LA-UR- 08-6837

Approved for public release; distribution is unlimited.

Title:	Reliability Models for Almost-Series and Almost-Parallel Systems
Author(s):	Todd L. Graves, CCS-6 Christine Anderson-Cook, CCS-6 Michael S. Hamada, CCS-6
Intended for:	Technometrics

Los Alamos
NATIONAL LABORATORY
EST. 1943

Form 836 (7/06)

Reliability Models for Almost-Series and Almost-Parallel Systems

Todd L. Graves

Christine Anderson-Cook

Michael S. Hamada

Statistical Sciences

Los Alamos National Laboratory

Abstract

When assessing system reliability using system, sub-system and component level data, assumptions are required about the form of the system structure in order to utilize the lower level data. We consider model forms which allow for the assessment and modeling of possible discrepancies between reliability estimates based on different levels of data. By understanding these potential conflicts between data, we can more realistically represent the true uncertainty of the estimates and gain understanding about inconsistencies which might guide further improvements to the system model. The new methodology is illustrated with several examples.

Key Words: Bayesian, discrepancy, Markov chain Monte Carlo, multi-level binomial data, parallel and series systems

1 Introduction

Realistic and representative reliability assessments of high consequence systems remain an important task, especially as the complexity of the systems increases. Examples include one-shot systems such as automotive airbags, safety systems in nuclear power plants, and missiles, for which it is desirable that they work when called upon to perform. Data at multiple levels may be available – few full system tests, more subsystem tests, and even more component tests – and analysts are asked to use these with other sources of information in the assessments. The system reliability assessment literature has mostly focused on methods for combining such multi-level data. A key aspect of such assessments is the assumed reliability structure of the system, e.g., as laid out by reliability block diagrams. This paper focuses on treating the stated structure as an assumption that may or may not be exactly correct, while using multi-level data to assess the assumption. We show how to assess system reliability when there is inconsistency between the assumed and actual system reliability structure.

While systems may have been designed to have a series or parallel structure, in practice there may be dependencies between components or connectivity issues that cause these assumed relationships to no longer be correct in practice. For example, consider a population of complex systems, such as car engines, kitchen appliances, or rockets. The engineers who design the system intend for it to have a series structure. However, this might not be the correct structure once the system is built and data are collected. (1) During production connectivity issues between components might make it possible for two components to each work separately, but not work when combined together. (2) If the performance of one component depends on receiving the correct signal or power supply from another component, we may observe apparent failures of some components even when that component is actually working. (3) The component level tests may be incorrectly calibrated relative to the demands on that component during the full system test. For example, it might be that a component will pass the component level test, but the system level test could fail because the component cannot perform adequately.

Hence before we confidently use component level test data to estimate system reliability, it

would be beneficial to estimate the compatibility of the different data sources. Understanding the reliability of a complex system is important in many applications. It allows us to predict performance and for many systems there are substantial consequences for the system not performing as intended. Most government and commercial applications have some type of assessment of reliablity included. As the complexity and cost of the system increases, it may be increasingly expensive or impractical to collect full system data as the primary means of assessing system reliability. Hence, supplementing this data with other less direct sources may be beneficial and cost-effective.

We begin by briefly reviewing the literature on system reliability assessment. There are a number of papers on Bayesian statistical assessments of system reliability, based on data from component and higher level tests, which assume series and parallel subsystems. Some important examples are Cole (1975), Mastran (1976), Mastran and Singpurwalla (1978), Natvig and Eide (1987), Martz, Waller and Fickas (1988), Martz and Waller (1990), Johnson, Graves, Hamada and Reese (2003) and Anderson-Cook, Graves, Hamada, Hengartner, Johnson, Reese and Wilson (2007). Only Johnson et al. (2003) and Anderson-Cook et al. (2007) provide a fully Bayesian solution using Markov chain Monte Carlo (MCMC). In all these papers, it is assumed that the components fail independently and that the assumed system structure as described by a reliability block diagram is correct. Lynn, Singpurwalla, and Smith (1998) present several interesting and useful types of prior dependence between the component reliabilities, including several highly reliable components, nearly identical components, and ordered component reliabilities: these ideas can be used to complement the ideas in our paper.

In this paper, our interest is in the case where a nominal structure of the system is given, but where the reliability of subsystems or the full system may deviate from the exact functional form given by the stated structure. The simplest example is a two-component series system, where nominally the system works properly if and only if both components work properly. Denote the indicator that the system works properly by X_0 (i.e., $X_0 = 1$ if it works properly) and the corresponding indicators for the two components by X_1 and X_2.

The *functional series system assumption* is

$$X_0 = X_1 X_2 \text{ (with probability 1).} \tag{1}$$

Consider a simple probability model with $E(X_1) = p_1$, $E(X_2) = p_2$, and where p_1 and p_2 are fixed constants. A further assumption is that X_1 and X_2 are independent given p_1 and p_2, and if this is true along with the series system assumption, then we have $p_0 = E(X_0) = p_1 p_2$; we call this the *probabilistic series system assumption.* See Anderson-Cook (2008) for a test of this assumption. (Note that we use the reliability block diagram definition of a series system, not the fault tree definition, which investigates failure mechanisms). We wish to allow for the possibility that p_0 is not exactly equal to $p_1 p_2$. In other words, we wish to model and understand the discrepancy between the reliabilities suggested by different data types. A classical example is of a "missing component," where a failure mode has not been characterized, in which case the system reliability p_0 is equal to $p_1 p_2 p_3$, where $1 - p_3$ is the probability of a failure from the unknown mode. However, if the two components are not required to be independent, it is also possible to have $p_0 > p_1 p_2$ as well as $p_0 < p_1 p_2$: for example, there may be an unrecognized environmental or manufacturing condition u with probability density function f_u that simultaneously affects the failure probability of both components: p_1 is by definition $\int p_1(u) f_u(u) du$, and $p_2 = \int p_2(u) f_u(u) du$, and now $p_0 = \int p_1(u) p_2(u) f_u(u) du$. We use this example to motivate defining models more generally than the missing component model, but typically we will not explicitly write down a model for dependence on u or try to impute it when it is unknown.

We focus on the case where simultaneous data (Graves et al., 2008) are not available, i.e., each observation of test data is a success or failure at exactly one component, subsystem, or full system, and the observations are independent of each other within a particular data type as well as across data types given the parameters. (An example of simultaneous data is a full system test of a two component series system, where both components are observed to work properly but where the system is also observed to fail.)

We consider modeling the discrepancy of reliability estimates based on the various types of data and information. For the previous two component example, we can estimate sys-

tem reliability using only the system data with $\hat{p}_0$, or using the component data under the assumption of a series system as $\hat{p}_1\hat{p}_2$. It is possible to use this methodology to investigate discrepancies in the data or to study the consistency of expert opinion at various levels.

1.1 Notation and definitions

In the following, we refer to an element of the system where data can be collected, such as the system, a subsystem or a component, as a *node*: denote the complete set of nodes in the system by J. The connection between nodes in the system structure are called gates, and are typically *and* or *or* gates. For example for the simple series system described earlier, we have 3 nodes: system, components 1 and 2 and one *and* gate. Denote the vector of unknown parameters by θ; this vector describes the process by which similar systems are built and tested, and from a Bayesian perspective we express uncertainty about the value of θ using a probability distribution $\pi(\theta)$. Figures 1 and 2 show a success tree representation (Stamatis, 1995) and block diagram for a two sub-system and five component system. The success tree representation in Figure 1 will be more useful in this paper. In Figure 1, each numbered oval represents a node, with the five nodes along the bottom representing the components. The AND symbol under node 2 indicates that components 5, 6, and 7 combine in series to form subsystem 2, the OR symbol under node 1 indicates that components 3 and 4 combine in parallel to form subsystem 1, and the AND symbol under node 0 shows that subsystems 1 and 2 combine in series to form full system 0.

