LA-UR- 10-01391

Approved for public release; distribution is unlimited.

Title: Resource Allocation: Sequential Data Collection for Reliability Analyses Involving System and Component Level Data

Author(s): C.M. Anderson-Cook

Intended for: JANNAF Meeting
Colorado Springs, CO
May 3, 2010

Los Alamos NATIONAL LABORATORY
EST. 1943

RESOURCE ALLOCATION: SEQUENTIAL DATA COLLECTION FOR RELIABILITY ANALYSES INVOLVING SYSTEM AND COMPONENT LEVEL DATA

Christine M. Anderson-Cook
Statistical Sciences Group
Los Alamos National Laboratory
Los Alamos, NM

ABSTRACT

In analyzing the reliability of complex systems, several types of data from full-system tests to component level tests are commonly available and are used. After a preliminary analysis, additional resources may be available to collect new data. The goal of resource allocation is to identify the best new data to collect to maximally improve the prediction of system reliability. While several possible definitions of "maximally improve" are possible, we focus on reducing the uncertainty or the width of the uncertainty interval for the prediction of system reliability at a user-specified age(s). In this paper, we present an algorithm that allows us to estimate the anticipated improvement to the analysis with the addition of new data, based on current understanding of all of the statistical model parameters. This quantitative assessment of the anticipated improvement can be helpful to justify the benefits of collecting new data. Additionally by comparing different potential allocations, it is possible to determine what new data should be collected to improve our understanding of the response. This optimization takes into account the relative cost of different data types and can be based on flexible allocation options, or subject to logistical constraints.

INTRODUCTION

For many populations of complex systems, there is a need to monitor, manage and maintained these systems over time. Typically, these systems are thought to degrade as a function of age and usage, and it is important to be able to estimate the fraction of units within a given population that will perform as expected at any given time. One approach to obtain an accurate estimate of the reliability is to rely on full system tests on a sample of units from the population, and use the results of full-system tests to obtain estimates. While this approach is commonly used and does provide the needed estimates, it tends to be an expensive approach to answering this question, given the expensive and frequently destructive nature of these tests.

Complex systems are frequently comprised of well-defined subsystems, which in turn are composed of many components. The reliability of such systems can be more effectively and efficiently assessed by using all possible relevant sources of data. In addition to the full system tests, which may be destructive (in the case of a weapon system) and very expensive, it may be possible to obtain measurements for subsystems or components to verify that key elements are in appropriate working order. With a system model that details how the various elements of the system are interrelated and affect the system performance, this subsystem data can be utilized to increase precision of the estimated system reliability. This combined experiment, or meta-analysis, using multiple types of data can be a cost-effective and efficient way to model the system characteristics of interest. See Anderson-Cook (2009a) for more details on meta-analyses.

Figure 1 shows a relatively simple sample system comprised of 5 components where different types of data are available at the system, subsystem and components levels. Typically different costs are also associated with the various data types. Anderson-Cook et al. (2007 and 2008) outline Bayesian methodology which can combine data from these various levels and expert knowledge in the form of priors to obtain an estimate of system reliability. 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 types of data that might be available to assess the individual boxes in the figure include:

1. pass/fail data evaluated at a given age of the part
2. degradation data consisting of a continuous measure with known operational limits, outside of which the part is not expected to work successfully
3. lifetime data\
4. pass/fail data not associated with any particular age for a component that is not believed to age over time

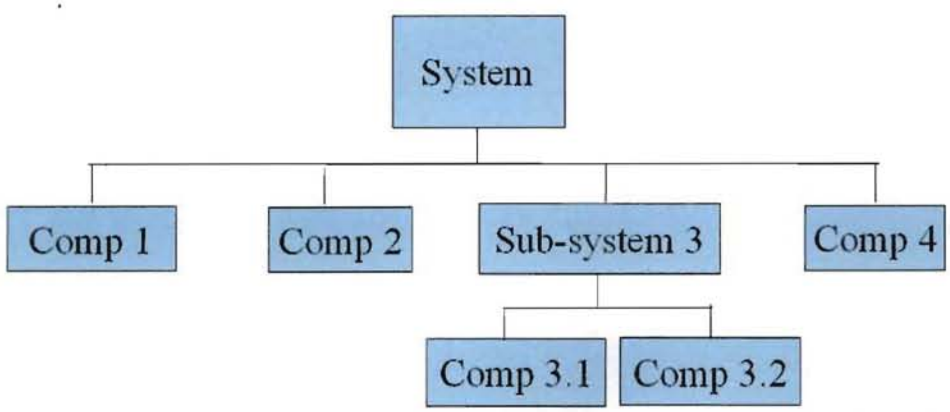

Figure 1: Sample series complex system with multiple data sources available at the various levels

