LA-UR- 08-5483

Approved for public release; distribution is unlimited.

Title: Resource Allocation for Reliability of a Complex System with Aging Components

Author(s): C.M. Anderson-Cook
T.L. Graves
M.S. Hamada

Intended for: Quality and Reliability Engineering International

Los Alamos
NATIONAL LABORATORY
EST. 1943

Resource Allocation for Reliability of a Complex System with Aging Components

Christine Anderson-Cook
Todd L. Graves
Michael S. Hamada

Statistical Sciences
Los Alamos National Laboratory
Los Alamos, NM 87545

Abstract:

To assess the reliability of a complex system, many different types of data may be available. Full-system tests are the most direct measure of reliability, but may be prohibitively expensive or difficult to obtain. Other less direct measures, such as component or section level tests, may be cheaper to obtain and more readily available. Using a single Bayesian analysis, multiple sources of data can be combined to give component and system reliability estimates. Resource allocation looks to develop methods to predict which new data would most improve the precision of the estimate of system reliability, in order to maximally improve understanding. In this paper, we consider a relatively simple system with different types of data from the components and system. We present a methodology for assessing the relative improvement in system reliability estimation for additional data from the various types. Various metrics for comparing improvement and a response surface approach to modeling the relationship between improvement and the additional data are presented.

Key words: Bayesian analysis, design of experiments, sequential experimentation, mixture experiments, meta-analysis, reduction of uncertainty

1. Introduction

When estimating the reliability of a complex system, different potential sources of data may be available. The full-system tests are the most direct assessment of this reliability, but other sources may be cheaper, more plentiful and can also be beneficial when appropriately combined with understanding of the system structure. Methodology to model system reliability as a function of component, sub-system and system level data is discussed in Wilson et al. (2006) and Anderson-Cook et al. (2007, 2008). This Bayesian analysis approach allows multiple types of data to be combined with subject matter expertise through prior distributions in a single analysis to provide a synthesized estimate of system reliability which reflects all sources of data.

The form of the system is incorporated into the model to reflect how the component and subsystem data should be combined to accurately reflect the connections between the components. Common structures used to capture the structure include series and parallel systems. For more details on types of system structures, see Rausand and Hoyland (2004) and Saunders (2007). Some of the types of data that might be available to assess portions of the system include:

1. pass/fail data evaluated at a given age of the component
2. degradation data consisting of a continuous measure with known operational limits, outside of which the component is not expected to work successfully
3. lifetime data for components tested and observed until they fail

Typically the estimate of system reliability is modeled as a function of the age and potentially the usage of the system. The data collection and analysis can be an ongoing process where new data are collected to help update the reliability estimates as the population of systems age. Prediction of system reliability beyond the observed ages of the system is common, and the target range of extrapolation may be expanded with sequential data collection over time.

Resource allocation is a form of sequential experimentation, where a formal process is used to determine how to best spend future resources available for collecting new data. The problem considers how to best determine which types of data are most advantageous for maximally improving the precision of our estimation and prediction conditional on the data already available and leveraging understanding of system reliability as a function of its components' reliabilities. Figure 1 illustrates the basic problem considered in the remainder of this paper. Phase 1 involves collecting initial data and information, which when combined with knowledge of the system structure allows for construction of an appropriate statistical model for an analysis to be performed. From the analysis, estimates of system and component reliabilities are available across the range of observed ages as well as for system ages not yet observed. Phase 2 involves collecting more data, and re-running the analysis with the combined data (with both the initial data and the newly collected data) to provide updated estimates of model parameters and system reliability. We assume that the same form of the model is being used, but just with additional data. The decision-making process of resource allocation should occur between Phases 1 and 2 of Figure 2, and guides the choice of new data to be collected as part of sequential experimentation. As with any design of experiment selection, the choice of which new data to collect must be made before the new data are available to show what improvements to the estimation of the system reliability their values provide. Hence we wish to use our current understanding of model parameters to help inform us about what data might be expected as well as how it will influence the uncertainty about system reliability. First, a few key points about the scope of our discussions:

1. We consider new data of the same types as those already collected in the Phase 1. This approach does not consider adding new potential data sources, which might model alternate failure mechanism not currently in the model or alternative data types that would complement existing data for a given component or subsystem.
2. Because we believe that the current model appropriately summarizes the system reliability, we focus on reducing the uncertainty in our estimation, rather than looking to reduce potential bias from an incorrectly specified model.
3. We assume that the available budget for Phase 2 data collection and cost of each type of data are known and fixed. Typically the costs for different types of data can vary greatly.
4. There may be restrictions and constraints on the types and amounts of new data that can be collected. These logistical or practical restrictions may limit the available choices for allocations. Initially, we assume that the user specifies possible allocations to be considered, and then the best of these will be identified. Later in the paper we present some extensions that allow for estimation of a global best allocation within a bounded allocation design space.
5. We assume that while good estimation of reliability for all of the components is helpful for understanding, the primary focus of the problem is to improve the precision of our system reliability estimate.

6. We assume that management of the systems depends on the estimation or prediction of system reliability at particular ages, perhaps in the range of systems already observed or involving extrapolation to older ages.

We therefore seek to find the best allocation of these fixed resources to maximally reduce the uncertainty of our prediction of system reliability for a user specified range of system ages.

We first present the algorithm to assess the potential allocations which have been identified. Suppose that we have d allocations to compare based on a user-selected metric for quantifying the uncertainty of our estimate (more details about this metric are given in the next section). Below we outline an algorithm for comparing the allocations:

1. Analyze currently available data.
2. For each of the d potential allocations,
   a. Use reliability estimates for each component, subsystem or system to generate *multiple new data sets* for each type of data in the amount required by the allocation.
   b. For each of the generated data sets, perform a new analysis using the same model as used in Step 1 above but with combined data (original + new simulated data).
   c. Summarize results for allocation using selected uncertainty metric.
3. Compare results for all allocations, and select best one.

