LA-UR- 00-4201

Approved for public release; distribution is unlimited.

Title: Estimating Reliability Trends for the World's Fastest Computer

Author(s): Kenneth J. Ryan
C. Shane Reese

Submitted to: Technometrics

Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Form 836 (8/00)

Estimating Reliability Trends for the World's Fastest Computer

Kenneth J. Ryan and C. Shane Reese *

September 6, 2000

ABSTRACT

Los Alamos National Laboratory is home to the World's fastest computer–Blue Mountain. This machine was created by parallelizing "desktop" computers. To determine whether or not this type of architecture represents the future of super-computing, reliability must be estimated. This paper presents and analyzes failure data of Blue Mountain. Non-homogeneous Poisson processes are fit to the data within a Bayesian hierarchical framework. The task of selecting hyperparameters is discussed, and Bayes factors are used to compare models.

Key Words: Bayesian Hierarchical Model, Bayes Factor, Poisson Process.

1 Introduction

When modeling failure time data, a distinction must be made between one-time-use and multiple-time-use or repairable systems. When a one-time-use system fails, it is simply replaced by a new system of the same type. A light bulb is an example of a one-time-use system. To study the failure-time properties of a one-time-use system, suppose n of these systems are tested. In

*Kenneth J. Ryan is a graduate student in the Department of Statistics at Iowa State University, Ames, IA, 50011-1210. C. Shane Reese is a Technical Staff Member in the Statistical Sciences group at Los Alamos National Laboratory, Los Alamos, NM, 87545.

this case, treating the failure times as an *i.i.d* sample from some population usually suffices. However, models for the failure times of a single repairable system need to be able to quantify the reliability growth or decay of the system. For example, consider a complex piece of computer software. When the program "fails" because it is presented with a set of inputs that it is not able to properly handle, programmers add and alter the code so that the bug that caused the failure is (hopefully) gone. The software should thus have a reliability growth–there should be fewer failures occurring with decreasing frequency.

To deal with failure time data from a single repairable system, we will need some notation. First, let T_i be the time at which the ith failure occurs. The failure times of a single repairable system satisfy $0 < T_1 < T_2 < \ldots$. Next, we define the inter-failure times as $Y_i = T_i - T_{i-1}$ (with $T_0 = 0$). Finally, let

- $N(a, b)$ be the number of failures in an interval $(a, b]$
- $N(t)$ be the number of failures in $(0, t]$.

A simple model for the failure times of a repairable system are renewal processes. For a renewal process, the inter-failure times Y_i are an *i.i.d* sample from some population. Under a renewal process, the time to the next failure has the same distribution whether the system is brand new or has just been repaired for the 100th time. The terminology "as-good-as-new" is thus associated with a renewal process. Although a renewal process may adequately describe the failure times of a repairable system, the point of studying many repairable systems is to determine whether the system is experiencing reliability growth or decay, and a renewal process is unable to capture these aspects of a repairable system.

Another class of models for failure times of a repairable system is that of the non-homogeneous Poisson processes (NHPP). An NHPP is defined by its intensity $\nu(t)$. For an NHPP,

- $N(a,b)$ is a Poisson random variable with mean $\mu(a,b) = \int_a^b \nu(t)dt$
- $N(0) = 0$
- for disjoint intervals (a_1, b_1) and (a_2, b_2) (i.e. intervals for which either $b_1 < a_2$ or $b_2 < a_1$), $N(a_1, b_1)$ and $N(a_2, b_2)$ are independent.

NHPPs have received much attention in the literature. Duane (1964) conducted an empirical study to determine a general relationship for the failure times of a repairable system undergoing a testing process that involves a repair with engineering modification at each failure. For five different electromechanical and mechanical repairable systems, Duane collected failure data and noted that plots of cumulative failure rates versus cumulative operating hours were approximately linear on log-log paper. Crow (1974) suggested modeling a repairable system under such "find it and fix it" conditions conforming to the Duane relationship using an NHPP with a power law intensity

$$\nu(t) = \frac{\phi}{\eta}\left(\frac{t}{\eta}\right)^{\phi-1},$$

where both η and ϕ are positive parameters. In the literature, this model is referred to as a Weibull or power law process (PLP). Note that for a PLP, the mean number of failures up to time t is

$$\mu(t) = \left(\frac{t}{\eta}\right)^{\phi}.$$

We will also use the notation that $\mu(a,b)$ is the mean number of failures in the interval (a,b).

Under a PLP, Crow discussed classical point and interval estimation and hypothesis testing. This was done in a variety of situations assuming both time truncated and failure truncated data. He also considered the case where properties of multiple repairable systems (of the same

type) need to be estimated or tested and provided numerical examples and discussed a variety of applications where a PLP may be a useful model.

Finkelstein (1976) provided exact $100(1-\alpha)\%$ confidence intervals for the two parameters of a PLP based on exact failure times where the data for a single repairable system are failure truncated. The intervals are based on pivotal quantities involving the maximum likelihood estimators. A table of necessary cut-off points (calculated by Monte Carlo methods) is provided.

Lee and Lee (1978) provided an exact $100(1-\alpha)\%$ prediction interval for T_{n+k} based on exact failure times for a single repairable system where the data are failure truncated (i.e., the first n failures are observed). T_{n+k} may be useful in determining the length of a failure truncated study with "sample size" $n+k$.

