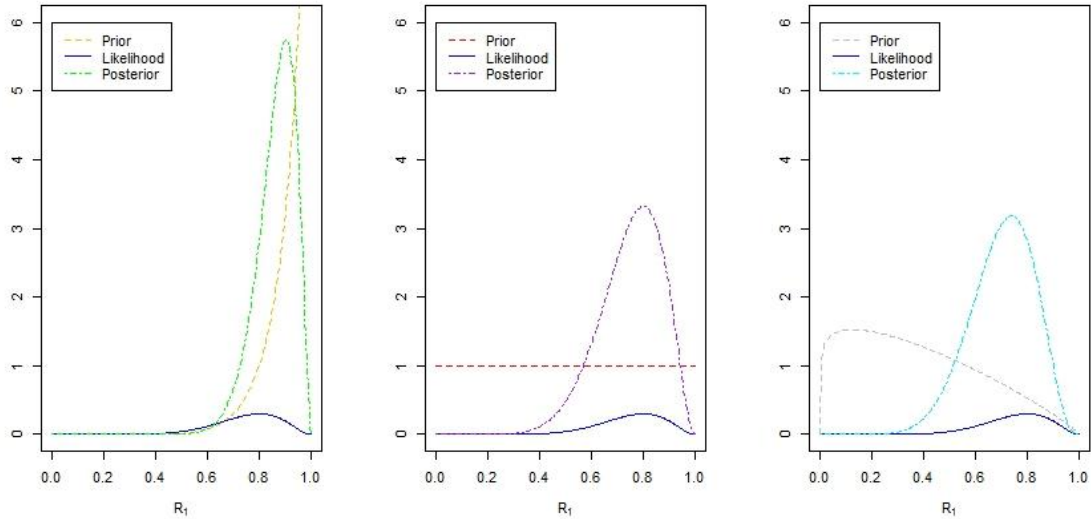

Figure 1. Prior (dashed), Likelihood (solid), and Posterior (dotted) Distributions of the Reliability of Component 1

Note that a thoughtfully specified prior distribution keeps the reliability prediction based on operational data from being either too optimistic or too pessimistic. Consider a situation in which you can run only one operational test. If it is a success, the frequentist estimate of reliability is 1. Similarly, if the single unit test failed, the data estimated reliability is zero. By taking the Bayesian approach, the reliability data estimate is mitigated based on the prior information. In our example, even a non-informative Beta(1,1) prior for the zero failures on Component 2 results in a posterior distribution of Beta(11,1); that is, the posterior mean for R_2 is 0.92 with an 80 percent credible interval of (0.811, 0.99).

Continuous Failure Data

For continuous failure time data, the reliability is generally defined as a function of the failure rate parameter, λ; i.e., MTBF is defined by 1/λ. The most common prior used for λ is the Gamma distribution proportional to

$$\lambda^{a-1}exp(-b\lambda),$$

where a,b>0 are the shape and scale parameters such that the expected value (mean) for λ is a/b. There are many ways to choose a and b. One approach is to make an educated guess at the MTBF, and proceed with a weak prior that will change rapidly based on test data. If the prior parameter a is set to 1, the Gamma has a standard deviation equal to its mean, which makes it spread out.   To ensure that the 50th percentile is set at λ_{50}=1/MTBF$_{guess}$, we have to choose b=log(2)×MTBF$_{guess}$.

For failure time, the Exponential distribution is commonly used for the likelihood, proportional to λexp(-λt), where t is the failure time and λ>0 is the failure rate. Here, the

reliability depends on time, t, given by R(t) = exp(-λt). The specific likelihood function is chosen based on the data to be analyzed, there are many options available. As with the Binomial case, the posterior distribution for λ is proportional to $\lambda^{\alpha+n-1}$exp(-(β+T)λ), where T is the sum of the failure times. This is of the same form as the Gamma distribution, with updated parameters.