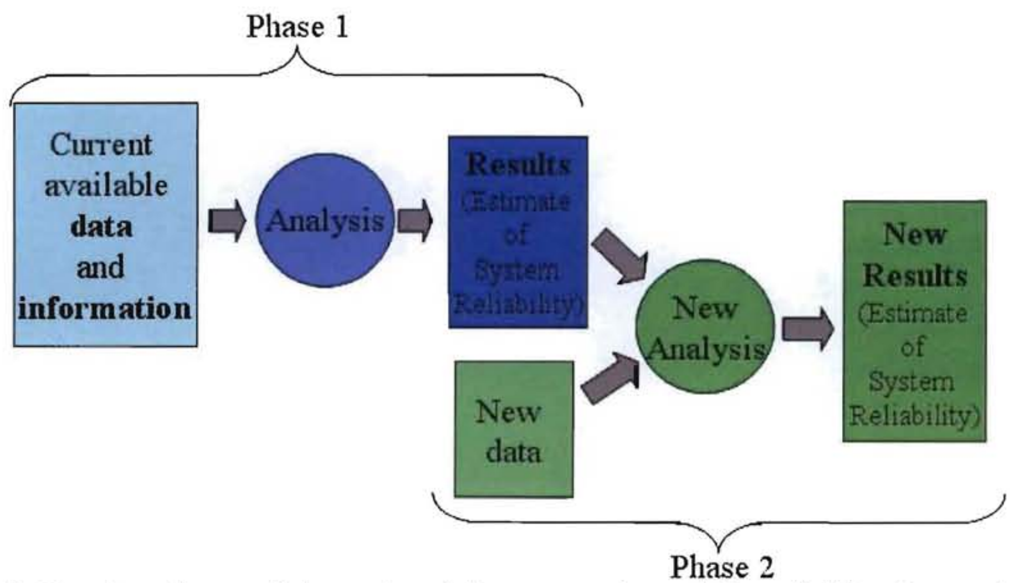

Figure 2: Overview of sequential experimentation process for modeling reliability of a complex system

In this paper we consider the situation where some data have been collected for each type of available data, and a preliminary analysis has been performed as part of Phase 1 shown in Figure 2. The goal of resource allocation is to determine the best collection of new data to obtain that will maximally improve our understanding of system reliability. This decision-making process 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. It should be clear that by obtaining data we cannot change the true value of the system reliability, but rather can only hope to estimate this true reliability more precisely. 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.

In the remainder of the paper, we discuss how potential allocations can be specified, and outline an algorithm for suggesting a best allocation given available resources. We present some of the considerations for selecting an appropriate metric for characterizing our understanding of system reliability, and the need for including a measure of discrepancy for conflicting data between different types of data. For a complete worked example based on the system described in Figure 1, see Anderson-Cook, Graves and Hamada (2009).

SELECTING ALLOCATIONS

In the introduction, we stated that one of the motivations for using component or subsystem level data was that full system tests can be prohibitively expensive. Hence it is natural to think of cost as being an important consideration in determining which allocations should be selected. Our approach has been simultaneously consider cost and reduction of uncertainty in our selection of a best allocation. In order to proceed, the relative cost of each of the sources of data should be assessed. For example for the system defined in Figure 1, the full system test might cost $200, while the costs for the subsystem 3 is $25, and the cost for data on each of the components is $10. As is typical for many of the systems we have considered, the cost of doing the full system test directly is more expensive than collecting information about all of the components separately (Here, data for Components 1, 2, 3.1, 3.2, 4 cost 5x$10 = $50, or components 1, 2, 4 and subsystem 3 cost 4x$10 + $25 = $55). Hence there may be considerable advantage to using more lower level data in the analysis, based on information per unit cost.