A single system (the ith such one) built according to this process is described by parameter vector θ^i, with $\theta^i \sim F(\cdot|\theta)$. Finally, suppose that node $j \in J$ of the ith system is tested; we then denote the indicator that the test is a success by X_j^i. The systems have an assumed structure ϕ, which means that one subset of nodes $C \subset J$ are called *components*. The set $C \subset J$ satisfies two properties: first, $\int P\{X_j^i = x_j \text{ for all } j \in C|\theta\}\pi(\theta)d\theta > 0$ for each choice of $\{x_j \in \{0,1\} : j \in C\}$, so that any combination of component successes and failures is possible in principle. Second, if $j \notin C$, then if the assumed structure ϕ is correct, then $X_j^i = \phi_j(X_C^i)$, where ϕ_j is a deterministic function and where we use the notation

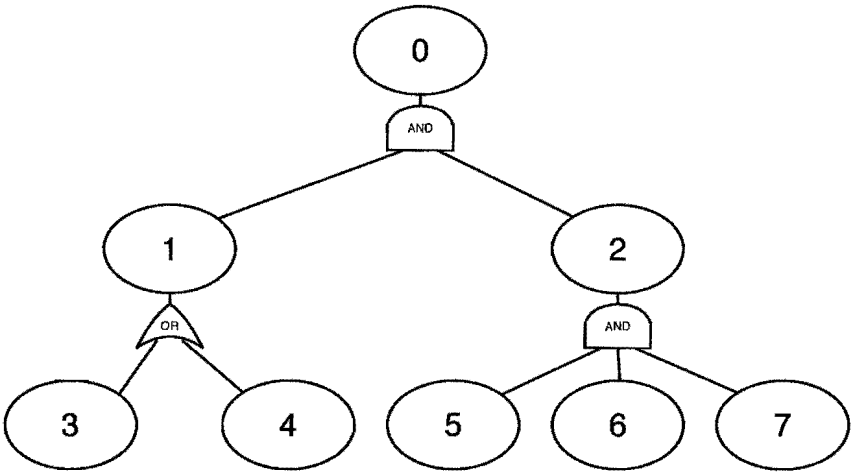

Figure 1: Success tree for system with two sub-systems and five components

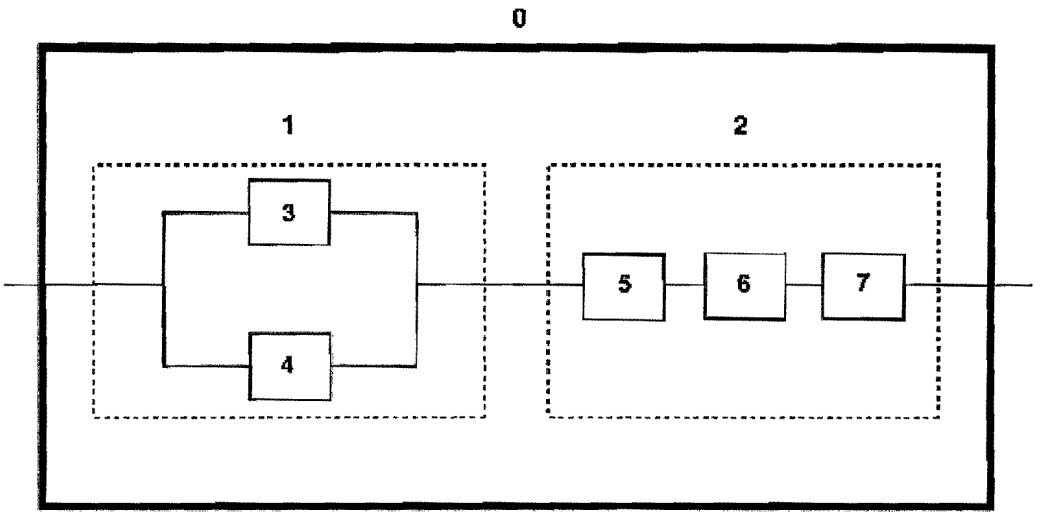

Figure 2: Simplified system reliability block diagram.

$X_C^i = \{X_j^i : j \in C\}$. For example, if node 0 is a series system consisting of components 1 and 2, then $\phi_0(x_1, x_2) = x_1 x_2$. Letting $p_j^i(\theta^i) = E(X_j^i|\theta^i)$, then $p_j(\theta) = \int p_j^i(\theta^i) dF(\theta^i|\theta)$ is the reliability of the jth node under unknown parameter θ. If X_j^i are tests of node j for $i = 1, \ldots, n_j$, then conditionally on θ, the statistic $\sum_{i=1}^{n_j} X_j^i$ has a binomial distribution with parameters n_j and $p_j(\theta)$. We will frequently write $p_j(\theta)$ simply as p_j. Lynn et al. (1998) write $P_j = \int p_j(\theta)\pi(\theta)d\theta$ and refer to it as the expected reliability of node j.

In the situations considered here, we do not separately model different instances of the system by explicitly calling out θ^i or p_j^i. In principle, one could measure covariates relevant to the ith system that would make it beneficial to have explicit models for these quantities, but we assume that such covariates are unavailable. (In Section 5, however, we consider the case where we collect data on covariates that predict component reliability, even while other unobserved covariates generate further relationships between component reliabilities.) We also restrict attention to the case where all observations are independent with X_j^i for a single j (i.e., not simultaneous data). In this case only the p_j's can be estimated from the data: we cannot in general estimate the full joint distribution of $\{X_j^i : j \in J\}$. Hence in this paper we discuss prior distributions on the p_j that attempt to quantify the notion that the assumed structure is approximately correct. If we had simultaneous observations (see Graves et al., 2008) of $\{X_j^i : j \in J_i\}$, for $J_i \subset J$, we would require more elaborate models with a Bayesian network flavor; see Wilson and Huzurbazar (2007).

In this formulation, we see that there are several possible forms of structural assumptions and approximate generalizations. Let $j \notin C$; writing $p_C = \{p_j : j \in C\}$ and similarly for P_C, three possible assumptions are:

(1) **Functional assumption:** $P\{X_j^i = \phi_j(X_C^i)|\theta\} = 1$ for π-almost all θ;

(2) **Probabilistic assumption:** $p_j(\theta) = \phi_j(p_C(\theta))$ for π-almost all θ;

(3) **Classical assumption:** $P_j = \phi_j(P_C)$.

(A statement is true for π-almost all θ if it is true for all θ in a set A such that $\pi(A) = 1$.) In the case where node 0 is a series system of the two components 1 and 2, these assumptions

are

(1) $P\{X_0^i = X_1^i X_2^i | \theta\} = 1$ for π-almost all θ;

(2) $p_0(\theta) = p_1(\theta) p_2(\theta)$ for π-almost all θ;

(3) $P_0 = P_1 P_2$.

In this paper, we will consider approximations of Assumption (2). One way that it can be satisfied is if the jth structure function is correct and the p_j^i for $j \in C$ are independent conditionally on θ, assuming standard structure functions that are multilinear in the p_C's (Barlow and Proschan, 1965). Hence, the probabilistic assumption can be approximately correct due to mild correlation between the p_j^i, or due to potential violations of the functional assumption. Our models will handle both sources of divergence from the assumption, without trying to distinguish between them since our data are not adequate for that task. The randomness in p_j^i and potential correlation between different j can be due to manufacturing variability and/or variability in test conditions, in which case we assume that our test data are representative of the population of interest. Lynn et al. (1998) present prior distributions π for θ that lead to correlations between the random variables $p_j(\theta)$ when $\theta \sim \pi$. In contrast, we consider that individual systems have randomly different component reliabilities, and the component reliabilities for a given system may be correlated, leading (for example) the average reliability of all systems to be different from what would be predicted from the average reliability of all components and exact truth of the probabilistic structure assumption.