Indeed, there are conjugate priors for many distributions. Some of the more common conjugate likelihood/prior combinations are displayed in Table 1. Note that these priors are chosen mainly for mathematical convenience. If a conjugate prior is not available (as in the case of logistic regression), choose one that addresses your current beliefs. While your posterior distribution will not have a nice analytic expression, there are many techniques to make inference and obtain credible intervals.

Table 1. Conjugate Prior Examples

LIKELIHOOD	PARAMETER	PRIOR	POSTERIOR
Binomial(s+f, R)	0 ≤ R ≤ 1	Beta(a,b) a > 0, b > 0	Beta(a',b') a'= a + s b'= b + f
Poisson(λ)	λ > 0	Gamma(a,b) a > 0, b > 0	Gamma(a',b') a' = a + Σ t b' = b + n
Exponential(λ)	λ > 0	Gamma(a,b) a > 0, b > 0	Gamma(a',b') a' = a + Σ t b' = b + n
** For more examples, see Wikipedia conjugate priors page or *Bayesian Reliability* pg. 48			

Given the posterior distribution for the reliability parameters, we can look at the mean and quartiles of the distributions to create credible intervals. These distributions can also be used to find the posterior distribution of any function of the parameters, including reliability of time dependent data, or system reliabilities.

Combining Reliabilities

The previous section focused on a single component or subsystem of a system. We want to assess the reliabilities of a system composed of multiple subsystems or components, as with the two-component example and with the multi-mission ship. We assume the components/subsystems of both systems of interest are in series, meaning that the system functions only if all components/subsystems are functioning. More general structures can be analyzed by following the same logic presented for series systems (See Meeker and Escobar, Chapter 15). Also note that we are assuming the subsystems and components are independent. This may not be a reasonable assumption in some cases.

Say we have N_c subsystems within a series system. At time t, we know whether each subsystem is functioning and, when applicable, whether the system is functioning. The reliability of the system at time t is defined as the probability of the system functioning at time t. By definition of a system functioning in series, this probability is equal to the probability that all N_c components are functioning at time t. Under an

assumption of independence of the subsystems, the system reliability is a product of the probability each component is functioning at time t, i.e., the product of the N_c component reliabilities. Mathematically,

$$R_{system}(t) = \Pi\, R_j(t)$$

where $R_j(t)$ denotes the reliability of component j.

This reasoning applies regardless of whether we take a frequentist or a Bayesian approach. One of the main differences is that there is no simple procedure to combine standard errors and create confidence intervals for the system reliability under the frequentist paradigm. Bootstrap intervals or intervals based on Normal-approximations are available, but they break down when zero failures are observed. Because we have a distribution for each of the component reliabilities with the Bayesian framework, we can easily find appropriate credible intervals. Looking at our two-component system with independent Beta(1,1) priors, the posterior distribution for R_1 is Beta(9,3) and Beta(11,1) for R_2. By generating a random value from the posterior of R_1 and multiplying it by a random value from the posterior distribution for R_2, we have a value from the system reliability. Repeating this step a large number of times, say 100,000, we obtain 100,000 samples from the distribution of the system reliability. We take the mean of the samples (0.69) and the 10th and 90th quantiles (0.52, 0.84) to characterize the distribution of the system reliability. Additionally, we may create a histogram of the samples (see Figure 2) for a graphical representation. From these distributions, we can make probabilistic statements about the parameters by counting the number of occurrences within the posterior samples. As an example, we can say that the probability the system reliability is over 0.8 is about 0.2, as about 20 percent of the posterior samples of the system reliability are above 0.8. This is not possible under the frequentist paradigm.
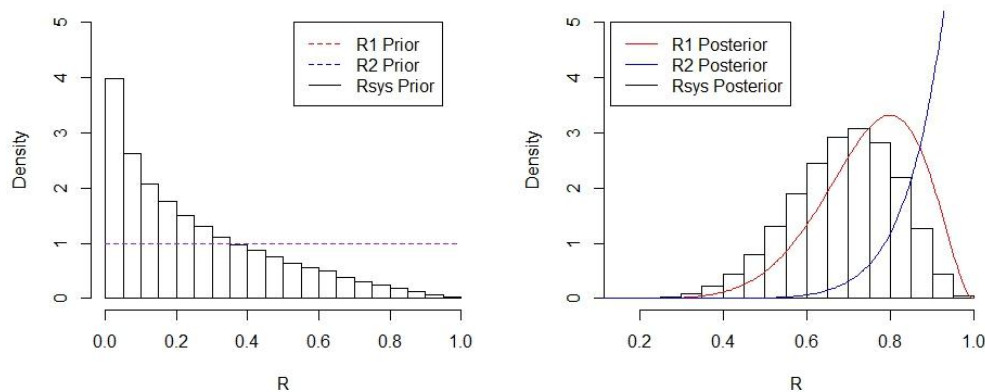