Note that the generation of new data uses the assumed model from the original analysis which specifies the distribution from which the data are generated as well as the current reliability estimates for that type of data. Note that the methodology outlined in Anderson-Cook et al (2007,2008) provides estimates of reliability for all components, subsystems and the system, which makes the required data generation possible. For example, suppose pass/fail data for a particular component are assumed to come from a probit model. If we wish to generate new data at a specific age, we would use the posterior distribution for the estimate the reliability at that age from the current analysis. From that, we would generate the required number of new pass/fail observations from a binomial distribution with the probability of a pass equal to the component reliability estimate.

Multiple data sets are required to capture both the uncertainty of the model parameters based on the first analysis and sample-to-sample variability expected if that allocation were selected and the actual data obtained. This generation and analysis of multiple data sets is beneficial for its more accurate assessment of variability, but is computationally quite intensive.

To illustrate the methodology we consider the series system shown in Figure 2 consists of 5 components (Components 1, 2, 3.1, 3.2, and 4) in series, with 8 different types of data available. For a series system, all components need to work for a successful full-system test, and the failure of one or more components will lead to a failed system test. Systems and components of ages 1 through 5 years are available for testing. Interest in prediction of system reliability is for ages 5 to 7 years.

Component 1

The data for Component 1 is Pass/Fail observed at particular ages of the system. We model the probability that Component 1 works at age t using a probit link as

$$\text{Prob}(C_1 = 1 \mid t_1) = \Phi(\alpha_0 + \alpha_1 t_1) \quad (1)$$

The observed data used in the initial analysis consists of Y_{1i} passes in N_{1i} tests at age t_{1i}, for $i = 1,\ldots,n_1$, with the units used at each age assumed to be different. Hence, we can model

$$Y_{1i} \sim Binomial(N_{1i}, \text{Prob}(C_1 = 1 \mid t_{1i})).$$

Component 2

Component 2 has two different types of data which can be used to assess its reliability: degradation data and pass/fail data at particular ages. The continuous measures or degradation data are compared to a specification limit to determine if the component would have passed. A value for the specification limit is known, but there is some uncertainty about whether this value accurately reflects when the component will actually fail during a different type of test which is thought to more accurately reflect how the component is exercised during the full-system test.

The degradation data are assumed to be distributed as

$$Z \mid t_2 \sim Normal(\gamma_0 + \gamma_1 t_2, \delta^2) \quad (2)$$

with observed data Z_{2ij} at age t_{2j} for $i = 1, \cdots, n_{2z}, j = 1, \cdots, n_{2t}$. That is, there are n_{2t} inspection times and at each inspection time, n_{2z} are destructively measured. Therefore, the n_{2z} units at each inspection time are different.

The pass/fail data at various ages for Component 2 is assumed to pass with probability,

$$\text{Prob}(C_2 = 1 \mid t_2) = \Phi\left(\frac{\gamma_0 + \gamma_1 t_2 - D_2}{\sqrt{\sigma^2 + \delta^2}}\right)$$

where the threshold D_2 is not known precisely. That is, the more that z exceeds D_2, the higher the probability of passing the test. We reparameterize so that

$$\text{Prob}(C_2 = 1 \mid t_2) = \Phi\left(\frac{(\gamma_0 + \gamma_1 t_2)/\sigma - D_2^{'}}{\sqrt{1 + \delta^2/\sigma^2}}\right) \quad (3)$$

where $D_2^{'} = D_2 / \sigma$. The observed data used in the initial analysis consists of Y_{2i} passes in N_{2i} tests at age t_{2i}, for $i = 1, \ldots, n_2$, with the units used at each age assumed to be different.

Sub-system 3

Sub-system 3 consists of 2 components, C_{31} and C_{32}, in series. Data for component C_{31} are Weibull lifetimes distributed as $\text{Weibull}(\lambda, \beta)$ where

$$\Pr ob(C_{31} = 1 \mid t) = \exp(-[t/\lambda]^{\beta}) \quad (4)$$

with observed lifetimes $Y_{31i}, i = 1, \cdots, n_{31}$.

Data for component C_{32} consists of degradation data distributed as

$$Y_{32t} \sim Normal(\eta_0 + \eta_1 t, \nu^2) \quad (5)$$

where $\text{Prob}(C_{32}=1\,|\,t)=\text{Prob}(Y_{32t}>D_{32})=1-\Phi\left(\frac{D_{32}-\eta_0+\eta_1 t}{\nu}\right)$. We assume that the threshold D_{32} is known. The observed degradation data are Y_{32ij} at age t_{32j} for $i=1,\cdots,n_{32z}, j=1,\cdots,n_{32t}$. That is, there are n_{32t} inspection times and at each inspection time, n_{32z} are destructively measured.

Data observed at the sub-system level are pass/fail data over time. Pass/fail data at various ages of the sub-section are also available. We observe Y_{3i} passes in N_{3i} tests at age t_{3i}, for $i=1,...,n_3$. The probability of passing at time t_3 is

$$\begin{aligned}\Pr ob(SS_3=1\,|\,t_3)&=\text{Prob}(C_{31}=1\,|\,t_3)\bullet\text{Prob}(C_{32}=1\,|\,t_3)\\&=\exp(-[t_3/\lambda]^{\beta})\bullet(1-\Phi\left(\frac{D_{32}-\eta_0+\eta_1 t_3}{\nu}\right))\end{aligned}\tag{6}$$

Therefore, $Y_{3i}\sim Binomial(N_{3i},\exp(-[t_3/\lambda]^{\beta})\bullet(1-\Phi\left(\frac{D_{32}-\eta_0+\eta_1 t_3}{\nu}\right)))$.

Component 4

Component 4 has pass/fail data and the component is assumed to not age over time. Therefore, we assume

$$\text{Prob}(C_4=1)=p_4\tag{7}$$

We have observed Y_4 passes in N_4 tests, and $Y_4\sim Binomial(N_4,p_4)$

System data

The system consists of five components in series, with the observed pass/fail data at various ages. We have Y_{Si} passes in N_{Si} tests at age t_{Si}, for $i=1,...,n_S$. We express system reliability as

$$\begin{aligned}\text{Prob}(S=1\,|\,t)=&\text{Prob}(C_1=1\,|\,t)\bullet\text{Prob}(C_2=1\,|\,t)\\&\bullet\text{Prob}(C_{31}=1\,|\,t)\bullet\text{Prob}(C_{32}=1\,|\,t)\bullet\text{Prob}(C_4=1)\end{aligned}\tag{8}$$

Hence, we model $Y_{Si}\sim Binomial(N_{Si},\text{Prob}(S=1\,|\,t_{Si}))$ at each time considered.