In many situations, there is a fixed budget available for the next set of data to be collected. In this case, reasonable allocations can be found by dividing the available budget completely between the available data types. For example, it may be the case that $10000 is available for additional data for improving our reliability estimate for the system given in Figure 1. Several possible allocations that would use the entire budget would include:

1. All full-system data: 50 test @ $200 / test = $10000
2. All component level data, with equal amounts for each component:
   200 tests x 5 components @ $10 / test = $10000
3. All subsystem 3 data: 400 test @ $25 / test = $10000
4. All component 4 data: 1000 test @ $10 / test = $10000

Finally, to completely specify the allocation, for some of the data types we need to specify the age of the system that we wish to test. Consider a particular component which yields pass/fail data when tested. In this case, an allocation would need to specify what age of system should be selected for testing. Hence a completely specified allocation for 1 above might take the form: 50 full system tests with 10 systems of age 2 years, 10 systems of age 4, 10 systems of age 6, 10 systems of age 8 and 10 systems of age 10.

In addition to wishing to spend the entire budget, there are often practical or logistical constraints on what data could be obtained. For example, it might be the case that the component level data requires some disassembly of the system. In this case it might make sense to obtain component level data on all components for that system once it has been disassembled. This would mean that allocation 2 above would be a sensible allocation to consider, while allocation 4 is not. Similarly, specifying the ages of systems to be tested is subject to their availability in the population of systems. There may also be logistical constraints about which systems can be brought in from the field.

ALGORITHM FOR FINDING THE BEST ALLOCAITION

In this section we outline an algorithm for comparing different potential allocations of future resources to determine a best candidate. We assume that several potential allocations have been identified, and we wish to select the best of these. At the conclusion of this section, we briefly outline some more general approaches that would allow a more comprehensive search for an optimal allocation.

Suppose that we have *k* 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 *k* potential allocations,
   a. Use reliability estimate 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 new analysis using the same model as used in 1. above but with combined data (original + new simulated data).
   c. Summarize results for uncertainty metric.
3. Compare results for all allocations, and select best one.

Note that the generation of new data uses the current reliability estimates for that type of data, and uses the assumed model from the original analysis specifies the assumed distribution from which the data are generated. Note that the methodology outlined in Anderson-Cook et al (2007 and 2008) provides estimates of reliability for all components, subsystems and the system, which makes the required data generation possible. The Statistical Sciences group at Los Alamos has developed a Python-based computer package, called SRFYDO, which can implement a system reliability analysis using multiple sources of data. The inputs for the software are Excel spreadsheets which capture the system structure, data and priors. See Anderson-Cook, Huzurbazar, Klamann and Morzinski (2008) for more details.

For example, suppose pass/fail data for a particular component are assumed to come from a probit model. New data for a given age would estimate the reliability at that age from the current analysis, and then 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 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 as running a single analysis can require a moderate to large amount of computer time.

We have assumed that the *k* possible allocations have been specified a priori, and our goal for resource allocation is to select the best possible allocation. However, it may be the case that we wish to solve a more general problem of finding the best allocation subject to some cost or logistical constraints. In this case, we can think of constructing a response surface which uses the results of our initial set of allocations to suggest alternate allocations which may be superior to those itemized. Once these allocations are suggested, then we can iteratively assess them, update our response surface, and find new allocations to explore. Since the allocations can be thought of as defining the proportion of our resources to spend on each type of data, a mixture response surface using a higher order Scheffe model (see Cornell, 2002) can be used to characterize how the reduction of uncertainty changes with different proportions for the various types of data.

METRIC FOR UNCERTAINTY OF RELIABLITY ESTIMATE

Determining the appropriate criterion for assessing the improvement in precision for our reliability estimate has several aspects to consider. Figure 3 shows a typical summary of the results of the analysis from Phase 1 of the data collection based on the preliminary data. The solid curves show the median of the posterior from a Bayesian analysis for system reliability across various ages. The dashed lines show the 90% (red) and 99% (blue) credible intervals for the reliability estimate. For this example, data from the various sources were available for systems of age 0 to 10 years, as noted by the vertical line at age 10.