Figure 2. Prior (left panel) and Posterior (right panel) Distributions of Component 1 Reliability (red), Component 2 Reliability (blue), and System Reliability (black)

As a side note, the assumption of diffuse priors is not always appropriate, especially for components within a series system. Diffuse component reliability priors do not induce diffuse priors on the system reliability (see the left panel of Figure 2). In fact, the strength of the information on system reliability increases as the number of components or subsystems increases. Moreover, assuming independent priors for

subsystems or components of a system is not always applicable. Be conscientious of the assumptions you are making because every analysis is different.

Case Study: Ship Mission Reliability

The reliability requirements for ships are often broken down into thresholds for the critical or mission-essential subsystems. A notional example of a ship system under test is presented here. Typical mission-essential subsystems could be: Propulsion and Electrical Distribution; Heating, Ventilation, and Air Conditioning (HVAC); Communications; and Survivability; each comprising multiple subsystems. The subsystems are either on-demand (resulting in count data of number of missions requiring the system with and without an operational mission failure), continuous underway (resulting in count data of number of failures in the 2,000 hours of testing while underway), or continuous full time (resulting in count data of number of failures in the 4,500 hours of testing including underway and docked time). Each system includes different types of subsystems, and some of these subsystems have no failures in the duration of the test. These facts introduce multiple challenges in addressing an overall ship or functional area reliability requirement. We focus on the Propulsion and Electrical Distribution functional area, which includes four subsystems: Main Propulsion system (MP) (full time), Ship Service Diesel Generators (SSDG) (underway), Machinery Control System (MCS) (full time), and Auxiliary Propulsion System (APS) (on-demand). The target reliability for Propulsion and Electrical Distribution is 0.80 in 720 hours. In this case, the Bayesian analysis allows the estimation of the functional area reliability with credible intervals even though multiple subsystems of this functional area have zero failures. This is not possible under a traditional reliability analysis.

Reliability data were collected over a 200-day period from the start of developmental testing through operational testing. Quantitative measures were considered for all four functional areas, but we will focus only on Propulsion and Electrical Distribution. Each of the operational mission failures (OMF) was recorded under the relevant subsystem. For Propulsion and Electrical Distribution, these subsystems were MP (full time), SSDG (underway), MCS (full time), and APS (on-demand). Table 2 summarizes data that are based on the results of the ship reliability testing. It should be noted that the data presented in Table 2 are notional data to illustrate the concept of this analysis.

Table 2. Ship Reliability Test Data

CRITICAL SUBSYSTEM	TOTAL SYSTEM OPERATING TIME	OMFs
Main Propulsion (MP)	4500 hours	1
Ship Service Diesel Generators (SSDG)	2000 hours	3
Machinery Control System (MCS)	4500 hours	0
Auxiliary Propulsion System (APS)	11 missions	2

We assume an Exponential distribution for the time of OMFs for MP, SSDG, and MCS, and a Binomial distribution for the number of OMFs out of 11 missions for the on-demand APS. We will assume independent Gamma(a,b) prior distributions for the continuous subsystems, with a = 1 and b = log(2)*$MTBF_{50}$, where $MTBF_{50}$ is set for each subarea through the requirement, as this is the only information we have about the subsystem. An estimate for the MTBF is found by setting the requirement (0.80 at 720 hours) equal to the function R(t) and solving for λ. That value is split between the two-full time systems and the underway system. Figure 3 gives the prior distribution for each failure rate. A uniform prior ($n_p = 0$, or $Beta(1,1)$) is chosen for the reliability of APS. Here, we do not have much information to include about the reliability of the APS, and a diffuse prior is deemed appropriate. We are interested in the reliability of each subsystem and the system as a whole at 720 hours.