Since the best allocation will be dependent on the data already collected, Table 1 describes the data observed for each of the eight data types for the various ages. Given the costs given in Figure 2, the total cost of the initial data is $40,000 (400 system tests each at $50 = 20,000; 400 component 1 tests each at $10 = $4,000, etc.). Note that the lifetime data for component 3.1 and the non-aging pass/fail data for Component 4 are not associated with any particular age, and hence were not performed on components of a specified age, as the other forms of data are. Figure 3 shows estimates of the various components and the system with uncertainty (90% credible intervals) based on the initial data using a Bayesian approach.

Suppose than an additional $30,000 is available to collect more data. Based on logistical and practical constraints, possible allocations are identified in Tables 2 and 3. Available systems are between 1 and 5 years old, and sufficient components and parts from a large population are available to allow for all of the suggested allocations. Twenty-one allocations were considered: seven divisions of the budget across data types are considered, with three distributions of ages for each of the seven divisions. Collecting the component data naturally grouped into two sets of components: Group 1 components are those with costs of $10 per observation (C1, SS3 and C4), while Group 2 components are those with $5 per observation costs (C2a, C2b, C31 and C32). For example, allocation 3B (half system data and half component data with more older data) would contain 300 new observations from all eight data types with 20 observations of each type of age 1, 40 observations of age 2, …, and 100 observations of age 5. If for a particular allocation the total number of observations is not divisible by 15, then the number of observations at each of the 5 ages is rounded to the nearest observation.

The objective of this resource allocation problem is to identify the best allocation to maximally improve prediction of system reliability for systems aged 5 to 7 years. In the following sections, we describe the criteria used to compare allocations (Section 2), and then provide additional details for the algorithm described above to estimate the expected improvement in precision for each allocation (Section 3), before providing results in Section 4. Section 5 discusses some strategies for generalizing the solution to give an optimal solution across the range of data types, beyond just the specific allocations identified.

2. Criteria for Comparing Allocations

In determining the best allocation, we want to select based on a reduction of the uncertainty of the estimation of system reliability. There are number of consideration in selecting a single number summary of the improvement. We should consider the metric of uncertainty, and where we wish to predict reliability. In addition, given that we will be simulating data from the estimated reliability distributions using the posteriors of the various parameters, we should select an appropriate summary across the range for samples generated.

Several possible measures of uncertainty are possible for any age at which we wish to predict. We could look at the variance of posterior distribution for system reliability, the width of a particular $(1-\alpha)100\%$ credible interval, or the entropy of the estimate. Wynn (2004) shows that under certain restrictive conditions which are not applicable to our particular situation, these different measures of uncertainty are all asymptotically equivalent. However, in general and in our example, these different measures of uncertainty may lead to different relative rankings of the allocations. Hence it is important to consider how the results of the new analysis will be used and select a metric that most appropriately summarizes that aspect of uncertainty. In our example, system reliability will be reported with the median of the posterior, and a 90% credible interval. Hence the width of the 90% credible interval at a particular age is a sensible measure of uncertainty.

The objective of the study is to predict reliability well for systems with ages between 5 and 7 years. Several potential choices of metric might make sense: we may wish to select a single age (say age 6) and compare the width of credible intervals at that age. Alternately, we may wish to integrate the area between the credible interval lines across the range of ages, or consider a weighted average of several specific ages (with weights selected to reflect our relative interest in different ages). While the integrated area may be the most precise summary of uncertainty across the range of ages, we select a relatively simple proxy for this by considering the arithmetic average of the widths of the credible intervals at ages 5, 6 and 7 years.

Finally, we need to select an appropriate summary across the various samples of data generated for each allocation. A couple of intuitive choices would be to look at the median or average width across the samples or an upper percentile. The median or average would represent a "typical" improvement with the new data, while an upper percentile would estimate a "worst case" improvement. In our case, we consider both the median and 90th percentiles of the average width of the 90% credible intervals, where we average over the system estimates at ages 5, 6 and 7 years.

As a baseline, based on the Table 1 data, the width of the intervals at ages 5, 6 and 7 years are 0.071, 0.136 and 0.214, respectively. The average of these widths is 0.140. When new data are added to the analysis, we would expect that each of these intervals would become narrower.

3. Details of Algorithm for Assessing Allocations

In this section we consider some of the details for performing the algorithm described in the Introduction for the example. Step 1 uses a Bayesian approach and assumes that we have a well-defined model which allows us to write down the likelihood for all the types of data that we have observed. Equations (1)-(8) provide a mechanism for including each type of observed data into a global likelihood which is a function of the 13 parameters of the model. Component 1 has parameters α_0, α_1, Component 2 has $\gamma_0, \gamma_1, D_2^{'}, \delta^2, \sigma^2$, Component 3.1 has λ, β, Component 3.2 has η_0, η_1, ν^2, and Component 4 has parameter p_4. Let Θ denote the vector of the 13 model parameters. The Bayesian approach combines prior information about Θ with the information contained in the data. The prior information is described by a prior density $\pi(\Theta)$ and summarizes what is known about the model parameters before any data are observed. Here, we assume that little is known, and therefore choose diffuse proper prior distributions, which allow for the possibility of a wide range of values for the model parameters. ($\alpha_0, \alpha_1, D_2^{'}, \gamma_0, \gamma_1, \eta_0, \eta_1 \sim Normal(0, 10^2)$, $\delta, \sigma, \lambda, \beta, \nu \sim Gamma(1,1)$, $p_4 \sim Beta(9,1)$) The information provided by the data is captured by the data sampling model $f_y(\mathbf{y} \mid \Theta)$ known as the likelihood, which is based on Equations (1)-(8). The combined information is described by the posterior density, $\pi(\Theta \mid \mathbf{y})$. We evaluate the posterior density using Bayes' Theorem [Degroot (1970)] as