Assumption (1) is also of interest: approximations would include the possibility that a nominal series system could work properly even if one of its components is observed to fail, or system failure when all components in the system description worked. However, with data on X_j^i for only one j for each i, we cannot appropriately examine it. Assumption (3) is rarely of interest: it may be true *a priori* for priors π under which the component reliabilities are independent, but even then it will generally be false *a posteriori*. For example, failed tests of a two-component series system will generate negative posterior correlations in the reliabilities of its components. Note that none of these three assumptions implies another.

2 Example models

In this section we introduce a number of potential models. We focus at first on the two-component series system, where the component reliabilities are p_1 and p_2, the system reliability is p_0, and the approximate probabilistic system structure assumption is that $p_0 \simeq p_1 p_2$. Each of our models features a parameter λ governing the typical discrepancy between p_0 and $p_1 p_2$. The parameter λ can be fixed by the analyst if its value is assumed known, or given a prior distribution. If λ is random, we do not expect to be able to estimate it reliably based on data from one two-level system, but the posterior distributions for the component and system reliabilities will reflect uncertainty in the accuracy of the probabilistic structure assumption.

Different structure functions or systems with more components are mostly immediate generalizations. Models for systems with more than two levels (e.g., systems, subsystems and components) can be built up from the component level by applying these ideas.

2.1 Logit process

One natural model assumes that, given the component reliabilities p_1 and p_2 and the scale parameter λ, the system reliability satisfies

$$\text{logit}(p_0) \sim N(\text{logit}(p_1 p_2), \lambda^2). \tag{2}$$

In this case we have

$$\pi(p_0|p_1, p_2, \lambda) = \frac{1}{\lambda p_0(1-p_0)\sqrt{2\pi}} \exp\left(-\frac{1}{2\lambda^2}\{\text{logit}(p_0) - \text{logit}(p_1 p_2)\}^2\right). \tag{3}$$

One can also use the probit link function instead of the logit, which should give similar results. These models are straightforward to generalize flexibly to the case where covariates affect the reliabilities of the components and hence the system. The normal distribution is not critical here; in particular, a double exponential (Laplace) structure may be used.

In this model, λ should not be chosen to be too large, ($\lambda \leq 2$ is a reasonable rough guideline) especially if there are no failures or no successes at the system level. Very large λ generates a U-shaped prior for p_0 given (p_1, p_2) and favors values of p_0 closer to 1 or 0 than may be desired.

2.2 Beta error model

Another possibility assumes that p_0 has a Beta distribution centered around the series system model $p_1 p_2$ as follows:

$$\pi(p_0|p_1, p_2, \lambda) = \frac{\Gamma(1/\lambda)}{\Gamma(p_1 p_2/\lambda)\Gamma((1-p_1 p_2)/\lambda)} p_0^{p_1 p_2/\lambda - 1}(1-p_0)^{(1-p_1 p_2)/\lambda - 1}. \quad (4)$$

In this model, λ should be constrained to be no more than one; otherwise an improper posterior distribution is a possibility. This model behaves similarly to the logit model and is not discussed further.

2.3 Extensions to more components and to parallel systems

Naturally, the generalization of the logit model to the case of a series system of d components is

$$\text{logit}(p_0) \sim N\left(\text{logit}(\prod_{j=1}^{d} p_j), \lambda^2\right).$$

In a parallel system with d components, we assume

$$\text{logit}(1-p_0) \sim N\left(\text{logit}\left\{\prod_{j=1}^{d}(1-p_j)\right\}, \lambda^2\right).$$

For a system with more than two levels, we can build a model using elements already discussed. In this case, there is potentially one λ parameter at each upper-level node. One can use fixed λs and they can be different or the same throughout the graph. Once a system has several subsystems, the possibility exists for assuming a common but unknown

λ throughout the network, and estimating this parameter. The common λ then governs the overall level of correctness of the structure function throughout the system.

2.4 Default priors for component reliabilities

If prior information is available on the component reliabilities p_i for $i \in C$, this should be used. Priors such as those given in Lynn et al. (1998) can be employed here. In the absence of such information, a number of options are available: the components can be given independent uniform priors, Beta(0.5,0.5) priors, for example. The negative log gamma distribution (Hamada et al. 2008) is also worthy of consideration; in this way independent priors on the component reliabilities yield a given reliability for the series system. We also recommend a hierarchical model for the component reliabilities, in which $p_i \sim \text{Beta}(\nu\tilde{p}, \nu(1-\tilde{p}))$ for $i \in C$, with a flat or Beta(0.5,0.5) prior for $\tilde{p}$, and, for example, a prior for ν proportional to $\nu^{-3/2}$ on $(\nu_{\min}, \infty)$. (This form of the prior for ν results from a uniform $(0, s)$ prior on $\nu^{-1/2}$, a prior recommended for standard deviation parameters by Gelman et al. (2008).) The hierarchical model for component reliabilities helps minimize the phenomenon in which the estimated reliability of a system depends on the number of components in the model of the system; see Johnson et al. (2003). In much of this paper we use the uniform $(0, 1)$ prior for component reliabilities.

3 Algorithms

To combine the data with the expert knowledge represented by the various prior distributions, we use Bayes' Theorem:

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) = \pi(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \; / \int \pi(\boldsymbol{x}|\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi}, \tag{5}$$

where $\boldsymbol{\theta}$ is the parameter vector (i.e., the node reliabilities $\boldsymbol{p_j}$ and λ), $\boldsymbol{x}$ is the data vector, $\pi(\boldsymbol{\theta})$ is the prior probability density function and $\pi(\boldsymbol{x}|\boldsymbol{\theta})$ is the data probability density function (i.e., the binomial probability mass function for binomial data) which viewed as

a function of the parameter vector given the data is known as the likelihood. The result of combining the data with expert knowledge is the posterior distribution, $\pi(\boldsymbol{\theta}|\boldsymbol{x})$. Since the 1990's, advances in Bayesian computing through Markov chain Monte Carlo (MCMC) have made it possible to sample from the posterior distribution (Gelfand and Smith (1990)). Consequently, a Bayesian analysis requires an implementation of an MCMC algorithm such as Metropolis-Hastings; see, for example, Chib and Greenberg (1995).

The models discussed in this paper do not lend themselves to routine MCMC algorithms. Consider a two component series system with reliabilities p_1, p_2 and $p_0 \simeq p_1 p_2$ and a Gibbs-sampler type algorithm that attempts to sample from conditional distributions $p_0|(p_1, p_2)$, then $p_1|(p_0, p_2)$, then $p_2|(p_0, p_1)$. If the discrepancy between p_0 and $p_1 p_2$ is small but there is nontrivial uncertainty in the near-common value of p_0 and $p_1 p_2$, steps taken by p_0 can only be very small. Since p_1 and p_2 are also taking small steps, it can take a long time to adequately explore the posterior distribution. In this case, one might prefer to remove p_0 from the model altogether, assuming that $p_0 = p_1 p_2$, but in a bigger model, it may not be obvious beforehand that this will happen and may be inconvenient to rewrite the model with fewer parameters. For that reason one needs Metropolis-Hastings steps that propose moves to all the p parameters simultaneously, while preserving the current amount of discrepancy in order that accepted moves are proposed more frequently.