By choosing independent, conjugate priors, we have posterior distributions in closed form for each of the individual subsystems. However, the full-system reliability does not have an analytical expression. To obtain samples from the posterior distribution, a Markov chain Monte Carlo (MCMC) algorithm is employed. MCMC techniques generate a sequence of random samples from a probability distribution for which direct sampling is difficult. This sequence can be used to approximate the distribution (i.e., to generate a histogram) or to compute an integral (such as an expected value). The quality of the sample improves as the number of draws gets larger. Essentially, we draw a sample from each closed form posterior of the individual subsystem parameters. Those samples are used to calculate the full system reliability (see expression below) to acquire a sample from the full system reliability posterior distribution. This process is repeated a large number of times.
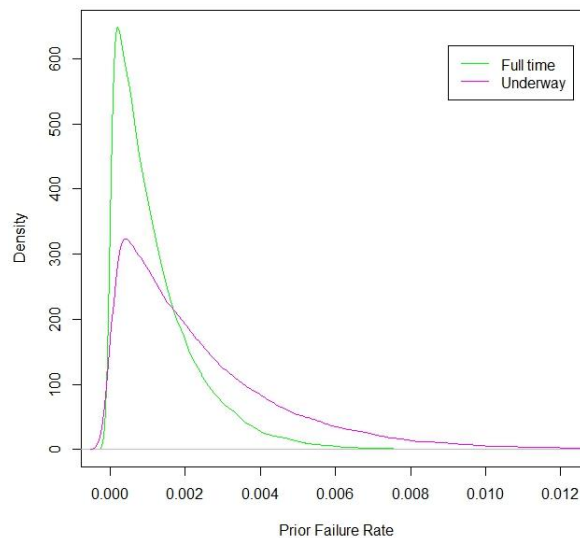

**Figure 3. Prior Distribution on the Failure Rate (λ) for the
Full-Time Systems (green) and the Underway System (purple)**

Table 3 provides results from the MCMC analysis. The posterior mean estimate and the 80 percent credible interval are shown for the MTBFs ($1/\lambda$) of MP, SSDG, and MCS, as well as the reliability of APS. Note that the average MTBFs are similar to those found from the classical analysis, save for MCS, where no failures occurred. Also, the intervals are fairly wide because of the small number of observed OMFs and the dispersed priors. The credible intervals are not necessarily symmetric around the mean because they are based on the quantiles of the estimated posterior distribution.

Table 3. Mean and 80 Percent Interval Estimates for the Subarea and Full-System Reliability Based on Frequentist and Bayesian Analyses

	FREQUENTIST MTBF	FREQUENTIST R(720 HOURS)	BAYESIAN MTBF	BAYESIAN R(720 HOURS)
MP	4500 hrs (1156 hrs, 42710 hrs)	0.85 (0.54, 0.98)	3630 hrs (1179 hrs, 6753 hrs)	0.73 (0.54, 0.90)
SSDG	667 hrs (299 hrs, 1814 hrs)	0.33 (0.09, 0.67)	697 hrs (332 hrs, 1172 hrs)	0.31 (0.11, 0.54)
MCS	** 2796 hrs	** 0.77	10320 hrs (1721 hrs, 18210 hrs)	0.83 (0.66, 0.96)
APS		0.82 (0.46, 0.87)		0.77 (0.62, 0.91)
OVERALL Propulsion and Electrical Distribution		0.18		0.15 (0.05, 0.27)