$$\pi(\Theta \mid \mathbf{y}) \propto f(\mathbf{y} \mid \Theta)\pi(\Theta).$$

When the form of the posterior density is well known, the distributional form of the posterior density can be obtained in closed form. For more general forms of the posterior density, we can use recent advances in Bayesian computing to approximate the posterior distribution via Markov chain Monte Carlo [Gelfand and Smith (1990), Casella and George (1992), Chib and Greenberg (1995)]. That is, Markov chain Monte Carlo (MCMC) algorithms produce draws (i.e., samples) from the joint posterior distribution of Θ by sequentially updating each model parameter conditional on the current values of the other model parameters. These draws of the posterior of Θ are easy to work with in evaluating system reliability at a given age which is a function of Θ. That is, we obtain a draw from the posterior of system reliability at a given age by evaluating the system reliability with a draw of the posterior of Θ.

As a result of the Bayesian analysis performed based on the initial data, we obtain posterior distributions for all 13 parameters. To incorporate the uncertainty in the estimates of model parameters, values are sampled from the posterior of each parameter and new data are generated based on these values. For example, for allocation 1A, we wish to generate 120 (600*1/5) new observations at each of ages 1, 2, …, 5. A draw from the joint posterior of all 13 parameters is selected, say draw *(k)*, yielding parameter values $(\alpha_0^{(k)}, \alpha_1^{(k)}, \cdots, p_4^{(k)})$. The probability of a successful system test is estimated at age 1 using $\text{Prob}(S=1 \mid t=1)$ in equation (8) with the draw values. Then a binomial would be generated $Y_{S(1)}^* \sim Binomial(120, \text{Prob}(S=1 \mid t=1))$. This data would then be combined with the original data to give the total number of successes $Y_{S(1)} + Y_{S(1)}^*$ out of (80 + 120) = 200 test of systems at age 1. The process would be repeated for ages 2 through 5.

For the different allocations, different types of new data are generated. All of the generated data are created using the estimated model parameters from the initial analysis.

For many of the allocations, new data for all or most of the data types are required at each of ages 1 through 5 years. For Component 3.1, new lifetime data are created consistent with the estimated parameter values $\lambda^{(k)}, \beta^{(k)}$. Since there is no aging for Component 4, new binomial data are generated using the draw estimate $p_4^{(k)}$. In our example, for each allocation, we sample 100 draws from the posterior, and generated a new data set for each, and then re-ran the analysis with the combined data. Once the new parameter posteriors are obtained, the width of the 90% credible interval for system reliability at each of ages 5, 6 and 7 years was calculated. Then an average of these three widths was obtained for each new analysis. The results from these 100 new analyses were used to find the median average width and the 90th percentile average width for each allocation.

It should be apparent from the description above, that this approach to estimating the improvement in precision for a particular allocation can be computationally very intensive. For our simple example where we are comparing 21 different allocations, 21 x 100 = 2100 different data sets need to be created and then combined analyses of the original data with a particular generated data set performed. If each analysis is time-consuming, then resource allocation will become a very time intensive procedure.

4. Results for Example

In this section we consider the results from our comparison of the 21 possible allocations. Table 4 summarizes the median and 90th percentile average widths based on the 100 generated samples and new analyses for each allocation. Recall that the average width for the original data was 0.140. By changing the total budget from \$40,000 to \$40,000+\$30,000=\$70,000, we are able to realize substantial reduction in the width of the intervals. The best allocation for both criteria is 5B, with the narrowest predicted median and 90th percentile average credible intervals.

As we might have expected, the "more older" data allocations (B) are consistently best for predicting in the age ranges 5 to 7 years. In addition, the allocations with more of the group 1 component data (allocations 4 and 5) perform well. By examining Figure 3, we can see that group 1 components tend to correspond to components with lower reliablities at older ages. Recall that the variability of proportions becomes larger as the probability of success moves away from one. Hence by obtaining more data for these data types we are able to reduce uncertainty more substantially than for a highly reliable component.

Based on these results the best available allocation to maximally reduce the uncertainty of our system reliability predictions for ages 5, 6 and 7 years is to collect new data as listed in Table 5. Recall that the average width of the 90% credible interval based on the original data was 0.140. By adding the $30,000 of additional data, we can expect that there is a 50% chance of reducing the average width of the interval to less than 0.079 (a 43% reduction) and a 90% chance of reducing it to less than 0.092 (a 34% reduction).

It should be noted that some of the allocations have an observed 90^{th} percentile that is actually wider than that observed in the original analysis. This suggests that if we choose a bad allocation of new data to collect, the uncertainty in our model estimates may not be improved. While this may not seem possible intuitively, one explanation might be that sampling variability from some new data may actually introduce some additional uncertainty, instead of reducing it. It should be noted that this is only occurs for the 90^{th} percentile, as this represents a "worst case" reduction of uncertainty across possible data sets consistent with the model parameter estimates.

5. Modeling Allocation Results as a Cost-Based Mixture Experiment

Sometimes instead of being asked to select a best allocation from a list of possible allocations, we are allowed to suggest a best allocation subject to some constraints. In this case, we may wish to characterize the allocation space with a response surface model, and use optimization techniques to find a best allocation. We now consider how this might work for our example. We can think of the allocation space as a mixture-process space (Cornell, 2002), where the mixture variables are the proportion of cost for the new allocation for each data type, and the process space allows us to consider the possible distribution of ages within a data type. In this case, since there are practical restrictions on how much data from any year can be considered, we create a single process factor with values -1 corresponding to "less older data", 0 for equal data at all years, and +1 for "more older data".

