Bayesian Evaluation of System Structure for Reliability Assessment

Lu Lu, University of South Florida

Abstract

When component data are used to assess system reliability, it is critical to validate the assumptions about the system structure for accurate prediction. This paper proposes a Bayesian approach for evaluating system structure based on measuring the multiplicative or additive discrepancy between the system and component test data under the assumed structure while quantifying uncertainty. The method is applied to multiple examples of series, parallel and complex systems for illustrating its implementation and flexibility for adapting to a variety of different structures. The impacts of sample size and the size of actual discrepancy on the power of the test and the precision of the estimated discrepancy are evaluated to provide insights on the test plan. In addition, the Bayesian probability of agreement is used to assess if the observed discrepancy is sufficiently large to be of practical importance, which offers useful information to support practical decisions for the assessment and management of system reliability.

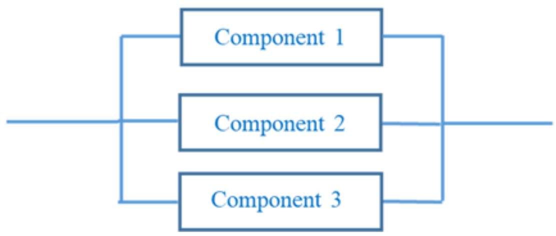
Keywords: reliability block diagram, series structure, parallel structure, complex systems, Bayesian analysis, discrepancy measure, sample size determination, power analysis, probability of agreement

Introduction

Reliability assessment and prediction are important to system design, management, and maintenance. Many complex systems are comprised of multiple components and/or subsystems, which are connected in different ways to achieve the desired functionality of the full systems. Assessing reliability through system level tests can be very costly with some full system tests being disruptive, destructive, or prohibitively expensive. Hence, it has been popular to leverage component and/or subsystem test data, which are often cheaper and easier to obtain to help improve the precision of the estimated system reliability when only limited full-system test data are available. Modeling system reliability from the component test data is referred to as a bottom-up approach (Tobias & Trindate, 2012; Chapter 10) and is dependent on assumptions about the system structure. See Martz et al. (1988, 1990), Wilson et al. (2006), and Anderson-Cook et al. (2007, 2008) for methods and examples of combining multiple sources and types of data to improve assessment of complex system reliability. The key to precise and accurate modeling and prediction of system reliability from component data is to have good fundamental understanding of the physical, structural, and operational characteristics of the complex system and how system reliability depends on the performance of individual components.

This paper develops a Bayesian method for evaluating assumptions on system structures based on examining the compatibility of the system and component test data. We focus on systems whose reliabilities do not change over time and the test data are obtained in the form of pass/fail observations for all the components and the system. In addition, the components and system tests are conducted separately and there is no information on which components fail when the system fails. For example, when performing a field test on a missile system, it is often difficult to obtain the component failure data for the test. The general strategy proposed in the paper should be easily adapted to other types of reliability data and systems with reliability changing over time by using appropriate reliability or lifetime models.

(a) A series system of $n = 5$ components



(b) A parallel system of $n = 3$ compoents

Figure 1: Reliability block diagrams for illustrating (a) a series system and (b) a parallel system.

Many common system configurations include the series and parallel structures. In a series structure, the system works if and only if all of its components function appropriately. While in a parallel configuration, redundancy is built into the system so that the system works if at least one of its components function. Reliability block diagrams (Hamada et al. 2008; Chapter 5) are often used for graphically representing the system structure. Figure 1 illustrates the reliability block diagrams for two simple systems. Figure 1(a) shows a series system consisting of $n = 5$ components. Figure 1(b) corresponds to a parallel system with $n = 3$ components. Let p_i denote the reliability of the ith component of a system, and p_s be the system reliability. Assuming all components function independently, the reliability of a series system can be calculated as

$$p_s = \prod_{i=1}^{n} p_i. \qquad [1]$$

Assuming a parallel structure suggests the system reliability can be summarized by

$$p_s = 1 - \prod_{i=1}^{n} (1 - p_i). \qquad [2]$$

Other more complicated system structures include the k-out-of-n systems (the system functions if at least k out of its n components function), standby systems (additional components are kept in the system in standby mode), failure dependent systems (the failures of some components are dependent on the failures of other components), and complex systems that are combinations of some of the above mentioned structures. Hamada et al. (2008; Chapter 5) and Tobias & Trindade (2012; Chapter 10) illustrate how to use the block diagrams, structure functions, minimal paths and minimal cuts approaches to represent the complex system reliability as functions of component reliabilities.

Since detailed knowledge of the system structure is the key to accurate estimation of system reliability utilizing component data, it is important to validate the structural assumptions to confidently use them for improved assessment. Anderson-Cook (2009) proposed a frequentist hypothesis test approach based on using the maximum likelihood estimates of the component and system reliabilities from the available data. Particularly, an empirical p-value for a proposed test statistic under the assumed relationship between the system and component reliability was calculated using a parametric bootstrap approach to evaluate the compatibility of results between the system and component level data. In this paper, we develop a Bayesian approach to directly quantify the discrepancy with associated uncertainty between the estimated system reliability and what is anticipated from the component data with the assumed structure. This is not only useful for calibrating the system reliability model from the component data, but also helpful for system structure diagnostics and failure detection.

The remaining of the paper is organized as follows. The next section describes the proposed Bayesian approach for evaluating system structure assumptions using a discrepancy measure between the test data. The method is first demonstrated with two examples with the simple system structures illustrated in Figure 1, and then followed with more advanced implementation for complex systems. The examples include discussion about how the analysis results can provide useful insights for the improved understanding of the system structure and operational conditions. The next section shows the results of a simulation study on sample size and power analysis to illustrate the impact of the amount of available data on our ability to assess system structure. The final section uses the Bayesian probability of agreement to evaluate if the discrepancy is practically meaningful in the context of engineering decision-making.

Bayesian Evaluation of System Structure

In this section, we propose a Bayesian method for quantifying the discrepancy between the system and component reliability under the assumed configuration structure. First, we illustrate the method for an example with the series structure shown in Figure 1(a). Suppose we have obtained pass/fail (binary) data for testing all five components and the full system. Let N_s and N_i represent the number of units tested at the system and the ith component, respectively. We use Y_s and Y_i to represent the corresponding observed number of passes from the N_s and N_i tests. Then the observed data can be model by

$$Y_s \sim Bin(N_s, p_s) \quad \text{and} \quad Y_i \sim Bin(N_i, p_i) \qquad [3]$$

at the system and component levels. In addition, we define metrics for describing the discrepancy between the system and component data in terms of the estimated system reliability. The discrepancy can be quantified in either a multiplicative or additive form. We define a multiplicative discrepancy parameter δ_1 of the form

$$\delta_1 = \frac{p_s}{\prod_{i=1}^{n} p_i}. \qquad [4]$$

This metric quantifies how compatible the observed results from system and component data are under the assumed structure in a multiplicative fashion. If $\delta_1 = 1$, then the system reliability is consistent with what is anticipated from component data given engineering understanding of the system configuration. If $\delta_1 < 1$, then the system reliability is worse (smaller) than the anticipated reliability from component data. This could be an indication of an additional unexpected pseudo component (which affects system performance without adding another physical component) that could cause failure of the system, or a connectivity issue between some components. It could also indicate that the full system places higher demands on certain component(s) than when it is operated or tested on its own, or the dependence of component reliabilities (e.g. the functioning of one component is dependent on the appropriate functioning of another component, or a common cause leading to simultaneous failures of multiple components). On the other hand, $\delta_1 > 1$ means system reliability is better (higher) than what is anticipated under the assumed system structure. This could be an indication of some component reliabilities not being appropriately calibrated, such that the requirement for a single component to function on its own is higher than what is needed to support the functioning of the full system.

An alternative additive discrepancy measure can be defined as

$$\delta_2 = p_s - \prod_{i=1}^{n} p_i. \qquad [5]$$

In this case, the evaluation of the system structure is based on comparing δ_2 to 0. If $\delta_2 = 0$, then the system structure is consistent with the assumptions. If $\delta_2 > 0$, then system data suggest higher reliability than the component data based on the assumed structure. This matches the scenario of $\delta_1 > 1$ and may suggest one or more of the possible deviations from the assumed structure discussed above. Also, $\delta_2 < 0$ matches $\delta_1 < 1$ which could be an indication of mis-calibration of system or component requirement or a connectivity issue (see the second complex system example). Both discrepancy measures can be effective in the evaluation of the system structure and the choice of which to use should depend on which metric offers more intuitive interpretation for practitioners in a given scenario.