** A conservative 80 percent lower confidence bound; frequentist MTBF does not exist

We can assess the reliability of the continuous subsystems over time through $R(t) = \exp(-\lambda t)$. The posterior mean (solid line) and 80 percent intervals (dashed lines) of $R(t)$ are given in Figure 4 for MP (blue), SSDG (red), and MCS (green), as a function of time. The Machinery Control System has the most uncertainty because there were no failures during the full testing time. MP has slightly narrower intervals and decline because of the one observed failure throughout the full test time. SSDG has a rapid decline because it is an underway system (tested for less than half of the full time systems) coupled with multiple observed failures.
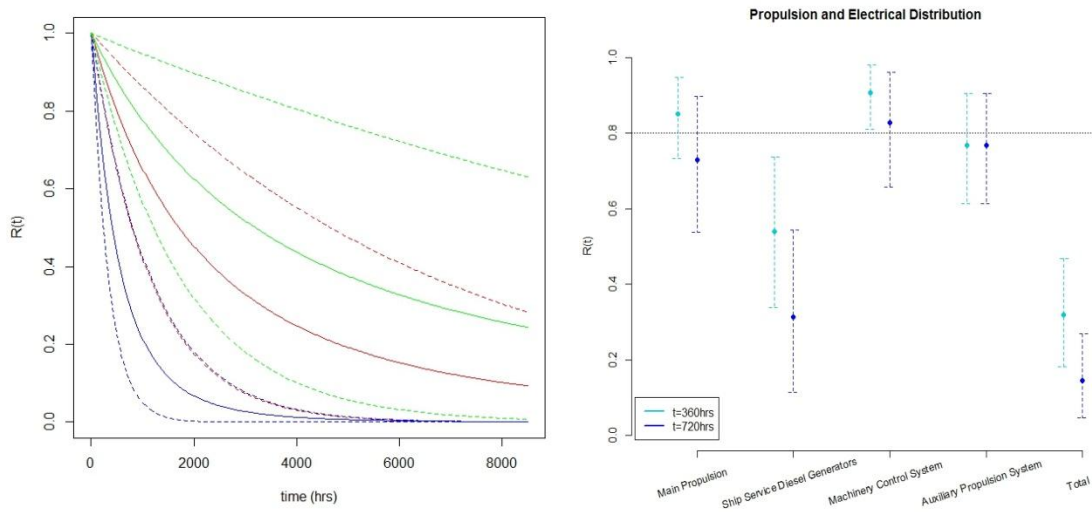
Figure 4. Left Panel: Posterior Mean (solid line) and 80 Percent Intervals (dashed lines) of R(t) for MP (red), SSDG (blue), and MCS (green). Right Panel: Mean and 80 Percent Intervals for Each Subsystem and the total System Reliability for 15 days (light blue) and 30 days (dark blue)

To find the system reliability at 720 hours, we will evaluate the following equation with the posterior distributions of the four parameters (three rates of failures and the one subsystem reliability)

$$Rsys \ = \ exp(-720\lambda_{MP})exp(-720\lambda_{SSDG})exp(-720\lambda_{MCS})\pi_{APS}.$$

To change the time frame from 30 days to say 15 days, for example, we just replace the 720 with 360. The right panel of Figure 4 graphically depicts the mean and interval estimates of both subsystem and combined system reliability at 15 and 30 days. As APS is an on-demand system, there is no dependence on time and the reliability estimates for that subsystem are equivalent at 15 and 30 days.

The series assumption is made in many cases, regardless of the actual functional structure of the whole system. This results in a relatively conservative estimate of the system reliability if the subsystems or components do not actually run in series. We may choose to weight the four subsystems based on user profiles, but results will be highly sensitive to the re-weighting. As seen in the last row of Table 3, the reliability of the system as a whole is estimated to be at 0.15, with an 80 percent credible interval of (0.05,0.27). The assessment of reliability for the Propulsion and Electrical Distribution functional area is significantly below the requirement of 0.80, whether we take a frequentist or Bayesian approach. However, we gain the ability to provide interval estimates of the system reliability in the Bayesian analysis. Worth noting is that the Bayesian analysis was done multiple times with different prior settings, and the resulting inference for the subsystem and combined reliability were robust to these settings.