The algorithm we use here cycles through several Metropolis-Hastings proposals, each of which is accepted or rejected according to the standard rule. In all of what follows, we write $\theta = (p, \lambda)$ as the unknown parameter and π for an unnormalized version of the posterior density function. In what follows we discuss our choices for T_j, using the symbol θ to represent a generic "current value" of the unknown parameter, θ' to denote a generic new proposal. A generic algorithm is a vector $(T_1, \ldots, T_N)$ of conditional probability density functions $T(\theta'|\theta)$, for some number N of proposal distributions. Let n be the number of MCMC of samples from π we wish to draw. The algorithm is as follows:

1. Initialize with $\theta = \theta_{0,0}$.

2. For $i = 1, \ldots, n$ {

3. For $j = 1, \ldots, N$ {

4. Propose a new value of θ by sampling $\theta'_{i,j} \sim T_j(\cdot|\theta_{i,j-1})$.

5. Calculate the acceptance probability

$$\alpha_{i,j} = \min\left(1, \frac{T_j(\theta_{i,j-1}|\theta'_{i,j})}{T_j(\theta'_{i,j}|\theta_{i,j-1})} \frac{\pi(\theta'_{i,j})}{\pi(\theta_{i,j-1})}\right).$$

6. Sample a $U(0,1)$ random variable $U_{i,j}$.

7. If $U_{i,j} < \alpha_{i,j}$, then let $\theta_{i,j} = \theta'_{i,j}$. Otherwise let $\theta_{i,j} = \theta_{i,j-1}$. } #end j-for-loop } #end i-for loop

Most commonly, Monte Carlo averages will be calculated using $\theta_{1,N}, \ldots, \theta_{n,N}$, with $\theta_{i,j}$ for $j < N$ discarded, although this is not required. We describe a proposal mechanism in terms of how it changes a subvector of θ, in which case it is to be understood that the remaining subvector remains fixed with probability one. When defining acceptance probabilities, we will report the ratio of T_j's and leave the ratio of π's unstated.

The first several T_j have a one-to-one correspondence with the p_j, and T_j attempts to change p_j alone, through the proposal $\operatorname{logit} p'_j \sim N(\operatorname{logit} p_j, s_j^2)$. The acceptance probability for this move is then $\frac{p'_j(1-p'_j)}{p_j(1-p_j)}$ multiplied by the ratio of posterior density values. The s_j are tuned in a burn-in period so that approximately 40% of the proposed moves are accepted; see Graves (2005). Following these moves, we propose new values for λ, according to $\log \lambda' \sim N(\log \lambda, s_\lambda^2)$, where the acceptance probability is the posterior ratio multiplied by λ'/λ.

Finally, we have a set of proposal mechanisms that attempt to move all of the p_j simultaneously. We have one of these beginning at each of the p_j's, and this p_j will generally be making the largest move, but we will then also traverse the graph in all directions, proposing moves to the other p's that compensate for the proposed move to the previously changed parameters. The first parameter changed follows a Gaussian move on the logit scale as above, which contributes a factor of $\frac{p'_j(1-p'_j)}{p_j(1-p_j)}$ to the acceptance probability.

In this way, whenever we alter a new set of parameters corresponding to a given gate, exactly one of the parameters has been changed already. For each of the four cases considered,

we describe how to propose a new value for each node reliability that has not yet been changed, and report by what factor we should multiply the acceptance probability for the move.

Step A If the gate is a series gate with d components, and the parent node 0 has been changed already, define $U = \log\{p_0'/p_0\}$, and let $(V_1, \ldots, V_d)$ be drawn from the Dirichlet distribution with parameters $(\beta/d, \ldots, \beta/d)$. Then let $p_i' = p_i \exp(UV_i)$ for $i = 1, \ldots, d$. Multiply the acceptance probability by $\prod_{i=1}^d p_i'/p_i$. In this way we preserve the amount of discrepancy as $\log(p_0'/\prod_{i=1}^d p_i') = \log(p_0/\prod_{i=1}^d p_i)$.

Step B If the gate is a parallel gate, and the parent node 0 is the node that has been already changed, then $U = \log\{(1-p_0')/(1-p_0)\}$. Draw Dirichlet V's as above, and let $p_i' = 1-(1-p_i)\exp(UV_i)$. Multiply the acceptance probability by $\prod_{i=1}^d (1-p_i')/(1-p_i)$. Again we preserve the amount of discrepancy as $\log\{(1-p_0')/\prod_{i=1}^d (1-p_i')\} = \log\{(1-p_0)/\prod_{i=1}^d (1-p_i)\}$.

Step C If the gate is a series gate, and child 1 has already been changed, let $U = \log\{p_1'/p_1\}$. Let $p_0' = p_0 \exp(UV_1)$, and for $i > 1$, let $p_i' = p_i \exp(-UV_i)$. Multiply the acceptance probability by $(p_0'/p_0)\prod_{i=2}^d p_i'/p_i$. Again this preserves $\log(p_0'/\prod_{i=1}^d p_i') = \log(p_0/\prod_{i=1}^d p_i)$.

Step D If the gate is a parallel gate, and child 1 has already been changed, let $U = \log\{(1-p_1')/(1-p_1)\}$. Let $(1-p_0') = (1-p_0)\exp(UV_1)$, and for $i > 1$, let $(1-p_i') = (1-p_i)\exp(-UV_i)$. Multiply the acceptance probability by $(1-p_0')/(1-p_0)\prod_{i=2}^d (1-p_i')/(1-p_i)$. This preserves $\log\{(1-p_0')/\prod_{i=1}^d (1-p_i')\} = \log\{(1-p_0)/\prod_{i=1}^d (1-p_i)\}$.

It is possible to propose $p_i' \notin (0,1)$, in which case the proposal algorithm is terminated with a rejected move. Moves like these continue until each of the p's in the entire graph has been proposed to be changed exactly once, at which point the acceptance/rejection step is attempted, and all those p's changed are either accepted or rejected at once.

For example, consider the system with five components, two subsystems given in Figure 1. Suppose we first change node 0: the next step is to change nodes 1 and 2 using **Step A**,

then (in either order) to change nodes 3 and 4 according to **Step B** and nodes 5-7 using **Step A**. If, on the other hand, the first node to change is node 3, our next step is to change nodes 1 and 4 using **Step D**; then we change nodes 0 and 2 according to **Step C**; finally changing nodes 5-7 again according to **Step A**.

4 Examples

As a simple example, consider the case of a two component system with flat (uniform) priors on their reliabilities, where each component is tested $n_1 = n_2 = 10$ times, resulting in nine successes. The system is also tested ten times but only $x_0 = 2$ successes were obtained. Assuming $p_0 = p_1 p_2$, the maximum likelihood estimate for both p_1 and p_2 is 0.717 and for p_0 is 0.514. Given the MLEs, the probability of two or fewer successes in ten system tests is 0.0454, while the probability of eighteen or more successful component tests in twenty trials is 0.0497, so these results are fairly unlikely for a true series system. We analyze these data for several values of (fixed) λ, then repeat the analysis with exponential prior distributions for λ with varying prior means. Figure 3 displays the results. The top plots show the results for the fixed λ analyses: for different values of $\log(\lambda)$, solid lines show the posterior medians for p_0 (left), p_1 (center; this plot also applies for p_2), and $p_0/(p_1 p_2)$ (right). The corresponding dashed lines show the 5th and 95th posterior percentiles for these quantities. The bottom plots display the same quantities for analyses with exponential prior distributions for λ, with the log of the prior mean on the $x-$axis. For the fixed λ analysis and for large λ, the results are consistent with the Beta(2,8) distribution for p_0 and with the Beta(10,2) distribution for p_1 and p_2; these beta posteriors would be the result if each node were analyzed separately, with uniform priors on p_1 and p_2, and with essentially a Beta(0,0) prior for p_0, since the large λ prior works like a U-shaped prior. Results change slowly with decreasing λ, with most rapid changes occurring for moderate λ. Results are similar when λ is random with an exponential prior distribution, with the prior mean for λ being slightly less influential than λ's value when it is fixed.