In a Bayesian analysis of reliability (Hamada et al. 2008), the distributions of model parameters are initially specified based on any available historical knowledge using prior distributions, and after the analysis, inference is based on the posterior distribution of the parameters conditional on the observed data. The use of prior distributions offers an effective mechanism for incorporating available engineering knowledge or historical data on the reliability performance. When there is little prior information, one can perform a Bayesian analysis with non-informative priors to give comparable results to a frequentist approach. Anderson-Cook (2009) proposed a frequentist maximum likelihood approach for testing assumptions on system structure for series and parallel systems, which is equivalent to testing $H_0: \delta_1 = 1$ or $H_0: \delta_2 = 0$ for the two discrepancy measures defined above. This paper tests and estimates δ_1 and δ_2 using a Bayesian framework to demonstrate the flexibility and broad applicability of the approach for evaluating complex system structures.

To illustrate the methodology, we consider using the multiplicative discrepancy δ_1 for a system where little prior information is available. We choose a conjugate prior distribution (Hamada et al. 2008; page 31) for the component reliability, $p_i \sim Beta(a_i, b_i)$, with a probability density function (PDF) as

$$f(p_i|a_i, b_i) = \frac{\Gamma(a_i+b_i)}{\Gamma(a_i)\Gamma(b_i)} p_i^{a_i-1}(1-p_i)^{b_i-1}\,. \qquad [6]$$

In the above PDF, $\Gamma(\alpha) = \int_0^\infty s^{\alpha-1}e^{-s}ds\,, \alpha > 0$ is a Gamma function. The hyperparameters a_i & b_i are the shape parameters of the beta distribution, which can be interpreted as one plus the number of obtained successes and failures from historical data. Given there were no historical data, we choose $a_i = b_i = 1$ for all the components, $i \in \{1, \cdots, n\}$. A lognormal prior distribution was chosen for the discrepancy parameter δ_1 . Namely, $\log(\delta_1) \sim N(\mu_1, \sigma_1^2)$, with a PDF given in the form of

$$f(\delta_1|\mu_1, \sigma_1^2) = \frac{1}{\delta_1\sigma_1\sqrt{2\pi}} exp\left\{-\frac{1}{2\sigma_1^2}[\log(\delta_1) - \mu_1]^2\right\}. \qquad [7]$$

The hyperparameters μ_1 and σ_1 represent the mean and variance of the normal distribution of $\log(\delta_1)$. We choose $\mu_1 = 0$ and $\sigma_1 = 1$ as a diffuse prior to reflect our lack of understanding of what to expect for the amount of discrepancy. The chosen prior distribution assumes the possible value to be in the range of $(0.1, 7.1)$ with a 95% probability, which is very wide for the multiplicative discrepancy.

The posterior distribution of the model parameters given the observed data is calculated by

$$\pi(p_1, \cdots, p_n, \delta_1|y_1, \cdots, y_n, y_s) \propto f(y_1, \cdots, y_n, y_s|p_1, \cdots, p_n, \delta_1) \prod_{i=1}^n \pi(p_i)\, \pi(\delta_1) \qquad [8]$$

where $\pi(p_i)$ and $\pi(\delta_1)$ denote the PDFs of the prior distributions given in Eqn. [6] and Eqn. [7], and $f(y_1, \cdots, y_n, y_s | p_1, \cdots, p_n, \delta_1)$ is the likelihood function of the form

$$f(y_1, \cdots, y_n, y_s | p_1, \cdots, p_n, \delta_1) = f(y_s | p_1, \cdots, p_n, \delta_1) \cdot f(y_1, \cdots, y_n | p_1, \cdots, p_n)$$

$$= \binom{N_s}{y_s} [\delta_1 \prod_{i=1}^{n} p_i]^{y_s} [1 - \delta_1 \prod_{i=1}^{n} p_i]^{N_s - y_s} \cdot \prod_{i=1}^{n} \left\{ \binom{N_i}{y_i} p_i^{y_i} (1 - p_i)^{N_i - y_i} \right\}. \quad [9]$$

Note in Eqn. [9], for the series system structure in Figure 1(a), the system reliability can be written as $p_s = \delta_1 \prod_{i=1}^{n} p_i$, where δ_1 is inserted to adjust the discrepancy between the system reliability and the aggregated value from the component reliability under the series structure. The likelihood of the system data is given by $f(y_s | p_s) = \binom{N_s}{y_s} p_s{}^{y_s} [1 - p_s]^{N_s - y_s}$, which is equivalent to the first term in Eqn. [9] by substituting p_s by $\delta_1 \prod_{i=1}^{n} p_i$. Assuming independence between the components and system data, the joint likelihood in Eqn. [9] is the product of the likelihoods of the system and individual component data.

If the additive discrepancy measure δ_2 is used in the model, we recommend using a normal prior distribution $\delta_2 \sim N(\mu_2 = 0, \sigma_2^2)$, with a PDF of the form

$$f(\delta_2 | \mu_2, \sigma_2^2) = \frac{1}{\sigma_2 \sqrt{2\pi}} exp \left\{ -\frac{1}{2\sigma_2^2} [\delta_2 - \mu_2]^2 \right\} \quad [10]$$

When we are unsure what to expect for the discrepancy value, a diffuse prior should be used to allow a wide range of possible values for δ_2 such that $(\delta_2 + p_s^0) \in (0,1)$ for the anticipated $p_s^0 = \prod_{i=1}^{n} p_i$ under the assumed system structure. Also, the joint likelihood, $f(y_1, \cdots, y_n, y_s | p_1, \cdots, p_n, \delta_2)$, used to obtain the posterior distribution in Eqn. [8] can be calculated as

$$\binom{N_s}{y_s} [\delta_2 + \prod_{i=1}^{n} p_i]^{y_s} [1 - \delta_2 - \prod_{i=1}^{n} p_i]^{N_s - y_s} \cdot \prod_{i=1}^{n} \left\{ \binom{N_i}{y_i} p_i^{y_i} (1 - p_i)^{N_i - y_i} \right\} \quad [11]$$

Where the system reliability is estimated by $p_s = \delta_2 + \prod_{i=1}^{n} p_i$ to adjust for the discrepancy.

Due to the lack of a closed form expression, the posterior distribution in Eqn. [8] can be approximated using a Markov Chain Monte Carlo (MCMC) simulation (Gelman et al. 2013). We use the Gibbs sampler approach (Casella and George, 1992) to obtain samples from the approximate posterior distribution from Eqn. [8]. The calculation was implemented in R by using the rjags package to provide an interface to the JAGS library (Plummer 2003) for implementing the Bayesian data analysis with the Gibbs sampler. After obtaining the samples of the individual model parameters from the approximated posterior distribution, the analysis is then performed using the summaries from the joint or marginal posterior distributions.

To estimate the multiplicative discrepancy, δ_1, a credible interval of δ_1 can help understand in which direction and to what level the system data is different from what is expected from the component data based on the structural assumptions. For example, a 95% credible interval can be obtained by finding the 0.025 and 0.975 quantiles of the posterior marginal distribution of δ_1 from the MCMC simulations. For testing $H_0: \delta_1 = 1$ vs. $H_T: \delta_1 \neq 1$, the posterior predictive p-value (Gelman, 2013) can be calculated as

$$p_T = 2 \times \min\{\Pr(\delta_1 > 1 | y_s, y_1, \cdots, y_n), \Pr(\delta_1 < 1 | y_s, y_1, \cdots, y_n)\}. \quad [12]$$

The posterior predictive p-value measures the probability that a test quantity, which can be a function of the unknown parameters as well as data, evaluated over the draws from the posterior distribution is more extreme than what is anticipated from the observed data. For a two-sided test, this Bayesian p-value should be twice the smallest one-sided p-values. If there is a priori interest in testing $H_U: \delta_1 > 1$ (e.g. it was suspected the demand on a component might be lower for supporting system functionality) or $H_L: \delta_1 < 1$ (e.g. it was suspected that the failure of some components are positively correlated), then the corresponding posterior predictive p-value will be calculated as $p_U = \Pr(\delta_1 > 1 | y_s, y_1, \cdots, y_n)$ or $p_L = \Pr(\delta_1 < 1 | y_s, y_1, \cdots, y_n)$, respectively.

Table 1: The component and system test data (Anderson-Cook, 2009) for the five-component series system shown in Figure 1(a) and the posterior estimates of the system and component reliability with quantified uncertainty based on the test data.