Recall that mixture variables are subject to the constraint that the proportion of "ingredients" must sum to 1. In our case, we require that the sum of the proportions of the costs sum to 1, which implies that each allocation that we consider has the same total cost. This is quite natural as we would likely want to examine the reduction in width of the credible intervals for comparable allocations with the same cost. If we consider the allocations 1 and 2, we can illustrate how to represent these in mixture experiment notation. Let X_i represent the proportion of the total cost from each data type. Then allocation 1 corresponds to 100% of the budget being spent on system data. Therefore, allocation 1 would be represented as

$$(X_S, X_{C1}, X_{C2a}, X_{C2b}, X_{C3}, X_{C31}, X_{C32}, X_{C4}) = (1, 0, \cdots, 0).$$

Allocation 2 has new data for each of the component types, with 600 new observations for each of C1 through C4. Some of the component data costs \$10 per observation, while for others the cost is \$5 per observation. Hence we would summarize the allocation 2 as a mixture as

$$(X_S, X_{C1}, X_{C2a}, X_{C2b}, X_{C3}, X_{C31}, X_{C32}, X_{C4}) = (0, .2, .1, .1, .2, .1, .1, .2)$$

Since 600*\$10 = \$6000 for component 1 represents 20% of the total available new budget. Hence, we can model any allocation considered as a mixture combination based on cost. We can then fit a second order response surface model to the observed responses, here the median or 90th percentile average width of the ages 5, 6 and 7 years with the model

$$Y = \sum_i \beta_i X_i + \sum \sum_{j \neq i} \beta_{ij} X_i X_j + \sum_i \theta_i X_i P + \varepsilon \qquad (9)$$

where the $\beta_i X_i$ are the mixture main effect terms, the $\beta_{ij} X_i X_j$ are the mixture second order effect terms, and the $\theta_i X_i P$ are the mixture-process interactions. Recall that because of the constraint that all ingredients must sum to 1, there are no intercept, process main effects and no pure quadratic mixture terms in the second order mixture-process model. The usual regression assumption of $\varepsilon \sim iid\, N(0, \sigma_\varepsilon^2)$ apply to the error term.

Because we have eight data types, this leads to a model in (9) with 8 + 28 + 8 = 44 terms, but we only have 21 allocations with which to estimate this model. Hence, we propose to consider a simplified version of the model. In this case, since there are restrictions on collecting equal numbers of observations for group 1 data types and equal numbers for group 2 data, it is natural to re-express the model in terms of three ingredients: system, group 1 components and group 2 components. This leads to mixture combinations of

$$(X_S, X_{CG1}, X_{CG2}) = (1,0,0)$$

and

$$(X_S, X_{CG1}, X_{CG2}) = (0,.6,.4)$$

for allocations 1 and 2, respectively. The proportions for allocation 2 are based on 60% (or $18000) of the budget being spent on data for components 1, 4 and sub-system 3. The remainder of the budget ($12000) is spent on component data of types 2a, 2b, 3.1 and 3.2. Allocations 3 through 7 correspond to the following mixture combinations based on the new groupings: (.5,.3,.2), (.5,.5,0), (0,1,0), (.5,0,.5), (0,0,1), respectively. This grouping of data types into 3 categories leads to a simplified model with just 3 + 3 + 3 = 9 terms, which is easily estimable with our 21 allocations.

Figure 4 shows the design space for possible allocations based on the reduced number of data types and the simplified process variable structure for the distribution of the different ages of systems explored. Figure 5 shows the 7 allocations from Table 2 with the new grouping of components. Note that these allocations are well distributed throughout the mixture region, and hence should allow good estimation of the model parameters.

The model was estimated based on the 21 allocations with the results for both criteria shown in Figures 6 and 7. The models fit the observed allocations well with $R^2_{adj} = 99.85\%$ and 99.85% for the median and 90^{th} percentile of the average widths, respectively. In this case, the best allocation for maximally reducing the uncertainty in the prediction of system reliability subject to the restriction of sampling by groups 1 or 2 for the component data corresponds to exactly what was sampled with allocation 5B. This allocation was observed to give a median average width of 0.79, and is predicted from our model to give a value of 0.80. Similarly for the 90^{th} percentile of the average width, the observed and predicted values are very similar.

Alternately, since it is clear from the initial results that the "more older data" are consistently best, we may wish to exclude the other distributions and only focus on appropriate modeling of the allocations with this distribution of ages in a mixture model of the form $Y = \sum_i \beta_i X_i + \sum \sum_{j \neq i} \beta_{ij} X_i X_j + \varepsilon$. We also note that since the data in the response surface modeling are simulated, it is possible to propose and evaluate allocations outside of the constraints required for actual allocations. This provides an opportunity to consider a designed experiment in the allocation space that allows for good estimation of the response surface. If we selected possible allocations that did not keep the proportions of component data fixed within group 1 and 2, we would be able to explore if the grouping of components into these practically convenient groups is advantageous from the prediction perspective.

Regardless of the approach, it is beneficial to evaluate the proposed optimal allocation directly as well. In our example, the best allocation turns out to be one that we have already evaluated. This will not be true in general. By comparing the responses from the confirmatory run with those predicted by the response surface model, we can assess the goodness-of-fit of the model and also verify that the suggested allocation matches expected model results.

6. Discussion and Conclusions

In the example considered, the relative cost of the system data relative to obtaining measures on the individual components is equal. Namely, both a single system observation and one observation for each component cost \$50. We also considered changing the relative cost of these, by changing the system data costs to \$25 and \$100 per observation. This did change the relative ranking of some of the allocations, but for each of the new scenarios, allocation 5B remained the best choice. This is perhaps not surprising since this allocation does not involve any system level data, and the system level data does not appear to be as beneficial as additional component level data. Typically, we would expect that changing the relative costs of the different data types would lead to different best allocations based on a fixed budget.

The suggested best allocation is highly dependent on what original data were used in the first analysis as well as on the reliability of the individual components. Data types that are already abundant represented in the original data are less likely to yield substantial improvement in the reduction of prediction uncertainty, as there are diminishing returns on additional data. As well, components that are highly reliable have less associated uncertainty in their estimation, and hence require less total data to feel confident that they are well understood. The ideal candidate for a large proportion of the new allocation is a data type that has relatively less data and has reliability that is less reliable or changing over time.

Given that there may be some uncertainty about the correctness of the system structure model, it may be beneficial to consider including some system level data when possible to validate the statistical model used to combine the component data into a system reliability estimate. See Anderson-Cook (2008) for details on testing the series assumption for a system. Additionally, including a measure of discrepancy in the model to allow estimation of potential differences in system reliability estimate from various data sources may also be advisable.