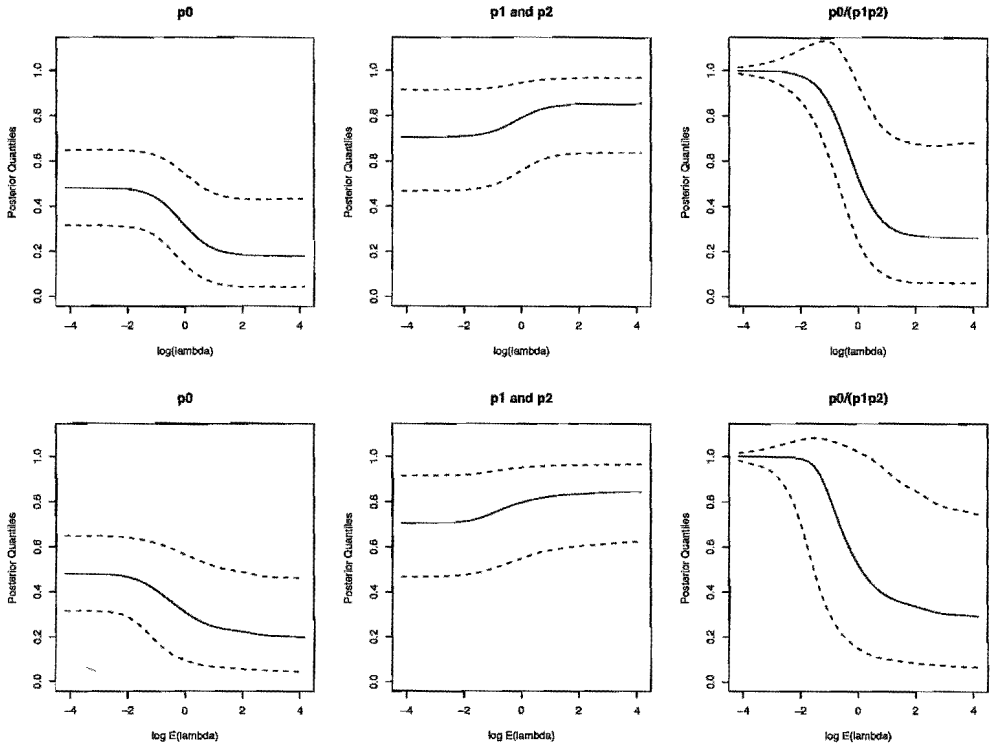

Figure 3: Results for discrepant data, fixed and random λ. Solid curves are posterior medians and dashed lines are 90% posterior intervals. Top row: posterior distributions for system reliability, component reliability, and discrepancy from series system model, for fixed λ analysis. Bottom row: same quantities for random λ analysis.

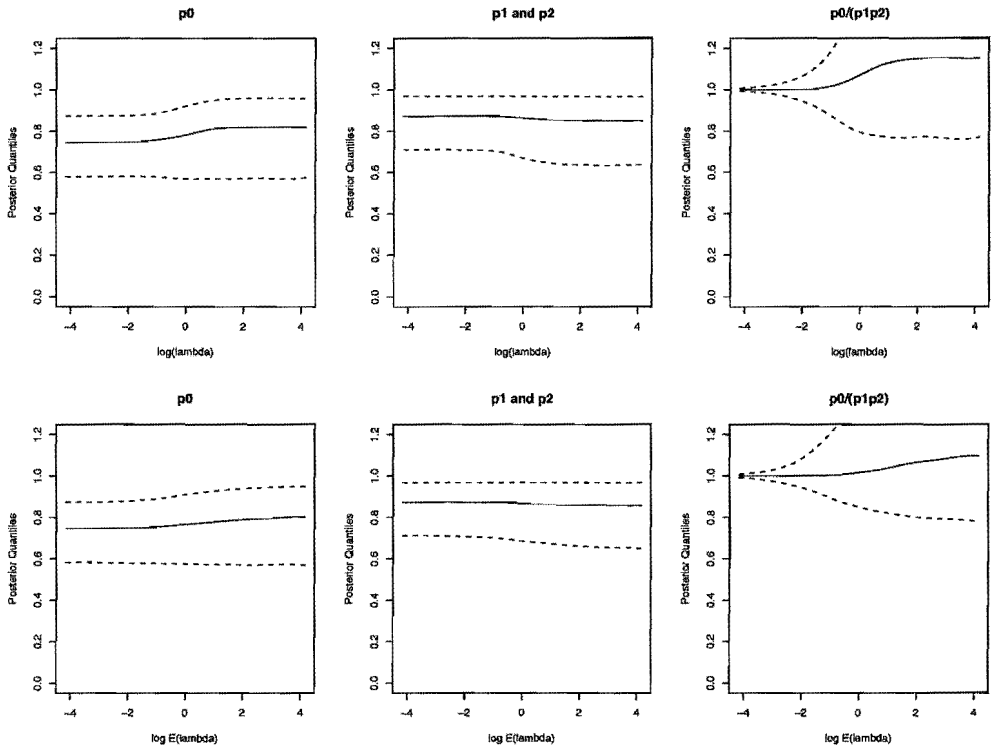

Figure 4: Results for non-discrepant data, fixed and random λ.

For another example, we assume data with minimal discrepancy: again $n_0 = n_1 = n_2 = 10$, with $x_0 = 8$, $x_1 = x_2 = 9$. With little evidence of discrepancy, we hope that allowing for the possibility of $p_0 \neq p_1 p_2$ does not greatly increase the posterior variance. As shown in Figure 4, The posterior medians for the p's are nearly independent of λ (or its prior mean), but the uncertainty in the p_i increases to some extent with increasing λ: the width of the 90% interval increases by about 30% from minimal λ (no discrepancy) to maximal λ.

In our next example, we consider a larger network with five components and two subsystems. We assume that a single λ applies to both subsystems and to the full system's relationship to the subsystems, with an exponential prior distribution with mean one. In principle one could consider a hierarchical model for three λ's, but even estimating a variance parameter based on three observations is ambitious. The assumed system structure is given in Figure 2; components 3 and 4, with reliabilities p_3 and p_4 combine in parallel to give subsystem 1, with reliability $p_1 \simeq 1 - (1 - p_3)(1 - p_4)$; components 5, 6, and 7 combine

Table 1: Data for Simplified System

Node	Data	$\tilde{p}$
0	15/20	0.8
1		0.9
2	10/10	0.9
3	34/40	0.9
4	47/50	0.9
5	3/5	0.95
6	8/8	0.95
7	16/17	0.95

in series to form subsystem 2, whose reliability satisfies $p_2 \simeq p_5 p_6 p_7$, and the system (node 0) is an approximate series system of subsystems 1 and 2. The observed data are given in Table 1. We use Beta(0.5,0.5) priors for each of the components.

We analyze these data with no discrepancy and with the $\lambda \sim$ Exponential(1) prior, and display quantile-quantile plots of posterior samples under the two models in Figure 5, with the no-discrepancy model on the y-axis. The amount of discrepancy in these data is modest: the posterior median of λ is 1.03, with a 90% interval $(0.09, 3.46)$.

We see that system reliability is more uncertain if one allows discrepancy (node 0), and uncertainty increases dramatically under discrepancy for node 1, for which there is no data. Nodes 3 and 4 have ample data so that their estimates are not affected by possible discrepancy. Node 2's reliability estimates increase when discrepancy is allowed, for three reasons: it has no failures of its own, the failures at nodes 5, 6, and 7 become less relevant, and the data at nodes 3 and 4 suggest that node 1 should be highly reliable, forcing node 2 to help explain the system failures. Nodes 5, 6, and 7 all appear at least a little more reliable when the perfect record of node 2 is not ameliorated by discrepancy.