Data Type	Sample Size (N_s or N_i)	Number of Successes (Y_s or Y_i)	Posterior Mean of p_s or p_i (95% Credible Interval)
System	17	25	0.657 (0.470, 0.821)
Component 1	111	120	0.918 (0.863, 0.960)
Component 2	96	100	0.951 (0.902, 0.984)
Component 3	76	80	0.939 (0.877, 0.980)
Component 4	98	100	0.971 (0.931, 0.994)
Component 5	79	80	0.976 (0.933, 0.997)

To illustrate the Bayesian method, we reevaluate the five-component series system example from Anderson-Cook (2009) (also shown in Figure 1(a)). The available data from both component and full system tests are summarized in Table 1. To obtain the MCMC samples from the joint posterior distribution in [8] approximated by the Gibbs sampler, three MCMC chains were run with the first 5000 iterations used as burn-in, after which a total of 30,000 MCMC samples were drawn. The samples are thinned by taking every 5th observation to reduce autocorrelation between adjacent samples. Diagnostic plots such as the trace and density plots shown in Figure 2 can help examine if the MCMC samples have converged and whether stable posterior distributions of the model parameters are available.

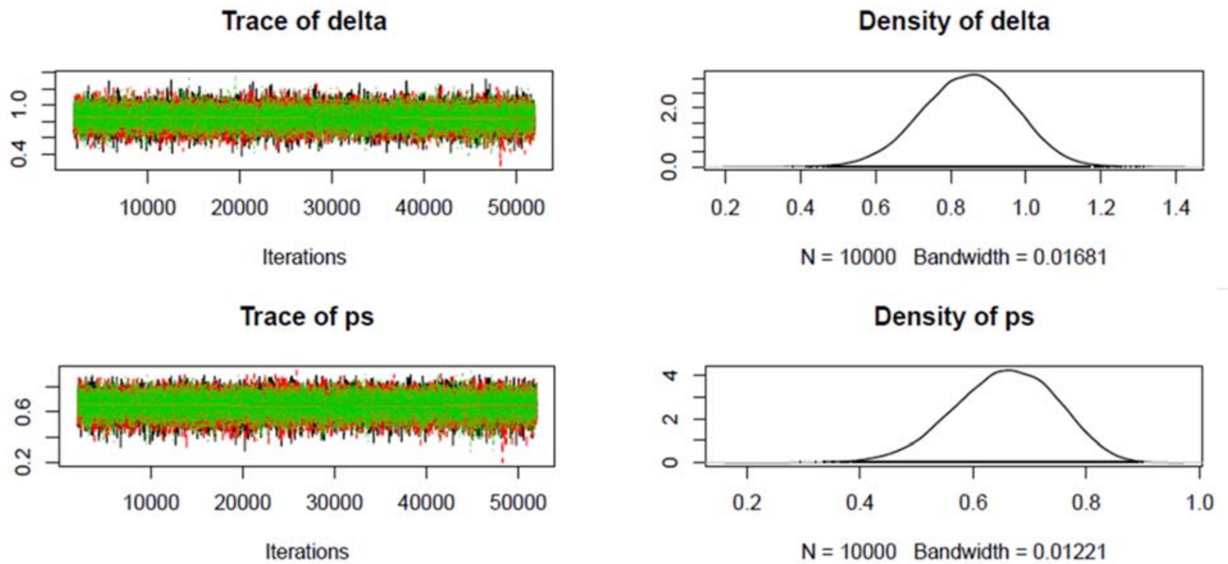


Figure 2: Plots of the trace of MCMC samples from the approximate joint posterior distribution given in Eqn. [8] and the density plots of the posterior marginal distributions of δ_1 and p_s.

The last column of Table 1 shows the Bayesian estimates (posterior mean) of the component and system reliabilities with their associated uncertainty summarized by 95% credible intervals. The posterior mean of the marginal distribution of δ_1 is 0.85 with the 95% credible interval given by (0.60, 1.08). Hence, it is estimated that there is 0.95 probability that the discrepancy δ_1 is between 0.60 and 1.08. Since the 95% credible interval includes 1, there is no statistically significant evidence indicating that the system data are inconsistent with the component data under the assumed system structure. A similar conclusion can be drawn from the Bayesian p-value approach for testing $H_0: \delta_1 = 1$ vs. $H_T: \delta_1 \neq 1$. The two-sided posterior predictive p-value is calculated as $p_T = 0.22$, which is large enough to not claim any significant discrepancy. The Bayesian analysis here offers the same conclusion as the frequentist approach from Anderson-Cook (2009). An advantage of the Bayesian approach is that the estimated discrepancy measure with quantified uncertainty can be used directly for calibrating system reliability estimated from future component test data.

Next, we consider the example of a parallel system with three components, which was also explored in Anderson-Cook (2009). The reliability block diagram of this parallel system is shown in Figure 1(b) and the component and system test data are summarized in Table 2. Given the assumed structure function from Eqn. [2], the discrepancy measures for this parallel system are defined as

$$\delta_1 = p_s / \{1 - \prod_{i=1}^{n}(1 - p_i)\} \quad [13]$$

for the multiplicative discrepancy and

$$\delta_2 = p_s - \{1 - \prod_{i=1}^{n}(1 - p_i)\} \quad [14]$$

for the additive discrepancy, respectively. The same diffuse prior distributions from Eqn.'s [6], [7], and [10] are used for the component reliabilities and the additive and multiplicative discrepancy measures. The joint likelihood of the system and component test data based on the additive discrepancy, δ_1, for this parallel system is given by

$$f(y_1, \cdots, y_n, y_s | p_1, \cdots, p_n, \delta_1) = \binom{N_s}{y_s} \{\delta_1 [1 - \prod_{i=1}^{n}(1 - p_i)]\}^{y_s} \{1 - \delta_1 [1 - \prod_{i=1}^{n}(1 - p_i)]\}^{N_s - y_s} \cdot \prod_{i=1}^{n} \left\{ \binom{N_i}{y_i} p_i^{y_i} (1 - p_i)^{N_i - y_i} \right\}. \quad [15]$$

If the additive discrepancy, δ_2, is used, then we can use $\delta_2 + [1 - \prod_{i=1}^{n}(1 - p_i)]$ to replace $\delta_1 [1 - \prod_{i=1}^{n}(1 - p_i)]$ in Eqn. [15] for quantifying the system reliability.

Table 2: The component and system test data (Anderson-Cook, 2009) for the three-component parallel system shown in Figure 1(b) and the posterior estimates of the system and component reliability with quantified uncertainty based on the test data.

Data Type	Sample Size (N_s or N_i)	Number of Successes (Y_s or Y_i)	Posterior Mean of p_s or p_i (95% Credible Interval)
System	40	33	0.806 (0.675, 0.909)
Component 1	200	136	0.678 (0.612, 0.742)
Component 2	80	54	0.671 (0.565, 0.768)
Component 3	40	31	0.516 (0.394, 0.638)

After running the MCMC simulation and summarizing the results following the same method described for the series system, the posterior mean and credible intervals of the system and component reliabilities are summarized in the last column of Table 2. The posterior mean of the multiplicative discrepancy, δ_1, is 0.85 with the 95% credible interval as (0.71, 0.96). Since the 95% credible interval excludes 1, there is a significant discrepancy between the system and component data under the assumed structure. The two-sided posterior predictive p-value for testing $H_0: \delta_1 = 1$ is $p_T = 0.003$. The very small p-value indicates a strong inconsistency between the different types of test data under the assumed system structure. Similar conclusions can be drawn with the additive discrepancy, δ_2, with the estimated 95% credible interval as (-0.27,-0.03) and the two-sided posterior predictive p-value as 0.004.

Examples of Complex Systems

An advantage of the proposed Bayesian method is that it is flexible and easier to adapt for evaluating different structures of more complex systems than the frequentist approach. Figure 3(a) illustrates the block diagram of a complex system, referred to as complex system I, which consists of six components A, B, C, D, E, and F. Based on the assumed system configuration, two pairs of components B & C and D & E are connected in parallel to form two subsystems, which are then connected in series with components A and F to form the full system. The reliability function for this complex system structure is given by

$$\begin{aligned} p_s &= p_A \cdot \{1-(1-p_B)(1-p_C)\} \cdot \{1-(1-p_D)(1-p_E)\} \cdot p_F \\ &= p_A[p_B + p_C - p_B p_C][p_D + p_E - p_D p_E]p_F \end{aligned} \qquad [16]$$

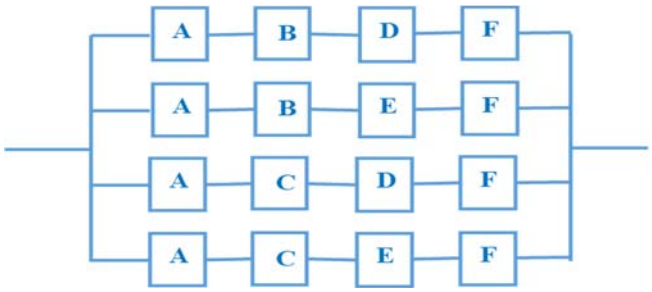
In Eqn. [16], $1-(1-p_B)(1-p_C)$ and $1-(1-p_D)(1-p_E)$ are the reliabilities of the two parallel subsystems, which are multiplied with the reliabilities of components A and F to calculate the system reliability under the series structure assumption.