Conclusions

This article provides an introduction to the Bayesian approach of reliability analysis. We describe how combining component or subsystem reliabilities into a full-system estimate is tractable and provides appropriate uncertainty quantification under this paradigm. Illustration of the method is given using an example based on a recent ship reliability test (see Dickinson et al (2015) for another example of combining information with a Bayesian reliability analysis). The methods used here can easily be applied to many other programs, such as programs with kill chains, complex system structures, and previous test data or SME opinion (information that is not incorporated into a frequentist analysis). Note that if the system does not work in series, the logic to find the system reliability estimate is not as intuitive, but still possible. There are also mechanisms to up- or down-weight the prior information based on how relevant it is to the current testing. The Bayesian approach also avoids unrealistic reliability estimates in cases where there are zero failures or zero successes, another common complication. By exploiting all available information and tools, we can obtain rich inference for very complex problems.

References

Anderson-Cook, C. M. (2009). Opportunities and issues in multiple data type meta-analyses, *Quality Engineering*, 21:3, 241-253.

Anderson-Cook, C. M., Graves, T., Hamada, M., Hengartner, N., Johnson, V., Reese, C.S., Wilson, A.G. (2007). "Bayesian stockpile reliability methodology for complex systems." *Journal of the Military Operations Research Society* 12 25-37.

Dickinson, R., Freeman, L.J., Simpson, B., Wilson, A.G. (2015) "Statistical Methods for Combining Information: Stryker Family of Vehicles Reliability Case Study." *Journal of Quality Technology*, to appear.

Graves, T., Anderson-Cook, C. M., and Hamada, M. (2010). Reliability models for almost-series and almost-parallel systems. *Technometrics.* 52(2): 160-171.

Guo J., Wilson A.G. (2013) "Bayesian methods for estimating the reliability of complex systems using heterogeneous multilevel information." *Technometrics* 55, 461-472.

Hamada, M., Martz, H. F., Reese C. S., Graves, T., Johnson, V., Wilson, A. G. (2004). A fully Bayesian approach for combining multilevel failure information in fault tree quantification and optimal follow-on resource allocation. *Reliability Engineering and System Safety*, 86(3): 297-305.

Hamada, M.S., Wilson, A.G., Reese, C.S., Martz, H.F. (2008). *Bayesian Reliability*. New York: Springer.

Johnson, V. E., Graves, T. L., Hamada, M. S., Reese, C. S. (2003). A hierarchical model for estimating the reliability of complex systems. *Bayesian Statistics*, 7, Oxford University Press.

Meeker, W. Q., Escobar, L. (1998). *Statistical Methods for Reliability Data*. New York: John Wiley & Sons.

Rausand, M., Hoyland, A. (2004). *System Reliability Theory: Models, Statistical Methods, and Applications, 2nd Editon.* New York: John Wiley & Sons.

Reese, C. S., Wilson, A.G., Guo, J., Hamada, M.S., Johnson V.E. (2011). A Bayesian Model for Integrating Multiple Sources of Lifetime Information in System-Reliability Assessments. *Journal of Quality Technology,* 43(2): 127-141.

Wilson, A. G., Graves, T. L., Hamada, M. S., Reese, S. C. (2006). Advances in Data Combination, Analysis and Collection for System Reliability Assessment, *Institute of Mathematical Statistics*, 21:4, 514-531.

Wilson, A., Anderson-Cook, C. M., Huzurbazar (2011). A case study for quantifying system reliability and uncertainty. *Reliability Engineering and System Safety*, 96(9): 1076-1084.

Wilson, A.G., McNamara, L., Wilson, G. (2007). Information integration for complex systems. *Reliability Engineering and Systems Safety.* 91(1):121-130.