4.1 Missile Series System Example

In this section, we apply the discrepancy model to the analysis of a published reliability assessment. Martz et al. (1988) considered the reliability of a certain air-to-air heat-seeking

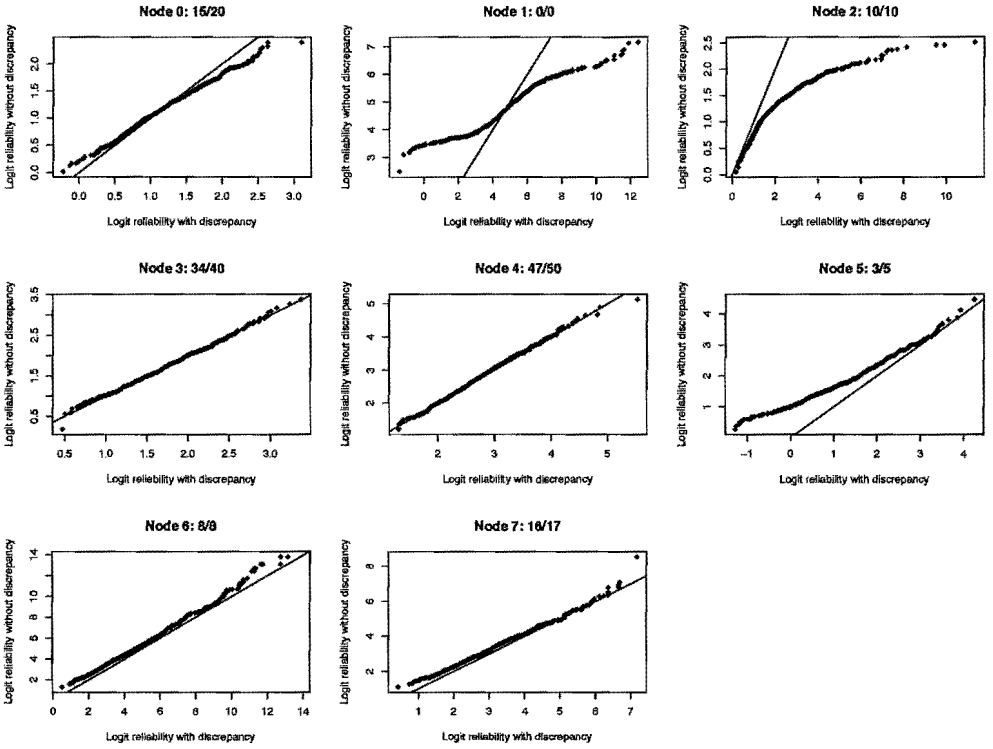

Figure 5: Comparison of results for larger system between (x-axis) a single λ for the entire model and (y-axis) no discrepancy.

missile system consisting of five subsystems in series each consisting of multiple components themselves combined in series. The data and prior information that Martz et al. (1988) used are presented in Table 2 as (successes/trials) and best guesses $\tilde{p}$ and precisions ν. Martz et al. (1988) did not provide details on how these data were obtained and how the prior information was arrived at. We treat the precisions as constants and use an Exponential(1) prior for λ.

This example highlights a modeling strategy that can be used in situations where analysts have multiple pieces of information derived from expert opinion, and wish to determine whether this information should be used at face value. First, we fit the approximate structure function model to the expert information alone, omitting the data. We then examine the results to evaluate whether there are substantial departures from the structure function. If so, we interpret this as evidence that the expert information is questionable rather than evidence that the structure function is faulty. We then remove the upper-level (subsystem and system) expert opinion from subsequent analyses; other approaches to resolving the inconsistent expert opinion can also be considered. Next, we add the data to the remaining expert opinion-based information, and fit the model again, and this time evidence of departure from the structure function is interpreted as just that. If there is no evidence that the structure assumption is violated, one can consider refitting a system reliability model with an exact structure function for final results. This is an iterative model-building approach to Bayesian analysis rather than an exact adherence to a narrowly defined version of the Bayesian paradigm.

First, we use the proposed method to assess the discrepancy of the expert opinion-based information, omitting the data from the analysis. The results are given in columns 2-4 of Table 3. For example for the system (node 0), we evaluate $\frac{p_0}{p_1 p_2 p_3 p_4 p_5}$ from the posterior. Similarly, for the subsystems (nodes 1-5), we compare their posterior reliabilities to the products of their respective component reliabilities. We see from Table 3 that the expert opinion do not constitute a consistent description of a series system: the evidence is especially strong for the system and subsystems 1 and 3. To resolve this inconsistency, here we remove the higher-level expert opinion, i.e., for the five subsystems and system. We now

Table 2: Data for Missile Series System Example

Node	Data	$\tilde{p}$	ν
0		115/265	265
1	8/8		
2	7/8		
3	191/205	257/269	269
4		55/66	66
5			
11	30/30	0.5	1
12	80/80	0.5	1
13	39/40	0.5	1
14	30/30	0.5	1
15	90/90	846/848	848
16	10/10	0.5	1
17	29/30	0.5	1
18	20/20	0.5	1
19	5/5	0.5	1
21	50/50	399/402	402
22	50/50	278/302	302
23	99/100	1098/1102	1102
24	23/25	654/690	690
25	50/50	299/301	302
26	55/55	348/352	352
31	129/130	246/250	250
32	130/130	245/250	250
33	129/130	247/250	250
34	129/130	272/276	276
35	130/130	357/360	360
36	247/250	254/257	257
37	129/130	250/252	252
38	249/250	250/252	252
39	330/330	341/352	352
41		797/802	802
42		796/802	802
43		794/802	802
44		791/802	802
45		386/402	402
51		1026/1122	1122
52		1087/1092	1092
53		1084/1092	1092

Table 3: Discrepancy Assessment Results for Missile Series System Example (0.05, 0.5, 0.95 quantiles)

	Expert Opinion At All Levels			Expert Opinion At Component Level			Component Priors With Data		
Node	0.025	0.50	0.975	0.025	0.50	0.975	0.025	0.50	0.975
0	1.18	4.06	50.39	0.00	1.00	22.39	0.31	1.01	1.58
1	1.29	11.44	554.32	0.00	0.93	16.31	0.92	1.02	1.36
2	0.26	0.88	1.15	0.18	1.00	1.17	0.83	1.00	1.12
3	1.03	1.08	1.14	0.27	1.00	1.14	0.98	1.01	1.06
4	0.79	0.90	0.98	0.35	1.00	1.08	0.71	1.00	1.07
5	0.30	0.88	1.04	0.42	1.00	1.05	0.79	1.00	1.04
λ	0.82	1.84	3.81	0.02	0.68	3.92	0.02	0.40	2.18

see from columns 5-7 of Table 3 that there is no longer any inconsistency, and λ is much smaller. In practice, once a discrepancy between higher level and component priors has been identified, engineering knowledge should be used to resolve the inconsistency. Now with a prior distribution consisting of expert opinion at the component level but not at higher levels, let us assess whether there is any discrepancy when the data are added (i.e., whether the data indicates that it is not a series system). We see from columns 8-10 of Table 3 that there is no inconsistency, as all the intervals include the value 1.

Historically, we would have used the results for the model without discrepancy. These results are provided in columns 5-7 of Table 4. We also display the results for the model with discrepancy, which show that they are more dispersed from assessing discrepancy. We also see some slight differences between the medians, especially for subsystem 1. Plots of the posterior densities for nodes 0-5 are given in Figure 6 and show slight differences. Subsystems 1 and 3 are estimated to be more reliable when we use the discrepancy model; the other subsystems get more uncertainty.

We also see the results for the components in Table 5 are less variable for the model without discrepancy. Recall that there are 9 components in subsystem 1, 6 components in subsystem 2, 9 components in subsystem 3, 5 components in subsystem 4, and 3 components in subsystem 5.