The same function from Eqn. [16] can be also calculated using the minimal paths or the minimal cuts approach (Tobias & Trindade, 2012; Chapter 10). For example, the equivalent system to complex system I based on all the minimal paths is shown in Figure 3(b). A minimal path is one where all the involved components are critical to the success of the system operation. If any of the components fail in a minimal path, then the path will fail. The equivalent system has four blocks of components, i.e. the minimal paths, connected in parallel. The system will function if at least one of the minimal path functions appropriately. However, the system will fail if all the minimal paths fail simultaneously. The components within each block, i.e. a minimal path, are connected in series. Hence, if any of the involved components fail, the path will fail. With this formulation, the system reliability can be calculated as

$$p_s = 1 - \sum_{i=1}^{4}(1-p_{bi}), \qquad [17]$$

where $p_{bi} = \prod_{j \in \text{Path } i} p_j$ is the reliability of the ith block of components in the minimal path i. It is easy to prove Eqn. [16] and Eqn. [17] are the same by expanding both equations and reorganizing the terms to their simplest expressions.

(a) The complex system I block diagram



(b) An equivalent system based on all the minimal paths

Figure 3: (a) Complex system I with six components A-F, where the pairs of B & C and D & E are connected in parallel and then connected with A and F in series; (b) Equivalent system based on all the minimal paths, which are series blocks of components, connected in parallel to form the full system.

To illustrate the methodology for evaluating the system structure based on different sets of data, synthetic system and component test data are generated from a hypothetical system comprised of the complex system I plus an additional pseudo component connected in series. The data are summarized in Table 3. Assuming the complex system structure shown in Figure 3(a) and the reliability structure function in Eqn. [16], the multiplicative and additive discrepancy measures can be defined as

$$\delta_1 = p_s / \{p_A[p_B + p_C - p_B p_C][p_D + p_E - p_D p_E]p_F\} \qquad [18]$$

and

$$\delta_2 = p_s - \{p_A[p_B + p_C - p_B p_C][p_D + p_E - p_D p_E]p_F\}. \qquad [19]$$

The same diffuse prior distributions as in previous examples are used for the component reliabilities and the discrepancy measures. The joint likelihood based on the multiplicative discrepancy in Eqn. [18] is given by

$$f(y_A, \cdots, y_F, y_s | p_A, \cdots, p_F, \delta_1) = \binom{N_s}{y_s} \{\delta_1[p_A(p_B + p_C - p_B p_C)(p_D + p_E - p_D p_E)p_F]\}^{y_s} \{1 - \delta_1[p_A(p_B + p_C - p_B p_C)(p_D + p_E - p_D p_E)p_F]\}^{N_s - y_s} \cdot \prod_{i \in \{A, \cdots, F\}} \left\{ \binom{N_i}{y_i} p_i^{y_i} (1 - p_i)^{N_i - y_i} \right\}. \qquad [20]$$

Similarly, when δ_2 is used, $\delta_1[p_A(p_B + p_C - p_B p_C)(p_D + p_E - p_D p_E)p_F]$ in Eqn. [20] is replaced by $\delta_2 + [p_A(p_B + p_C - p_B p_C)(p_D + p_E - p_D p_E)p_F]$ for quantifying the system reliability.

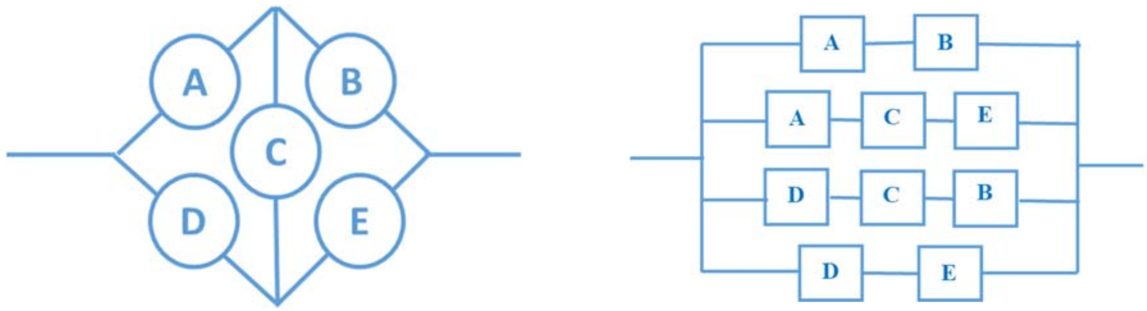
Three chains of MCMC simulations using Gibbs sampler were run with the first 5000 samples used as burn-in and then 10,000 samples were taken afterward from each chain and thinned by every 5$^{\text{th}}$ sample to reduce autocorrelation. The posterior mean of the multiplicative discrepancy, δ_1, summarized over the 30,000 MCMC samples is 0.75. The 95% credible interval is (0.61, 0.92) which excludes 1, and hence a conclusion based on these results is that the test data are not consistent with the assumed system structure. The two-sided posterior predictive p-value is $p_T = 0.005$, which also suggests a strong discrepancy. If the additive discrepancy, δ_2, is used, similar conclusions can be drawn, with the 95% credible interval for δ_2, being estimated as (-0.33, -0.06) and the two-sided posterior predictive p-value of 0.006. The analysis indicates that the system reliability is substantially lower than what is expected from the assumed configuration. This could be an indication of a missing component in the assumed structure, which matches the scenario used to generate the synthetic test data.

Table 3: The component test data for complex system I shown in Figure 3(a) and the system data simulated from a hypothetical system including complex system I and a pseudo component connected in series.

Data Type	Sample Size (N_s or N_i)	Number of Successes (Y_s or Y_i)	Posterior Mean of p_s or p_i (95% Credible Interval)
System	90	53	0.584 (0.483, 0.682)
A	60	56	0.919 (0.838, 0.973)
B	130	115	0.879 (0.818, 0.929)
C	110	92	0.830 (0.754, 0.893)
D	130	113	0.864 (0.801, 0.916)
E	90	76	0.837 (0.755, 0.905)
F	60	54	0.887 (0.799, 0.952)

Next, we consider a different example with complex system II, which has a bridge structure of five components, A, B, C, D, and E, and has been evaluated in Tobias & Trindade (2012; p. 357). Figure 4(a) shows the block diagram for the bridge structure. By using the minimal path approach, Figure 4(b) shows the equivalent parallel system with all the minimal paths. Based on these diagrams, the reliability function for this system structure can be expressed as

$$p_S = 1 - (1 - p_A p_B)(1 - p_A p_C p_E)(1 - p_D p_C p_B)(1 - p_D p_E). \qquad [21]$$



(a) Bridge structure block diagram (b) Equivalent structure

Figure 4: (a) Complex system II with 5 components A-E in a bridge structure; (b) The equivalent structure based on the minimal paths.

To illustrate the method, we generate two sets of synthetic test data with different deviations from the assumed system structure. The first scenario assumes that component C is not appropriately connected in the real system, in which case the actual system structure, referred to as complex system II-a, simplifies as what is shown in Figure 5(a). The data generated from this scenario are summarized in Table 4. By using the multiplicative discrepancy measure

$$\delta_1 = p_s/\{1-(1-p_Ap_B)(1-p_Ap_Cp_E)(1-p_Dp_Cp_B)(1-p_Dp_E)\}, \qquad [22]$$

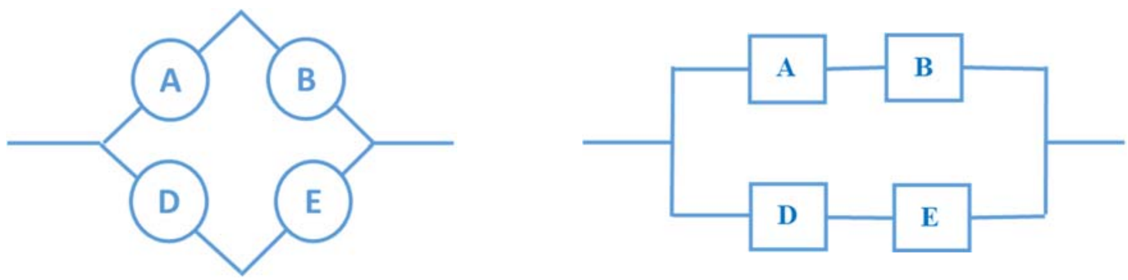
with the Lognormal prior distribution from Eqn. [7], the likelihood function is given by

$$f(y_A,\cdots,y_E,y_s|p_A,\cdots,p_E,\delta_1) = \binom{N_s}{y_s}\{\delta_1[1-(1-p_Ap_B)(1-p_Ap_Cp_E)(1-p_Dp_Cp_B)(1-p_Dp_E)]\}^{y_s}\{1-\delta_1[1-(1-p_Ap_B)(1-p_Ap_Cp_E)(1-p_Dp_Cp_B)(1-p_Dp_E)]\}^{N_s-y_s}\cdot \prod_{i\in\{A,\cdots,E\}}\left\{\binom{N_i}{y_i}p_i^{y_i}(1-p_i)^{N_i-y_i}\right\},$$