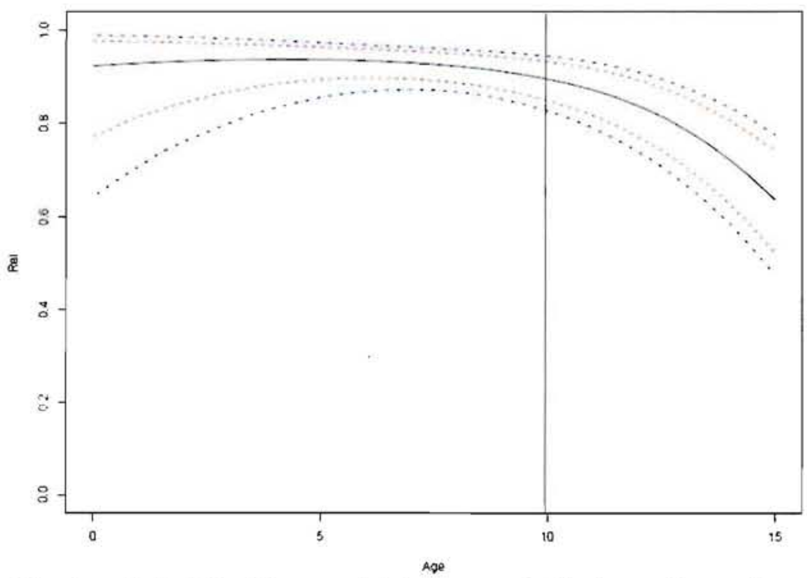

Figure 3: Plot of System Reliability. Observed data are available for systems of age 0 to 10 years, while interest lies in predicting reliability up to age 15 years.

The goal of the analysis is to be able to predict system reliability for ages 0 to 15 years. While the extrapolation of our estimates contradicts a fundamental rule of statistical analyses, policy decisions must be made about how to manage the population of systems several years in advance. Hence, our objectives for the analysis require us to focus beyond the range of observed data, and make assumptions that the pattern of aging will continue to follow similar mechanisms to those previously observed. The decision to continue this same pattern of aging is based on scientific or engineering understanding of the process.

Standard measures of precision tend to focus on ranges where data have been observed, and so some adaptation of these metrics is needed. In particular, the following choices should be considered when selecting a metric:

1. Measure of uncertainty (variance, width of credible interval, or entropy (see Lindley, 1956 or Chaloner & Verdinelli, 1995 for more details on entropy))
2. Summary of the measure to use from the multiple analyses (for example, the median value of the variances across the analyses would show a "typical" reduction of

uncertainty that could be expected if the allocation were selected, while the 10th percentile would give a lower bound on the reduction that could be expected)

3. Range of ages on which to base measure (some choices for this example might be 0-15 years or 10-15 years, depending on the ages of the available systems and where decision about their management are needed)
4. Number of ages to consider within this range (for example, if we select the age range 10-15 years in 2., then we could look at just ages 10 and 15, or ages 10, 12.5 and 15)

The choice of metric should be made based on how the results of the analysis will be used. Under certain restrictive assumptions about the nature of the data, these three metrics should all lead to the same selection of the best allocation (Wynn, 2004). However, in general, these metrics may lead to different suggested allocations, and hence a metric that is natural to the application should be chosen. Similarly, the range of ages on which to base the selection also may suggest different preferred allocations.

THE NEED FOR A DISCREPANCY MEASURE IN THE ANALYSIS

In the original and subsequent analyses, it may be beneficial to include a discrepancy term in the model that allows for different sources of data to suggest different patterns in the reliability. Consider the example of Figure 1 with the system structure suggesting that all of the components need to work correctly for the system to work. This would suggest a series system with the reliability of the system being appropriately describe with the following equation

$$rel_{sys} = rel_{Comp1} \bullet rel_{Comp2} \bullet rel_{Comp3.1} \bullet rel_{Comp3.2} \bullet rel_{Comp4}\ .$$

However, the estimates for the left-hand side and the right-hand sides may show considerable differences. This might be the result of additional engineering details existing for how the various components are connected that have not been included in the initial model, or from additional structures that have been introduced during the manufacturing process. One way that this might occur would be if the component level tests are incorrectly calibrated relative to the demands on that component during the full system test. For example, it might be that a component could pass the component level test, but the system level test could fail because the component cannot deliver the full capability required during the system test. A second way that inconsistencies could be introduced would be if there are some connectivity issues between components. For example, if all 5 of the components for a particular system were in working order, but there were problems with how the parts interacted, then we could observe a system failure. Hence before we confidently use component level test data to estimate system reliability, it would be beneficial to assess the compatibility of the different data sources. See Anderson-Cook (2009b) for more details on how to evaluate the compatibility of data from different levels of the system.