In this paper we propose an algorithm for evaluating different allocations of resources. With a focus on predicting reliability well at user-specified ages of interest, and by accounting for the uncertainty associated with both our parameter estimates and sampling variability from the new data, we are able to rank competing allocations and their potential value. As an initial approach, we compare a fixed number of possible allocations and determine a best allocation. However, it may also be desirable to estimate an optimal allocation within the ranges of allocations consistent with logistical and practical constraints. By fitting a mixture or mixture-process response surface model to the results of the evaluated allocations, we can find optimal proportions of the budget to be spent on each data type. A further extension would be to use an automated search algorithm, such as a genetic algorithm, to find a best allocation, subject to the constraints.

The methodology proposed would be appropriate for any meta-analysis where we have different data types which we wish to combine in a single analysis to estimate a primary response of interest. Integrating the relative cost of the data is important, as for different data types the associate costs can vary substantially and this will impact the amount of a particular data type that can be obtained subject to the available resources. The general nature of the algorithm makes it widely applicable to various scenarios, but the computationally intense nature of the algorithm may mean that it is not practical for some applications where a single analysis takes considerable time and computer power.

Acknowledgement

We that C.C. Essix for support and encouragement of this work.

References

1. C.M. Anderson-Cook, "Evaluating the Series or Parallel Structure Assumption for System Reliability" *Quality Engineering* (in press), 2008.
1. C.M. Anderson-Cook, T. Graves, M. Hamada, N. Hengartner, V. Johnson, C.S. Reese, A.G. Wilson, "Bayesian Stockpile Reliability Methodology for Complex Systems" *Journal of the Military Operations Research Society*, Vol. 12, pp. 25-37, 2007.
2. C.M. Anderson-Cook, T. Graves, N. Hengartner, R. Klamann, A.K. Wiedlea, A.G. Wilson, G. Anderson, G. Lopez, "Reliability Modeling Using Both System Test and Quality Assurance Data" *Journal of the Military Operations Research Society* (in press), 2008.
3. G. Casella, E. George "Explaining the Gibbs Sampler," *The American Statistician*, 46, pp. 167-174, 1992.
4. S. Chib, E. Greenberg, "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49, pp. 327-335, 1995.
5. J. Cornell, *Experiments with Mixtures: Design, Models, and the Analysis of Mixture Data*. New York: Wiley, 2002.
6. M.H. DeGroot, *Optimal Statistical Decisions*. New York: McGraw-Hill, 1970.
7. A.E. Gelfand, A.F.M. Smith, "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, pp. 398-409, (1990.
8. M. Rausand, A. Hoyland, System Reliability Theory: Models, Statistical Methods, and Applications. New York: Wiley, 2004.
9. S.C. Saunders, *Reliability, Life Testing, and Prediction of Service Life*. New York: Springer, 2007.
10. A.G. Wilson, T.L. Graves, M.S. Hamada, C.S. Reese, "Advances in Data Combination, Analysis and Collection for System Reliability Assessment" *Statistical Science*, Vol. 21, pp. 514-531, 2006.
11. H.P. Wynn, "Maximum Entropy Sampling and General Equivalence Theory" Advances in Model-Oriented Design and Analysis, Contributions to Statistics, Physica, Heidelberg, pp. 211-218, 2004.

Table 1: Data Types for System

		Amount of Initial Data					
	Data Type	**Age 1**	**Age 2**	**Age 3**	**Age 4**	**Age 5**	**No Age**
System	P/F at age	80	80	80	80	80	
Comp 1	P/F at age	80	80	80	80	80	
Comp 2 (a)	Testset degradation	80	80	80	80	80	
(b)	P/F at age	80	80	80	80	80	
Sub-Sys 3	P/F at age	80	80	80	80	80	
Comp 3.1	Lifetime						400
Comp 3.2	Testset degradation	80	80	80	80	80	
Comp 4	P/F						400

Table 2: Allocations Considered

Allocation	System	C1	C2a	C2b	SS3	C31	C32	C4
1 (all system)	600							
2 (all component)		600	600	600	600	600	600	600
3 (½ sys, ½ comp)	300	300	300	300	300	300	300	300
4 (½ sys, ½ grp 1 comp)	300	500			500			500
5 (all grp 1 comp)		1000			1000			1000
6 (½ sys, ½ grp 2 comp)	300		750	750		750	750	
7 (all grp 2 comp)			1500	1500		1500	1500	

Table 3: Distribution of Ages for Aging Data (System, C1, C2a C2b, SS3, C32) as a Proportion of Total Number of Observation for That Data Type

Allocation	Age = 1	Age = 2	Age = 3	Age = 4	Age = 5
A (equal ages)	1/5	1/5	1/5	1/5	1/5
B (more older)	1/15	2/15	3/15	4/15	5/15
C (less older)	5/15	4/15	3/15	2/15	1/15

Table 4: Results for Two Criteria for Five Component System Based on Average Width of Credible Interval at Ages 5, 6 and 7 Years. The best allocation for each criterion is shown in bold.

Allocation	Median Width			90th Percentile Width		
	A	**B**	**C**	**A**	**B**	**C**
1	0.121	0.106	0.127	0.137	0.125	0.144
2	0.095	0.088	0.106	0.110	0.104	0.122
3	0.104	0.095	0.116	0.123	0.113	0.132
4	0.095	0.090	0.109	0.109	0.101	0.128
5	0.083	**0.079**	0.099	0.099	**0.092**	0.114
6	0.126	0.123	0.134	0.143	0.139	0.146
7	0.134	0.135		0.142	0.145	

Table 5: Best Available Allocation for Minimizing the Median and 90th Percentile of the Average Width of the Age 5, 6 and 7 Year System Reliability Prediction

	Age = 1	Age = 2	Age = 3	Age = 4	Age = 5	No Age	Total
System	0	0	0	0	0		0
C1	67	133	200	267	333		1000
C2a	0	0	0	0	0		0
C2b	0	0	0	0	0		0
SS3	67	133	200	267	333		1000
C31						0	0
C32	0	0	0	0	0		0
C4						1000	1000

Figure 1: Overview of resource allocation problem.