Three chains of MCMC simulations were run with 30,000 samples selected for approximating the posterior distribution. The estimated posterior mean of δ_1 is 0.89, the 95% credible interval is (0.78, 0.97), and the two-sided posterior predictive p-value is 0.007. This indicates the system test data provide an estimate of system reliability that is significantly below what is expected from using the component data with the bridge structure shown in Figure 4(a). To understand these results, we examine the structure of the actual complex system II-a used to generate the example data. As shown in Figure 5(a), when component C is not appropriately connected, both the paths through C including A-C-E and C-D-B fail. Hence, the actual system includes only two minimal paths, connected in parallel as shown in Figure 5(b). Based on this structure, we have the actual reliability function given by

$$p_s = 1-(1-p_Ap_B)(1-p_Dp_E) \qquad [23]$$

Note the system reliability in Eqn. [23] must be smaller than the anticipated system reliability from the assumed bridge structure in Eqn. [21]. This is because the original complex system II (see Figure 4(b)) has two extra redundant subsystems (minimal paths A-C-E and C-D-B) compared to complex system II-a (see Figure 5(b)), and hence has higher overall system reliability. This matches our conclusion from evaluating the system and component test data. Hence, the proposed method for estimating reliability provides an effective way to note a potential connectivity issue with the real complex system.



(a) Block diagram for complex system II-a (b) Equivalent structure based on minimum paths

Figure 5: (a) The actual complex system II-a used to generate the synthetic data in Table 4 with component C in Figure 4(a) disconnected; (b) Equivalent structure based on the minimal paths.

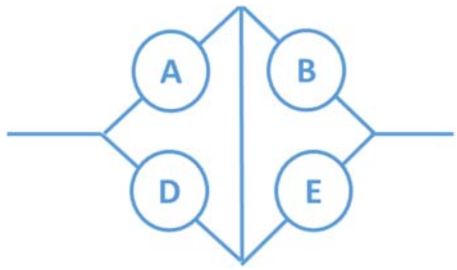
Table 4: The synthetic component and system test data for the actual complex system II-a shown in Figure 5(a) with component C in Figure 4(a) disconnected.

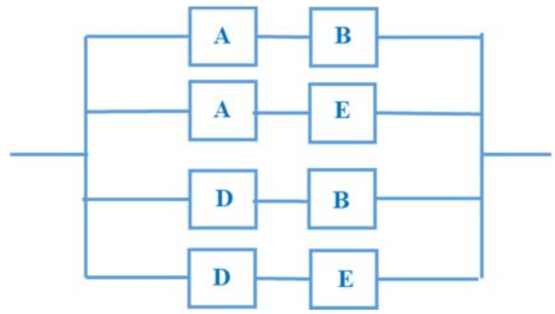
Data Type	Sample Size (N_s or N_i)	Number of Successes (Y_s or Y_i)	Posterior Mean of p_s or p_i (95% Credible Interval)
System	70	59	0.831 (0.737, 0.908)
A	100	70	0.696 (0.603, 0.781)
B	100	84	0.833 (0.756, 0.899)
C	160	113	0.704 (0.632, 0.771)
D	120	92	0.762 (0.683, 0.833)
E	160	123	0.766 (0.698, 0.828)

The second set of data was generated from a scenario assuming component C in complex system II shown in Figure 4(a) has a by-pass. In this case, component C should be replaced by a line without connecting to a device as shown in Figure 6(a) and the actual system, in this case, is referred to as complex system II-b. Based on the minimal path approach, complex system II-b is equivalent to a parallel system of all 4 minimal paths as shown in Figure 6(b), whose reliability function is given by

$$p_S = 1 - (1 - p_A p_B)(1 - p_A p_E)(1 - p_D p_B)(1 - p_D p_E). \qquad [24]$$

The data generated from this scenario is shown in Table 5, with the same component test data as in Table 4 for complex system II-a.





(a) Block diagram for complex system II-b (b) Equivalent structure based on minimum paths

Figure 6: (a) The actual complex system II-b which was used to generate the data in Table 5 with component C in Figure 4(a) having a by-pass; (b) Equivalent structure based on the minimal paths.

Table 5: The synthetic component and system test data for the actual complex system II-b shown in Figure 6(a) with component C in Figure 4(a) having a by-pass.

Data Type	Sample Size (N_s or N_i)	Number of Successes (Y_s or Y_i)	Posterior Mean of p_s or p_i (95% Credible Interval)
System	150	140	0.974 (0.943, 0.993)
A	100	70	0.696 (0.603, 0.781)
B	100	84	0.833 (0.756, 0.899)
C	160	113	0.704 (0.632, 0.771)
D	120	92	0.762 (0.683, 0.833)
E	160	123	0.766 (0.698, 0.828)

To analyze the data, the multiplicative discrepancy, δ_1, in Eqn. [22] is used, along with the same Bayesian models and prior distributions for analyzing the first set of data in Table 4 , and the same setting of the MCMC simulations. The estimated posterior mean of δ_1 for the second set of data from Table 5 is 1.044. The 95% credible interval is (1.003, 1.086), and the two-sided posterior predictive p-value is 0.037. This indicates the system reliability is above what is expected from the assumed bridge structure shown in Figure 4(a). This matches with our expectation as the resulting system reliability in Eqn. [24] must be higher than the assumed system reliability in Eqn. [21]. This can be observed from comparing the two systems shown in Figures 4(b) and 6(b). Two minimal paths in Figure 6(b) have fewer components (component C is removed) connected in a series structure than Figure 4(a), and hence have higher subsystem (series block) reliabilities and consequently higher system reliability for complex system II-b. If the additive discrepancy is used, the same conclusion is drawn with the estimated posterior mean of δ_2 as 0.041, and the 95% credible interval as (0.003, 0.078). For either choice of the discrepancy measure, the estimated discrepancy not only allows us to detect the inconsistencies between the test data and the assumed system structure, but also is useful for indicating how they are different and diagnosing the direction and possible source of the deviation.

Sample Size Determination

The amount of data available has an impact on the ability to assess the accuracy of the estimated system reliability based on its system diagram. In this section, a simulation study that was conducted is described to explore how sample size influences the power of the test and the precision of the estimated discrepancy measure for the proposed method. Consider the series system in Figure 1(a) as an example. Assume all five components have the same reliability at $p_i = 0.921$ and the same amount of observed test data at $N_i = 100$ for $i = 1, \cdots, 5$. Since the component data are generally cheaper and easier to obtain than the full system data, we assume there are more component test data available than the full system test data, and explore the impact of different sample sizes for the full system data that are less than or equal to component data size. Also, the size of the discrepancy between the system and component data has an impact on our capability to detect the inconsistencies. Hence, we explore multiple scenarios with full system testing sample size, $N_s \in [20,100]$ and the multiplicative discrepancy, $\delta_1 \in [0.5,1.5]$. Particularly, we examine the cases with N_s as multiples of 10 and δ_1 as multiple of 0.1 within the specified range, For each scenario, we generate $M = 1000$ data sets of $(Y_1, \cdots, Y_5, Y_s)$ for the component and system test data based on Eqn. [3] and the true reliability relationship, $p_s = \delta_1 \prod_{i=1}^{5} p_i$. Then the Bayesian analysis described above was conducted for each dataset with the posterior mean, 95% credible interval bounds, and the two-sided posterior predictive p-values for testing H_0: $\delta_1 = 1$ recorded for each case. By choosing $p < 0.05$ as the decision rule for rejecting H_0, the power of the test when $\delta_1 \neq 1$ and the size of the test when $\delta_1 = 1$ were calculated as the fraction of rejections out of the total $M = 1000$ simulations.

Figures 7(a) shows the power of the hypothesis test of H_0: $\delta_1 = 1$ changing as a function of the true discrepancy $\delta_1 \in [0.5,1.5]$ when the system test sample size is set at $N_s = 100$. We can see that when $\delta_1 = 1$, the size of the test (i.e. the probability of falsely rejecting H_0) is 0.056, which is pretty close to the chosen threshold of the p-value for rejecting H_0 at $\alpha = 0.05$. When $\delta_1 \neq 1$, the power is low when δ_1 is close to 1 and increases as the system model becomes increasingly less accurate with δ_1 moving farther away from 1. At $\alpha = 0.05$, the power for 100 system tests is around 0.4 for $\delta_1 = 0.8$, and increases to 0.8 for $\delta_1 = 0.7$, and is at least 0.9 for $\delta_1 < 0.65$. When the system reliability is higher than what is expected based on the component data, the power at $\delta_1 = 1.2$ is around 0.75 and is above 0.9 when $\delta_1 > 1.26$. An alternate way of looking at the results is to consider different amounts of full system data for an expected discrepancy. Figure 7(b) shows how the power changes with the system sample size $N_s \in [20,100]$ at a fixed discrepancy value of δ_1=0.6. We can see the power increases from 0.5 to about 0.95 as N_s increases from 20 to 100. The minimum size of the system test data to detect $\delta_1 = 0.6$ with at least 0.9 power is around 70.