Table 4: Comparison of System and Subsystem Posteriors for Series System Example (0.05, 0.5, 0.95 quantiles) for Discrepancy and No Discrepancy Models

	Discrepancy			No Discrepancy		
Node	0.025	0.50	0.975	0.025	0.50	0.975
0	0.145	0.583	0.905	0.460	0.579	0.655
1	0.672	0.873	0.987	0.673	0.842	0.938
2	0.710	0.862	0.960	0.825	0.858	0.888
3	0.893	0.924	0.955	0.893	0.915	0.935
4	0.655	0.925	0.986	0.901	0.926	0.946
5	0.748	0.949	0.991	0.927	0.949	0.966

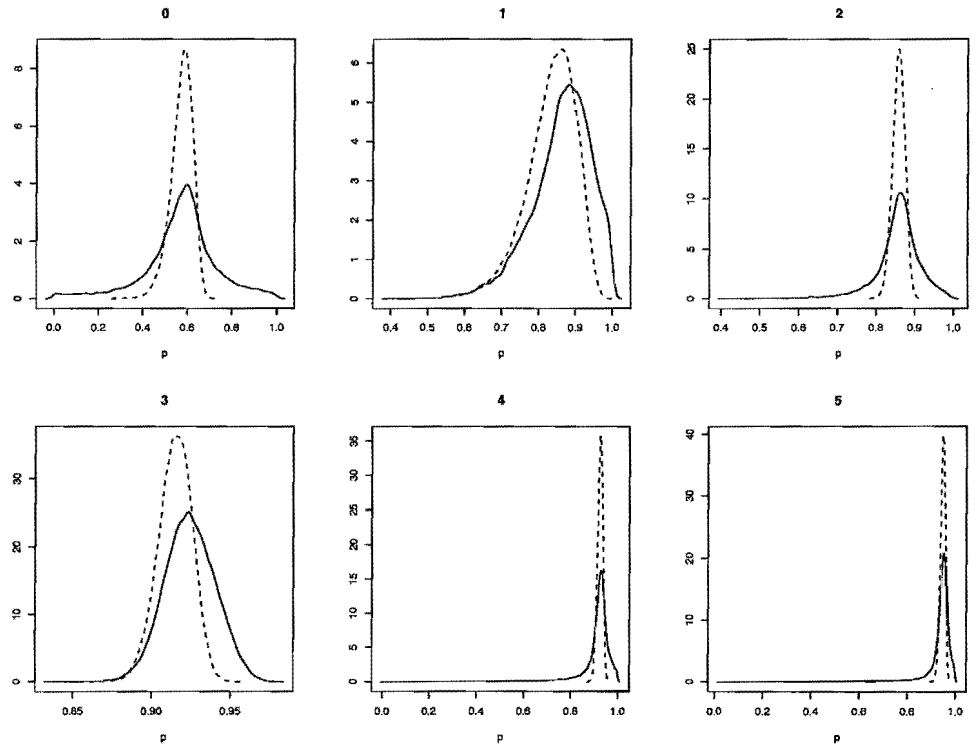

Figure 6: Posterior densities for nodes 0-5 based on models with discrepancy (solid line) and without discrepancy (dashed line).

Table 5: Comparison of Component Posteriors for Series System Example (0.05, 0.5, 0.95 quantiles) for Discrepancy and No Discrepancy Models

Node	Discrepancy			No Discrepancy		
	0.025	0.50	0.975	0.025	0.50	0.975
11	0.930	0.994	1.000	0.937	0.994	1.000
12	0.971	0.997	1.000	0.972	0.997	1.000
13	0.901	0.975	0.998	0.905	0.976	0.998
14	0.935	0.994	1.000	0.937	0.994	1.000
15	0.995	0.999	1.000	0.995	0.999	1.000
16	0.858	0.987	1.000	0.873	0.987	1.000
17	0.879	0.968	0.997	0.885	0.969	0.997
18	0.909	0.992	1.000	0.915	0.992	1.000
19	0.788	0.979	1.000	0.828	0.984	1.000
21	0.985	0.994	0.999	0.985	0.995	0.999
22	0.904	0.933	0.956	0.904	0.933	0.957
23	0.991	0.996	0.998	0.991	0.996	0.998
24	0.929	0.947	0.962	0.930	0.948	0.962
25	0.980	0.992	0.998	0.980	0.992	0.998
26	0.979	0.991	0.998	0.979	0.991	0.997
31	0.974	0.988	0.996	0.976	0.989	0.996
32	0.975	0.988	0.996	0.976	0.989	0.996
33	0.978	0.991	0.997	0.979	0.991	0.997
34	0.976	0.989	0.996	0.978	0.990	0.996
35	0.986	0.995	0.999	0.987	0.995	0.999
36	0.978	0.989	0.996	0.979	0.989	0.996
37	0.982	0.993	0.998	0.983	0.994	0.999
38	0.986	0.995	0.999	0.987	0.995	0.999
39	0.974	0.985	0.992	0.975	0.985	0.993
41	0.988	0.994	0.998	0.988	0.994	0.998
42	0.986	0.993	0.998	0.986	0.993	0.998
43	0.982	0.990	0.996	0.982	0.990	0.996
44	0.977	0.986	0.993	0.977	0.986	0.993
45	0.938	0.961	0.977	0.939	0.961	0.977
51	0.939	0.961	0.977	0.939	0.961	0.977
52	0.990	0.995	0.998	0.990	0.995	0.998
53	0.987	0.993	0.997	0.987	0.993	0.997

5 Models Incorporating Covariates

We have discussed the possibility of an unobserved covariate affecting the reliabilities of multiple components, leading to dependence between component performance and violation of the probabilistic structure assumption. We also wish to consider the case where a measured covariate is known or expected to influence component reliability, but where even after accounting for this covariate, the structure assumption is not exactly satisfied. First, consider a simple model where a single deviation from the specified structure affects the reliabilities in the same way regardless of the value of the covariate. Given the single covariate age, denoted by t, let the reliabilities of the components at age t be $p_1(t)$ and $p_2(t)$. The simplest model for an almost-series system is

$$\text{logit } p_0(t) = \text{logit}\{p_1(t)p_2(t)\} + Z, \tag{6}$$

with $Z \sim N(0, \lambda^2)$, the same random variable Z applying for all t.

A more complicated model is

$$\text{logit } p_0(t) = \text{logit}\{p_1(t)p_2(t)\} + Z(t), \tag{7}$$

where $Z(t)$ is a Gaussian process, with $E\{Z(t)\} = \zeta$ and $\text{Cov}\{Z(s), Z(t)\} = \lambda^2 \exp\{-\nu(s - t)^2\}$.

For example, consider a latent covariate model in a two-component series system, with each component aging according to a logistic regression model. For component $j = 1, 2$ and unit i, we have

$$\text{logit } p_j^i(t) = \alpha_j + \beta_j(t_i - 10) + \gamma_j u_i. \tag{8}$$

Here t_i is the age of the ith unit, which is assumed to range from 0 to 20, so that we center the variable by subtracting 10. Unit i has an observed manufacturing or testing effect given by u_i. We simulated data from $\alpha_1 = 3$, $\alpha_2 = 2.2$, $\beta_1 = -0.1$, $\beta_2 = -0.08$, and $\gamma_1 = \gamma_2 = 2.5$, and $u_i \sim N(0, 1)$. The (simulated) data are shown in Table 6. We collect 50 pass/fail tests for each component and for the system, at each of ages $(0, 5, 10, 15, 20)$. The true reliabilities at various ages are given in Table 7. The effect of u_i on both components insures positive

Table 6: Data for Covariate Example (Successes in 50 Tests)

Age	Comp 1	Comp 2	System
0	46	41	43
5	49	44	41
10	42	40	28
15	38	41	33
20	38	31	32

Table 7: True Parameter Values In Covariate Example

Age	p_1	p_2	p_0	$\text{logit}(p_0) - \text{logit}(p_1 p_2)$
0	0.905	0.837	0.798	0.237
5	0.874	0.803	0.755	0.270
10	0.835	0.763	0.704	0.303
15	0.791	0.719	0.649	0.335
20	0.745	0.678	0.597	0.373

correlation and hence higher system reliability p_0 than would be expected from $p_1 p_2$. The magnitude of this discrepancy (on the logit scale) is dependent on age, so a model that assumes a common discrepancy for each age cannot be expected to work perfectly.