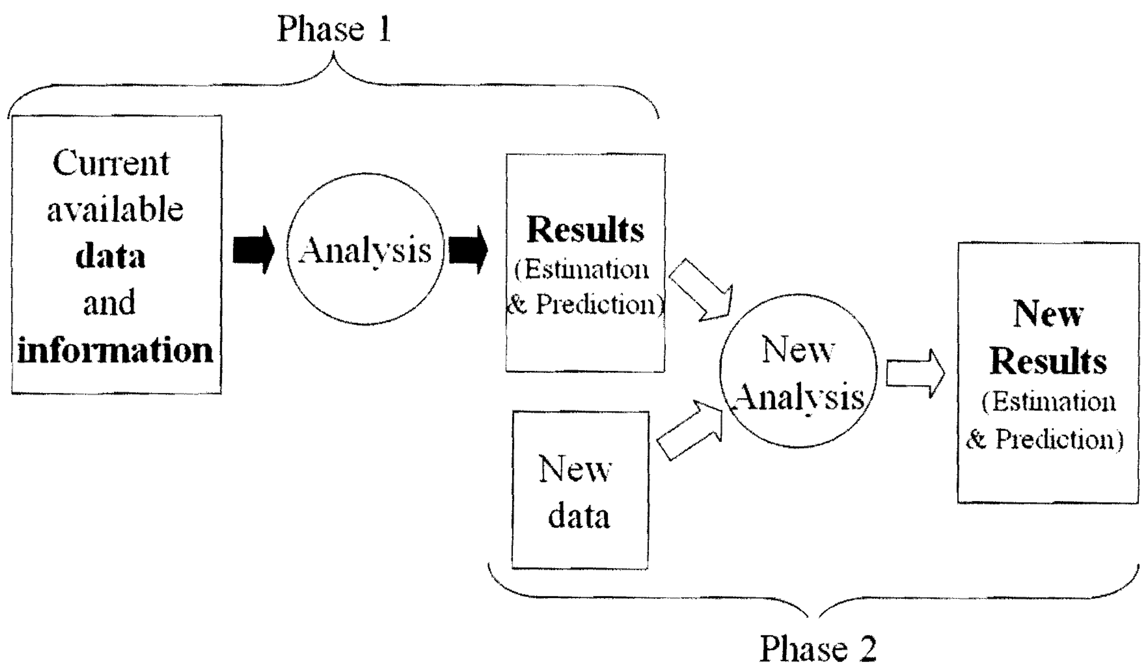

Figure 2: A simple example of a series system with five components and eight possible data types. The data types and cost per observation are shown in grey.

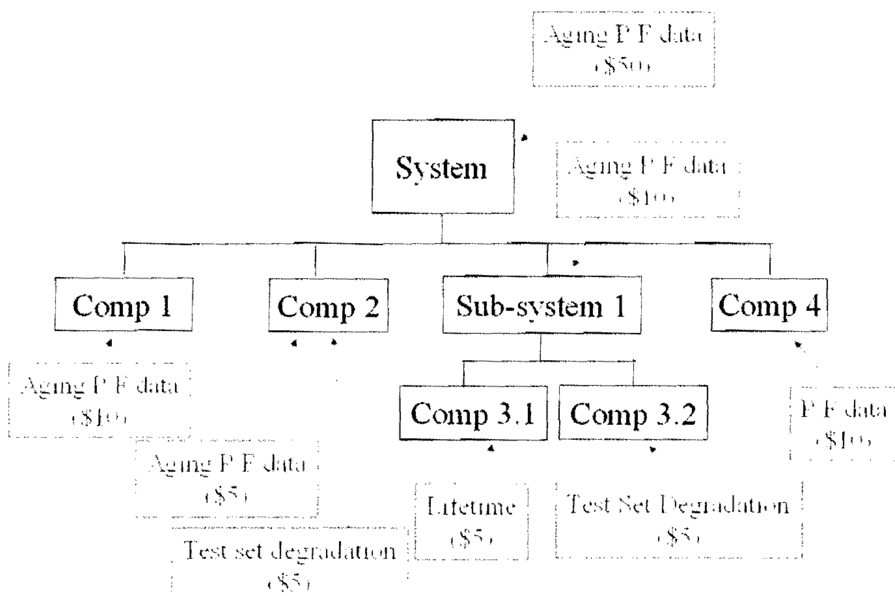

Figure 3: Reliability estimates (median – solid line, 90% credible interval – dashed lines) for components and system for ages 0 to 7 years based on initial data described in Table 1.