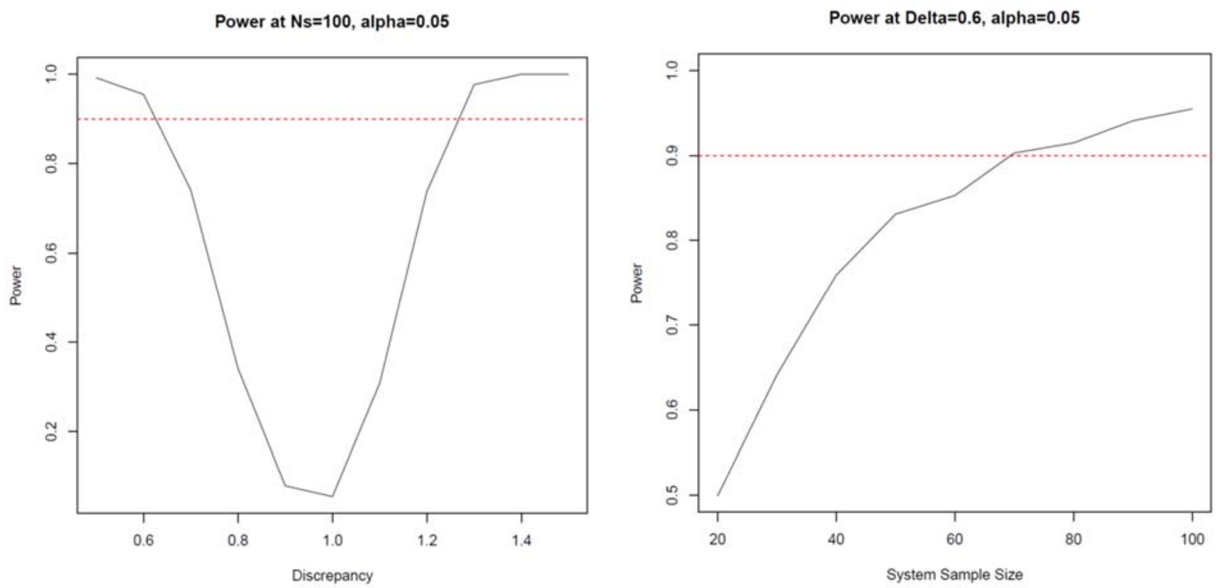
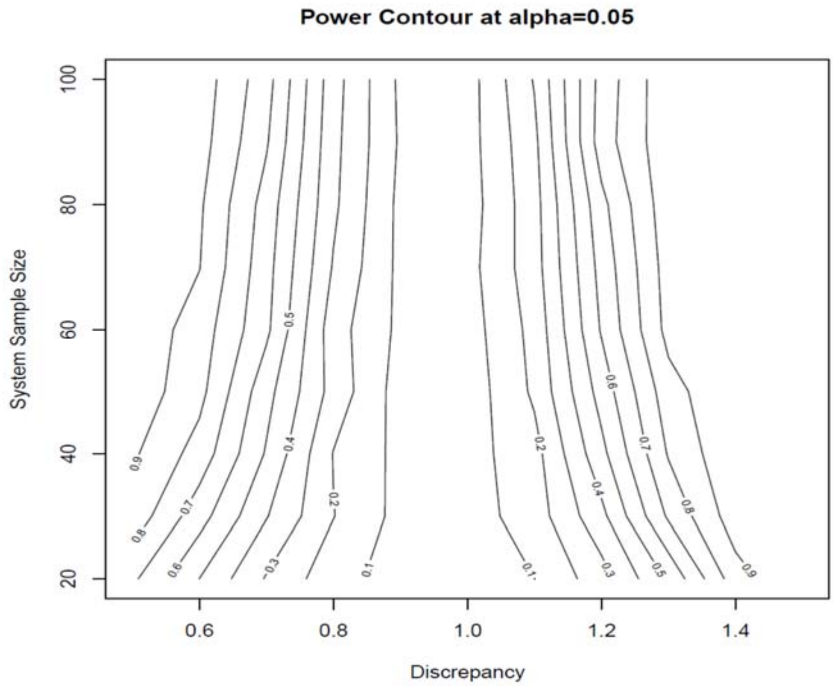
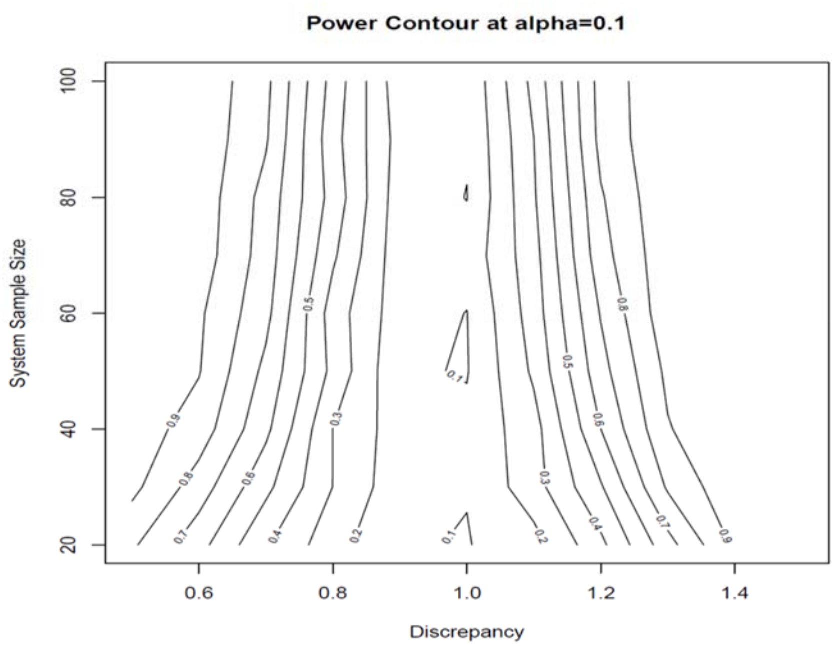


Figure 7: The plots of the power changing as a function of (a) the true discrepancy δ_1 at $N_s = 100$ and (b) the system sample size N_s at $\delta_1 = 0.6$.

The contour plot in Figure 8(a) shows how the power changes simultaneously with the true discrepancy $\delta_1 \in [0.5,1.5]$ and the sample size of the system test data $N_s \in [20,100]$. Each line corresponds to the combinations of δ_1 and N_s associated with the same labeled power value. For any fixed δ_1, the power increases as N_s increases (moving upwards along a vertical line for a fixed δ_1); while for any fixed system sample size N_s, the power is the smallest around $\delta_1 = 1$ (with its minimum achieved at the size of the test at $\delta_1 = 1$) and increases as the discrepancy becomes larger (moving towards both ends of the horizontal line for a particular amount of system data, N_s). Note the power is also dependent on our choice of the threshold value α. Figure 8(b) shows the contour plot of the power at the choice of $\alpha = 0.1$. We can see the power increases as the α increases, i.e. using a less stringent test increases the ability to detect the same amount of discrepancy.