The inclusion of a discrepancy measure in the model, perhaps something of the form

$$rel_{sys} = \delta \bullet \prod rel_{Comp\ i}$$

Allows for estimation of any inconsistencies in the various levels of data. If this term is included, then the resource allocation approach would select an allocation that provides some natural balancing of the different data sources. This would occur because in order to reduce overall uncertainty, the discrepancy parameter needs to be well estimated. This requires adequate data from both the system level and component level tests.

SUMMARY AND CONCLUSIONS

Determining a next set of data to be collected and analyzed can be a valuable part of sequential experimentation. By using current understanding of the system reliability, we can leverage this knowledge into making a more informed decision about where future resources should be spent. Because the variability of a reliability estimate is a function of its estimated value, using estimates based on available data can help us spend our resources most effectively to maximally reduce the uncertainty of our system reliability estimates. By performing this evaluation of anticipated results for different allocations before the data are actually collected, a best choice of new data can be identified and the projected gains from that data quantified. This information can be valuable when justifying the benefits of additional data.

The methodology outlined in this paper has provided an overview of the approach for assessing potential new allocations. Some of the key aspects to consider when implementing this approach include the careful specification of which allocations to compare, what metric should be used to determine a best allocation, and how best to formulate a model that best captures what is known about the system structure. The implementation of the algorithm has considerable flexibility to exploit knowledge about the system. Because of the computationally intensive nature of the approach, methods for streamlining are being sought to improve the timeliness of solutions.

REFERENCES

1. Anderson-Cook, C.M. "Opportunities and Issues in Multiple Data Type Meta-Analyses" (with discussion and rejoinder) ***Quality Engineering*** 21 241-253,260-261 (2009a).
2. Anderson-Cook, C.M. "Evaluating the Series or Parallel Structure Assumption for System Reliability" **Quality Engineering** 21 88-95 (2009b).
3. Anderson-Cook, C.M., Graves, T.G., Hamada, M.S. "Resource Allocation for Reliability of a Complex System with Aging Components" ***Quality and Reliability Engineering International*** 25 481-494 (2009).
4. Anderson-Cook, C.M., Graves, T., Hamada, M., Hengartner, N., Johnson, V., Reese, C.S., Wilson, A.G. "Bayesian Stockpile Reliability Methodology for Complex Systems" ***Journal of the Military Operations Research Society*** 12 25-37 (2007).
5. Anderson-Cook, C.M., Graves, T., Hengartner, N., Klamann, R., Wiedlea, A.K., Wilson, A.G., Anderson, G., Lopez, G. "Reliability Modeling using Both System Test and Quality Assurance Data" ***Journal of the Military Operations Research Society*** 13 5-18 (2008).
6. Anderson-Cook, C.M., Huzurbazar, A.V., Klamann, R.M., Morzinski, J. "Calculating System Reliability with SRFYDO A User Guide" **Los Alamos National Laboratory Technical Report** LA-UR 08-06161 (2008).
7. Chaloner, K., Verdinelli, I. "Bayesian Experimental Design: A Review" ***Statistical Science*** 10273-304 (1995).
8. Cornell, J. ***Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*** New York: Wiley (2002).
9. Lindley, D.V. "On a Measure of the Information Provided by an Experiment" ***The Annals of Mathematical Statistics*** 27 986-1005 (1956).
10. Rausand, M., Hoyland, A. (2004) ***System Reliability Theory: Models, Statistical Methods, and Applications*** New York: Wiley (2004).
11. Saunders, S.C. ***Reliability, Life Testing, and Prediction of Service Life*** New York: Springer (2007).
12. Wilson, A.G., Graves, T.L., Hamada, M.S., Reese, C.S. "Advances in Data Combination, Analysis and Collection for System Reliability Assessment" ***Statistical Science*** 21 514-531 (2006).
13. Wynn, H.P. "Maximum Entropy Sampling and General Equivalence Theory" **Advances in Model-Oriented Design and Analysis, Contributions to Statistics, Physica**, Heidelberg 211-218 (2004).