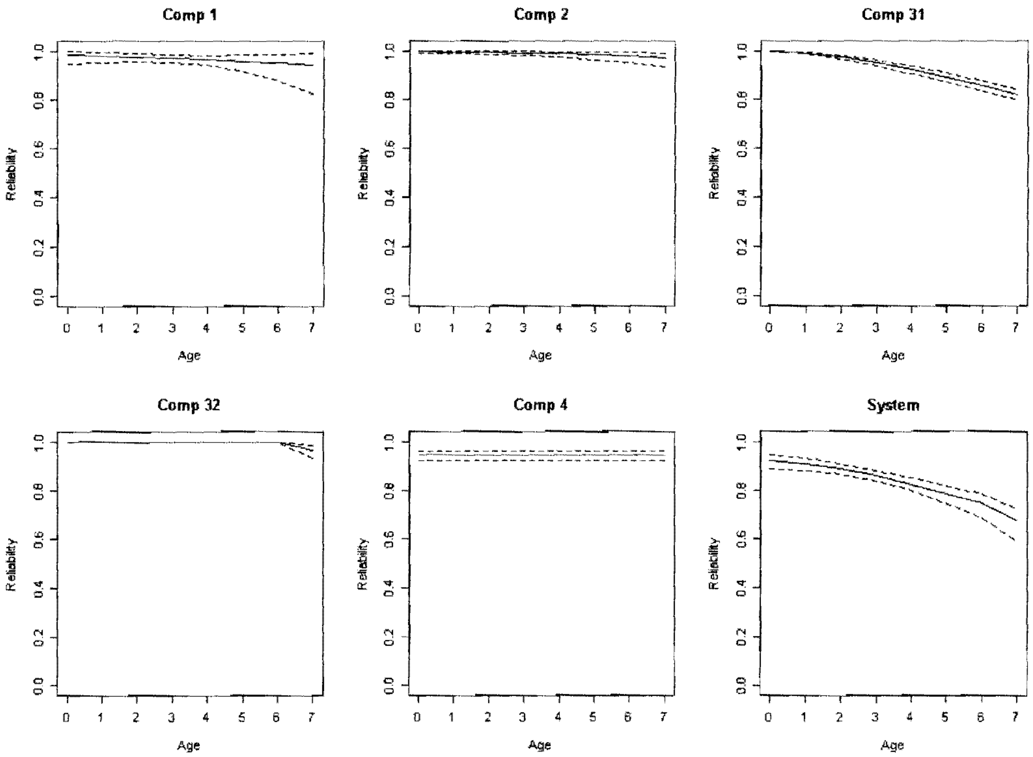

Figure 4: Mixture-Process space for simplified version of the model with three data types (system, group 1 components, group 2 components) and distribution of ages as a single process variable.

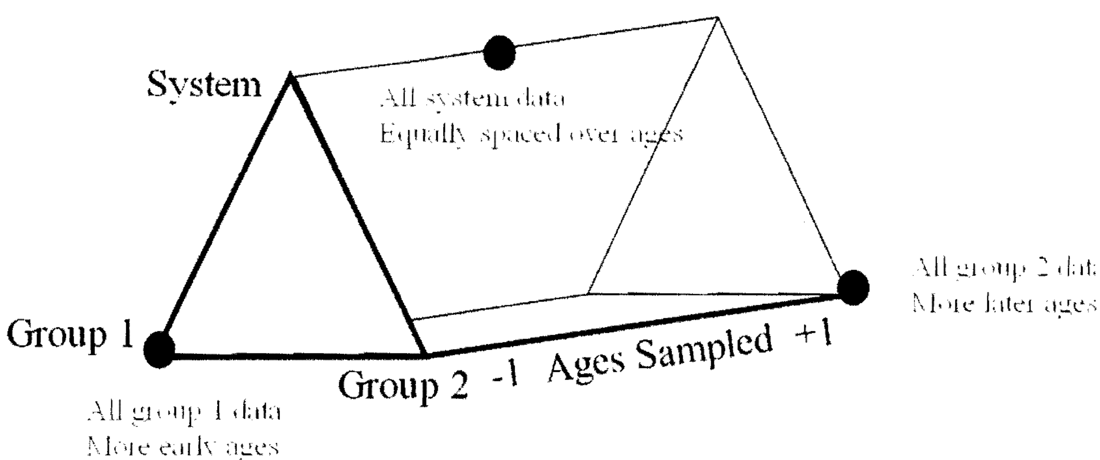

Figure 5: Allocations re-expressed in terms of System data, and Group 1 and 2 components.

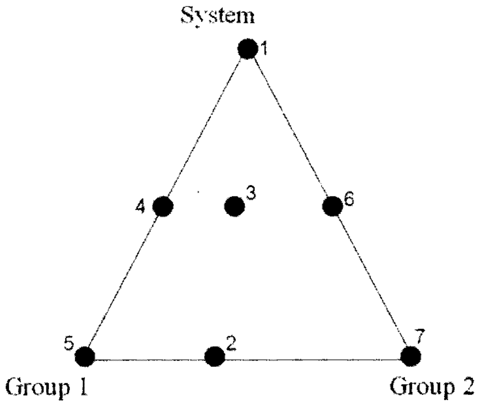

Figure 6: Contour plot of response surface model for median of average widths for ages 5, 6 and 7 years. The optimal allocation corresponds to 100% of the data being sampled from Group 1 components.

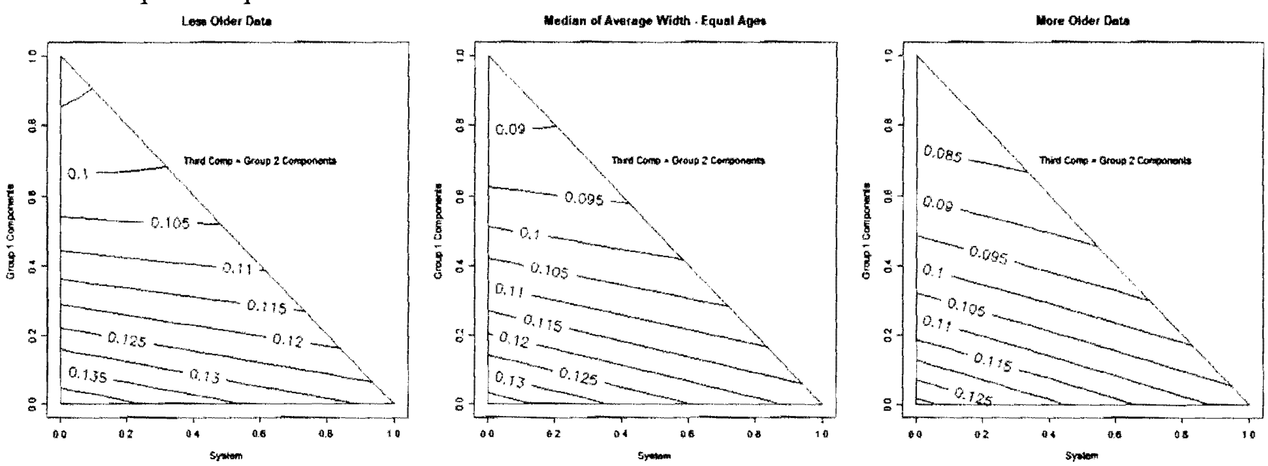

Figure 7: Contour plot of response surface model for 90th percentile of average widths for ages 5, 6 and 7 years. The optimal allocation corresponds to 100% of the data being sampled from Group 1 components.