(a) Power contour at $\alpha = 0.05$



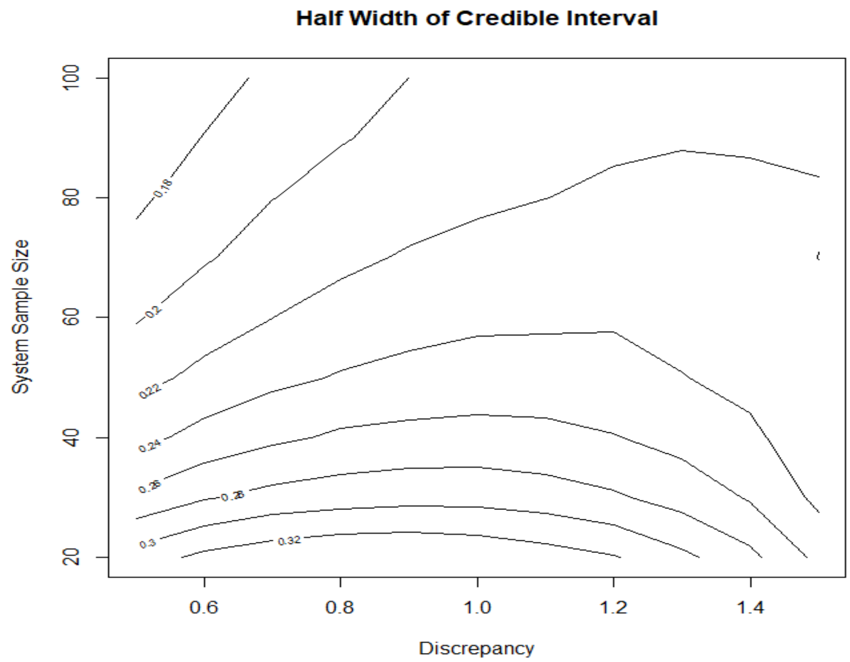
(b) Power contour at $\alpha = 0.1$

Figure 8: The contour plots of the power of the two-sided test for $H_0: \delta_1 = 1$ for different combinations of $N_s \in [20,100]$ and $\delta_1 \in [0.5,1.5]$ for different choices of (a) $\alpha = 0.05$ and (b) $\alpha = 0.1$.

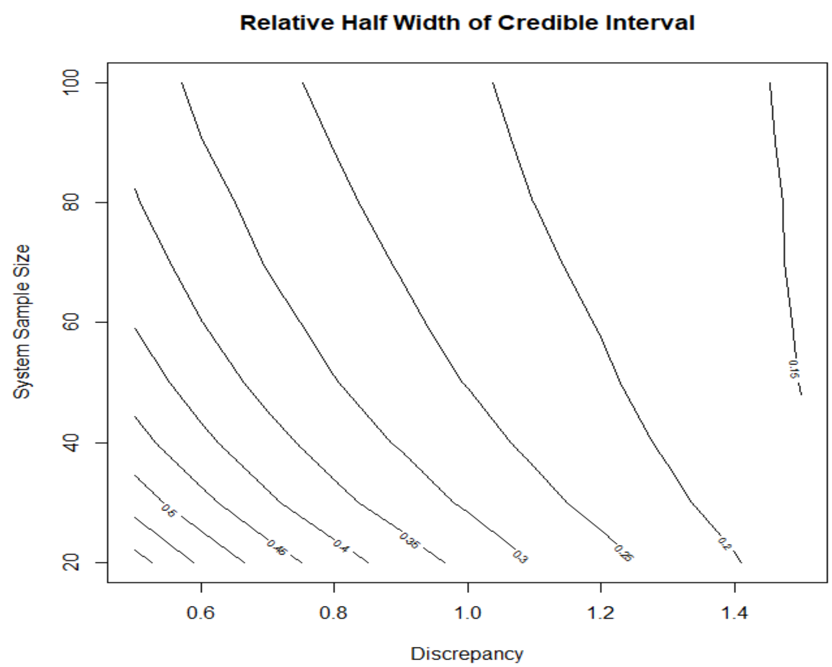
To quantify the precision of the estimated multiplicative discrepancy measure, we calculate the half width of the 95% credible interval of δ_1, $(l_{\delta_1}, u_{\delta_1})$, which is given by $HW_{\delta_1} = (u_{\delta_1} - l_{\delta_1})/2$ and measures the margin of error of the estimated discrepancy. The smaller HW_{δ_1} is, the more precise the estimated discrepancy is. Figure 9 (a) shows the contour plot of the average HW_{δ_1} over the $M = 1000$ simulations. It is quite prominent that in general, increasing the sample size increases the precision of the estimated discrepancy. For a fixed sample size at the system level, the credible intervals are generally narrower for discrepancy values further from 1. In other words, the larger the discrepancy is in either direction, the more precise the estimate of it is. However, increasing sample size is more effective for improving the precision of the estimated discrepancy when the system reliability is substantially lower than the expected performance from the component data (corresponding to small δ_1 values).

To adjust the impact of the size of δ_1 on the measured precision of HW_{δ_1}, we also calculated the relative half width of the credible interval, given by $RHW_{\delta_1} = HW_{\delta_1}/\delta_1$. Figure 9(b) shows the contour plot of RHW_{δ_1}. Increasing the system sample size N_s generally improves the relative precision of the estimated discrepancy. However, there is a diminishing effect as δ_1 increases. As δ_1 increases to 1.5 or larger, increasing the sample size has a negligible impact on the relative precision. Also, larger N_s and δ_1 values are generally associated with the better relative prediction of the estimated δ_1. This is evidenced by the smallest RHW_{δ_1} at the top right corner of Figure 9(b) which reduces as N_s and δ_1 reduce towards the bottom left corner. Note the change in the contour pattern between Figures 9(a) and 9(b) is mainly due to the choice of different metrics for quantifying precision vs. relative precision. Since the relative precision is higher for larger δ_1 value at the same precision level, the relative precision is consistently better for large δ_1 values. If the additive discrepancy, δ_2 is used in the simulation, we would expect a similar pattern for the half width of credible intervals, HW_{δ_2}, which in general reduces as $|\delta_2|$ and N_s increase. However, if relative precision given by $RHW_{\delta_2} = HW_{\delta_2}/|\delta_2|$, is considered, then the contour plot is expected to remain approximately symmetric around $\delta_2 = 0$.

The illustrated method and graphical summaries can be used for choosing an appropriate sample size for a particular application given the available testing resources across a range of potential discrepancy levels. For example, in many real applications, the amount of component data available is often known ahead of time, and then the amount of system data required to achieve a certain power for detecting some anticipated range of discrepancy is determined later. In other applications when some test data from the component and system levels are already collected, we may want to explore what sizes of discrepancy are likely to be detected with the available test data.



(a) Contour plot of HW_{δ_1}



(b) Contour plot of RHW_{δ_1}

Figure 9: The contour plots of the average (a) HW_{δ_1} and (b) RHW_{δ_1} summarized over the $M = 1000$ simulations for different combinations of $N_s \in [20,100]$ and $\delta_1 \in [0.5,1.5]$.

The Probability of Agreement between System and Component Reliabilities

An alternative to the regular hypothesis test proposed is the equivalence test (Wellek, 2010). Instead of assuming the system reliability and the component aggregated reliability are the same and seeking statistically significant evidence for contradicting the null hypothesis, the equivalence test assumes the two ways of calculating the reliabilities are different. The test then focuses on if the difference is large enough to be practically meaningful and worthwhile to be treated as different. In our case, if the discrepancy between the system and component data are not sufficiently large to be of practical importance in terms of assessing the system reliability, then we can comfortably use the cheaper and easier method to obtain component test data to help predict the system reliability. Stevens et al. (2015, 2017) propose the probability of agreement (PoA) as an alternative to the equivalence test for comparing measurement systems. Stevens and Anderson-Cook (2017a,b) extend the PoA to assess the similarity between reliability curves and surfaces. Stevens et al. (2018) develop the Bayesian PoA for measuring the predictive probability of two surfaces being practically equivalent with the differences within a specified threshold value.

We adopt the Bayesian PoA for assessing the consistency between the system and component data under the Bayesian framework. Particular, we conclude that the system and component data are sufficiently consistent for practical purposes when the multiplicative discrepancy, $\delta_1 \in [\Delta_{1l}, \Delta_{1u}]$. Usually, $[\Delta_{1l}, \Delta_{1u}]$ is chosen to be centered around 1. Then the Bayesian predictive PoA is defined as

$$BPoA_1 = \Pr(\Delta_{1l} \leq \delta_1 \leq \Delta_{1\mathrm{u}} | y_s, y_1, \cdots, y_n), \qquad [25]$$

which measures the likelihood that the discrepancy δ_1 is within the specified region of practical equivalence given the observed test data. The $BPoA_1$ can be estimated by the fraction of MCMC draws of δ_1 within the specified threshold values $[\Delta_{1l}, \Delta_{1u}]$ out of a total of M samples from the MCMC simulations. Similarly, if the additive discrepancy δ_2 is used to measure the discrepancy with the system and component data are considered practically equivalent if $\delta_2 \in [\Delta_{2l}, \Delta_{2u}]$, with $[\Delta_{2l}, \Delta_{2u}]$ often chosen to be centered around 0, then the Bayesian PoA is defined as

$$BPoA_2 = \Pr(\Delta_{2l} \leq \delta_2 \leq \Delta_{2\mathrm{u}} | y_s, y_1, \cdots, y_n). \qquad [26]$$

Consider the complex system II as an example. For the two sets of synthetic data generated in Tables 4 & 5, we calculated the Bayesian PoA values for different threshold value choices. For the first set of data generated from the system in Figure 5(a), the true discrepancy between system and component data is relatively large with $\delta_1 = 0.86$ and $\delta_2 = -0.13$ compared to the specified choices of threshold values at $[\Delta_{1l}, \Delta_{1u}] = [0.95, 1.05]$ or [0.9, 1.1] and $[\Delta_{2l}, \Delta_{2u}] = [-0.05, 0.05]$ or [-0.1, 0.1]. As a result, the calculated Bayesian PoAs are all small (<0.3), indicating that it is unlikely that the difference between the reliability estimation approaches are as similar as needed. However, for the second set of data generated from the system in Figure 6(a), the true discrepancy between the system and component data is small with $\delta_1 = 1.02$ and $\delta_2 = 0.02$, which are generally smaller than the specified thresholds values. When the more demanding thresholds at $\delta_1 \in [0.95,1.05]$ and $\delta_2 \in [-0.05,0.05]$ are used, the Bayesian PoAs are between 0.6 and 0.7, which indicates moderate agreement under the chosen practical difference values. When the wider thresholds at $\delta_1 \in [0.9,1.1]$ and $\delta_2 \in [-0.1,0.1]$ are considered, the Bayesian PoAs are above 0.99, which indicate extremely high agreement in terms of the specified practical difference.