In our analysis, we use $N(0, 10^2)$ priors for the α_i and β_i and a $N(0, 1)$ prior for the common value of $\text{logit}(p_0) - \text{logit}(p_1 p_2)$ (i.e., $\lambda = 1$).

The results are shown in Figure 7, with posterior medians and 90% intervals for reliability for both components and the system at ages ranging from zero to twenty-five. Points are used to display true reliabilities and are shown for ages at which data were available. The model mildly overestimates system reliability at young ages and underestimates it at older ages, since true discrepancy changes with age, but is not allowed to by the model.

We also fit the model (7), which features a different amount of discrepancy at each value of the covariate t, with discrepancies being drawn from a Gaussian process, to these data. Because the amount of noise in these data is such that it is not obvious from the data that

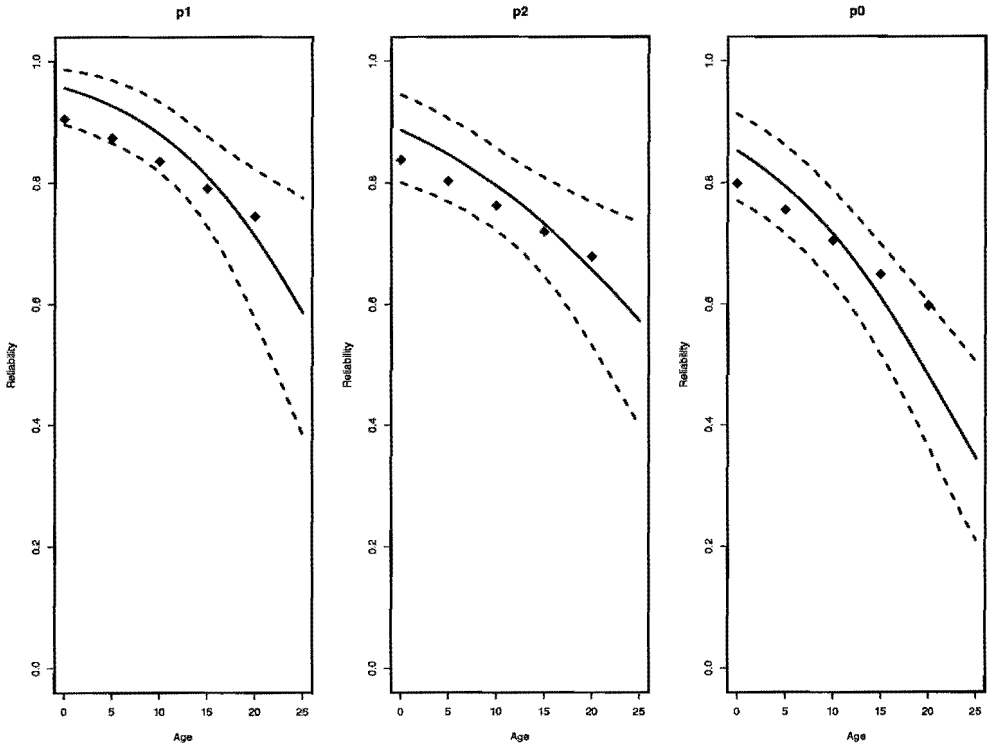

Figure 7: 90% intervals and true reliabilities for two components and system in logistic aging example with latent covariate.

different discrepancies are necessary, the extra complexity of the model does not improve the results.

6 Conclusions

By considering possible discrepancies between estimates of system reliability based on different sources of data, we are able to gain better understanding of the system structure and possible inconsistencies. The inclusion of a discrepancy term in the model can also help to give more realistic summaries of the true uncertainty in our estimates. Different possible mechanisms for modeling discrepancy are provided, as well as several examples to illustrate its use and effect on results.

The methodology described can be used both to assess any inconsistencies between expert opinion-based information included in priors for various levels of the system, and for examining whether data from the various sources indicate potential discrepancies. The remedy for different types of discrepancy should be based on engineering understanding of the system and on the type of information which is in conflict. If contradictions are revealed in the priors, a re-examination of expert opinion should be considered. If discrepancies in the data are found, then the system structure or the calibration of different data types should be examined.

Acknowledgment

The authors thank C.C. Essix for her encouragement of this work.

References

Anderson-Cook, C.M. (2008). "Evaluating the Series or Parallel Structure Assumption for System Reliability". *Quality Engineering*, (in press).

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.

Barlow, R.E. and Proschan F. (1965). *Mathematical Theory of Reliability.* New York: John Wiley and Sons.

Chib, S. and Greenberg, E. (1995). "Understanding the Metropolis-Hastings Algorithm". *The American Statistician*, 49, 327–335.

Cole, P.V.Z. (1975). "A Bayesian Reliability Assessment of Complex Systems for Binomial Sampling". *IEEE Transactions on Reliability*, R-24, 114–117.

Gelfand, A.E. and Smith, A.F.M. (1990). "Sampling-Based Approaches to Calculating Marginal Densities". *Journal of the American Statistical Association*, 85, 398–409.

Gelman, A., Jakulin, A., Pittau, M.G., and Su, Y.-S. (2008). "A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models". *Annals of Applied Statistics*, to appear.

Graves, T.L. (2003). "An Introduction to YADAS". `yadas.lanl.gov`.

Graves, T.L. (2005). "Automatic Step Size Selection in Random Walk Metropolis Algorithms". Los Alamos National Laboratory Report LA-UR-05-2359.

Graves, T.L. (2006). "The YADAS Reliability Package". Los Alamos National Laboratory Technical Report LA-UR-06-7739. Available for download at the YADAS website at `http://www.stat.lanl.gov/yadas/node1.html#download`.

Graves, T.L., Hamada, M., Klamann, R.M., Koehler, A.C., and Martz, H.F. (2008). "Using Simultaneous Higher-level and Partial Lower-level Data in Reliability Assessments". *Reliability Engineering and System Safety*, 93, 1273–1279.

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

Johnson, V., Graves, T., Hamada, M. and Reese, C.S. (2003). "A Hierarchical Model for Estimating the Reliability of Complex Systems (with discussion)". *Bayesian Statistics 7*, Bernardo, J.M., Bayarri, M.J., Berger, J., Dawid, A.P., Heckerman, D., Smith, A.F.M. and West, M. (Eds.). Oxford University Press, London, 199–213.

Lynn, N., Singpurwalla, N., and Smith, A. (1998). "Bayesian Assessment of Network Reliability". *Siam Review*, 40, 202–227.

Martz, H.F., Waller, R.A. (1990). "Bayesian Reliability Anaysis of Complex Series/Parallel Systems of Binomial Subsystems and Components". *Technometrics*, 32, 407–416.

Martz, H.F., Waller, R.A. and Fickas, E.T. (1988). "Bayesian Reliability Analysis of Series Systems of Binomial Subsystems and Componenents". *Technometrics*, 30, 143–154.

Mastran, D.V. (1976). 'Incorporating Component and System Test Data into the Same Assessment: a Bayesian Approach". *Operations Research*, 24, 491–499.

Mastran, D.V. and Singpurwalla, N.D. (1978). "A Bayesian Estimation of the Reliability of Coherent Structures". *Operations Research*, 26, 663–672.

Natvig, B. and Eide, H. (1987). "Bayesian Estimation of System Reliability". *Scandinavian Journal of Statistics*, 14, 319–327.

Stamatis, D.H. (1995). *Failure Mode and Effect Analysis: FMEA from Thoery to Execution.* New York: American Society for Quality.

Wilson, A.G. and Huzurbazar, A.V. (2007). " Bayesian Networks for Multilevel System Reliability". *Reliability Engineering and System Safety*, 92, 1413–1420.