Table 6: Bayesian PoA for the complex system II synthetic data generated in Tables 4 & 5.

Complex System II-a (Scenario 1: C is disconnected)			Complex System II-b (Scenario 2: C has a by-pass)		
Discrepancy	$[\Delta_{il}, \Delta_{iu}]$	$BPoA_i$	Discrepancy	$[\Delta_{il}, \Delta_{iu}]$	$BPoA_i$
$\delta_1 = 0.86$	[0.95,1.05]	0.009	$\delta_1 = 1.02$	[0.95,1.05]	0.619
	[0.9,1.1]	0.189		[0.9,1.1]	0.995
$\delta_2 = -0.13$	[-0.05,0.05]	0.012	$\delta_2 = 0.02$	[-0.05,0.05]	0.689
	[-0.1,0.1]	0.253		[-0.1,0.1]	0.999

Conclusions and Discussion

Engineering understanding of the system structure can be leveraged to use often cheaper component test data to improve the assessment of system reliability. This usually requires the reliability engineers to make assumptions about the failure mechanisms and the relationships between system and component functionalities. It is important to test these assumptions about the system structure with available data to ensure an accurate prediction of system reliability. This paper proposes a Bayesian approach for evaluating the system structure based on system and component test data. The method estimates the multiplicative or additive discrepancy between the system and component test data under the assumed structure with quantified uncertainty to understand both the direction and the degree of the discrepancy indicated by the observed test data. Reliability block diagram, the minimum path, and the minimum cut approaches are used to develop the reliability functions that connect the system and component reliabilities. By using MCMC simulation based on Gibbs sampler, we can generate a large number of MCMC samples from the approximate posterior distribution to estimate the discrepancy measures, system and component reliabilities, and Bayesian posterior predictive p-value for testing the significance of the inconsistency. The method is flexible and straightforward to adapt to a variety of different system structures, which was demonstrated through several examples of series, parallel, and more complex system structures.

One important decision to make when using this method is to choose between the multiplicative and additive discrepancy measures. Both measures have demonstrated similar effectiveness for testing the discrepancy for the illustrated examples. We recommend making a choice based on considering the interpretability of the measures for a particular application. For example, if the engineering knowledge suspect there might be a potentially omitted pseudo-component that is connected with other components in a series structure, then the multiplicative discrepancy could directly provide information about the reliability of the pseudo-component, since a component in a series structure contribute to the system reliability in a multiplicative form as given by $p_s = \prod_{i=1}^{n} p_i$. On the other hand, if the system has a parallel structure, then the additive discrepancy could be more informative since the contribution from each of the redundancy component is more close to an additive form by ignoring the higher order terms which are generally smaller than the linear term in the reliability function given by $p_s = 1 - \prod_{i=1}^{n}(1 - p_i) = \sum_{i=1}^{n} p_i + \sum_{i \neq j}(-1)^{n+1} p_i p_j + o(p_i p_j)$. This also applies for more complex systems which can be simplified to either a parallel system of minimal paths or a series system of minimal cuts. The gained understanding about the discrepancy is also useful for system structure diagnostics and failure detection.

It is important to determine how much test data is needed to detect a certain amount of discrepancy between the different types of data. For the system evaluation, the component test data are generally cheaper and easier to obtain. Therefore we used a simulation study to explore the impact of the amount of system test data and the true discrepancy value on the power of the test and the precision of the estimated discrepancy for a series system. The graphical summaries shown are useful to determine how much system test data is needed to detect a range of discrepancy of interest to achieve a desired power or precision of the estimate. In addition, it is helpful to consider what range of discrepancy we can detect with sufficient power or precision.

The Bayesian probability of agreement (BPoA) was illustrated as an alternative strategy to the hypothesis test-based summary. A statistically significant discrepancy in a large sample case could still be too small to be of practical importance. On the other hand, a practically meaningful discrepancy could be deemed statistically insignificant without sufficient test data. Instead of relying only on statistical significance which is a function of sample size, the BPoA provides a useful summary on whether the discrepancy is of practical importance for a particular application. We recommend examining alternative summaries for obtaining a comprehensive understanding of the system structure to make appropriate, practical and data-driven decisions about the assessment and management of system reliability.

About the Author

Dr. Lu Lu is an Assistant Professor of Statistics in the Department of Mathematics and Statistics at the University of South Florida in Tampa. Before joining the department, she was a postdoctoral research associated in the Statistics Sciences Group at Los Alamos National Laboratory. She earned a doctorate in Statistics from Iowa State University in Ames, IA. Her research interests include reliability analysis, design of experiments, response surface methodology, survey sampling, multiple objective/response optimization.

References:

1. Anderson-Cook, C.M. (2009) "Evaluating the Series or Parallel Structure Assumption for System Reliability", *Quality Engineering*, 21: 88-95.
2. 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.
3. Anderson-Cook, C.M., Graves, T., Hengartner, N., Klamann, R., Wiedlea, A.K., Wilson, A.G., Anderson, G., Lopez, G. (2008) "Reliability modeling using both system test and quality assurance data", *Journal of the Military Operations Research Society*, 13: 5-8.
4. Casella, G., and George, E.I. (1992) "Explaining the Gibbs sampler", *The American Statistician*, 46: 167-174.
5. Gelman, A. (2013) "Two Simple Examples for Understanding Posterior P-values Whose Distributions Are Far From Uniform", *Electronic Journal of Statistics*, 7: 2595-2602.
6. Gelman, A. Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (2013) *Bayesian Data Analysis*, Chapman and Hall/CRC, 3rd Edition.
7. Hamada, M.S., Alyson, G.W., Reese, C.S., and Martz, H.F. (2008) *Bayesian Reliability*, Springer.
8. Martz, H.F., Waller, R.A., and Fickas, E.T. (1988) "Bayesian Reliability Analysis of Series Systems of Binomial Subsystems and Components", *Technometrics*, 30(2): 143-154.
9. Martz, H.F., Waller, R.A. (1990) "Bayesian Reliability Analysis of Complex Series/Parallel Systems of Binomial Subsystems and Components", *Technometrics*, 32(4): 407-416.
10. Plumber, M. (2003) "JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling", *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20–22, Vienna, Austria. ISSN 1609-395X.

11. Stevens, N.T., Steiner, S.H., and MacKay, R.J. (2015) "Assessing Agreement Between Two Measurement Systems: An Alternative to the Limits of Agreement Approach", *Statistical Methods in Medical Research*, in press.
12. Stevens, N.T. and Anderson-Cook, C.M. (2017a) "Comparing the Reliability of Related Populations with the Probability of Agreement", *Technometrics*, 59(3), 371-380.
13. Stevens, N.T. and Anderson-Cook, C.M. (2017b) "Quantifying Similarity in Reliability Surfaces Using the Probability of Agreement", *Quality Engineering*, 29(3), 395-408.
14. Stevens, N.T., Rigdon, S.E., and Anderson-Cook, C.M. (2018) "Bayesian Probability of Predictive Agreement for Comparing the Outcome of Two Separate Regressions", Quality and Reliability Engineering, https://doi.org/10.1002/qre.2284.
15. R Development Core Team. 2016. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Available at http://www.R-project.org.
16. Tobias, P.A. and Trindade, D.C. (2012) *Applied Reliability*, CRC Press, 3rd edition.
17. Wilson, A.G., Graves, T.L., Hamada, M.S., Reese, C.S. (2006) "Advances in data combination, analysis and collection for system reliability assessment", *Statistical Science*, 21: 514-531.
18. Wellek S. (2010) *Testing Statistical Hypotheses of Equivalence and Noninferiority*. CRC Press, London. 2nd Edition.