If the "find it and fix it" stage of a development process concludes at the end of a study, it may be reasonable to assume that thereafter the system can be modeled as a homogeneous Poisson process (HPP) (i.e., an NHPP with constant intensity). In this case, $\nu(T_n)$ would be the constant future intensity and directly related to the system's "production time" reliability. Lee and Lee also provided an exact $100(1-\alpha)\%$ confidence interval for $\nu(T_n)$. Cut-off points (calculated by numerical integration) for inference for both T_{n+k} and $\nu(T_n)$ are provided.

A log-linear NHPP is an NHPP with intensity

$$\nu(t) \quad = \quad exp(\gamma + \delta t),$$

where the parameters δ, $\gamma \in \mathbb{R}$. Meeker and Escobar (1998) plotted maximum likelihood fits of $\mu(t)$ versus t for both the PLP and log-linear NHPP for two different data sets. They noted that the fits for these two models are similar for both data sets. This suggests that the log-linear intensity is flexible like the power intensity. Aven (1989) derived uniformly most power tests to compare the parameters of two repairable systems modeled using log-linear NHPPs.

We have noted that a renewal process is not capable of modeling reliability decay or growth in a repairable system because (under a renewal process) the system is "as-good-as-new" after a repair. A parallel criticism of the PLP is that after repair the system is "as-bad-as-old". The reason for this is that if system repair time is negligible then the intensity before the failure is the same as that after the repair. A compromise seems in order as it may be the case that a repair does not make the system brand new but does improve it.

To model this possibility, Black and Rigdon (1996) suggested what he called a modulated PLP (MPLP) which is a special case of a class of models presented by Berman (1981). An MPLP is essentially a PLP, but has another positive parameter κ, called the shock parameter. Berman presented methods for simulating failure times from a MPLP and provided an interpretation for the shock parameter in the case when it is a positive integer. For example, if $\kappa = 4$, a failure occurs at every fourth occurrence of a PLP with parameters ϕ and η. Thus, if the intensity of an MPLP is increasing (i.e. $\phi > 1$) and $\kappa > 1$, then the probability of a failure in a small interval just after a failure is smaller than the probability of a failure in an interval of the same length before the failure but larger than the probability of a failure in an interval of this length when the system was brand new. Black and Rigdon stated that an MPLP reduces to

- a gamma renewal process when $\phi = 1$
- a PLP when $\kappa = 1$
- an HPP when $\phi = \kappa = 1$.

Black and Rigdon also considered inference for MPLPs. He presented approximate $100(1-\alpha)\%$ confidence intervals for the parameters of a MPLP based on the asymptotic properties of maximum likelihood estimators. A simulation study assessed the coverage probabilities of these

interval estimators. Approximate likelihood ratio tests were used as goodness of fit tests to compare PLPs, renewal processes, and HPPs to MPLPs. A simulation study assessed the power function for these tests. Numerical examples of both the intervals and tests are provided.

Sen (1998) introduced a reliability growth model where the inter-failure times Y_i are independent exponential random variables with hazard function

$$\mu_i = \left(\frac{\mu}{\delta}\right) i^{1-\delta},$$

where $\mu > 0$ and $\delta > 1$. Note that μ_i is a strictly decreasing sequence in i. Thus, the system after a repair is "better than brand new". Obvious extensions to the parameter space of δ would allow the μ_i to be an increasing sequence in i. This would allow Sen's model to exhibit the worse than "brand new" but better than "as-good-as old" property of some MPLPs. Sen restricted his parameter space to model only reliability growth because the system being studied was in a design phase. Thus, a repair with engineering modification was only supposed to make the system "better than brand new".

Sen also derived maximum likelihood and least squares estimators for μ and δ. Large-sample properties of these estimators were also derived. As a test of robustness for the statistical methods he crafted, Sen used the maximum likelihood estimator of the PLP intensity to estimate the hazard function for his own model. This estimator–even under the wrong model–was found to be consistent. But, under this model misspecification, Sen found that precision may be underestimated.

Bayesian approaches for the PLP also have a history in the reliability growth literature. Higgins and Tsokos (1981) proposed a quasi-Bayesian estimator of $\nu(t)$ for a PLP. This estimator is easy to compute, allows for use of prior knowledge, and performs well compared to the corresponding maximum likelihood estimator. Littlewood and Verrall (1989) introduced a

Bayesian reliability growth model for computer software.

Kyparisis and Singpurwalla (1985) presented a data analysis on software failure times using a Bayesian PLP model. Guida *et al.* (1989) and Calabria *et al.* (1990) presented both non-informative and informative priors for a PLP. The description of the informative priors allows for an easy transfer from informal prior knowledge to a prior distribution. Guida *et al.* compared Bayesian point and interval estimation for the PLP parameter to maximum likelihood methods. Calabria *et al.* considered the problem of predicting T_{n+k} from $T_1, T_2, \ldots, T_n$ and compared their methods to maximum likelihood methods. Coverage probabilities and relative average interval length were used as comparison criteria.

Using priors presented by Guida *et al.* and Calabria *et. al*, Bar-Lev *et al.* (1992) fit a Bayesian PLP to two different data sets. The data and posterior summaries are provided. Bar-Lev *et al.* pointed out that the Bayesian method provides a "unified methodology" for dealing with exact failure data, since the method is the same regardless of whether the data are time or failure truncated. Maximum likelihood methods, however, depend on the data collection method.

We present a reliability study of a supercomputer consisting of 48 sub-computers. Essentially, these 48 sub-computers are 48 repairable systems in series. A job submitted to this computer will finish only if none of the sub-computers requested fail while the job is processing. (Jobs on this supercomputer are not always run on all 48 sub-computers.) Since the sub-computers are of the same brand, a Bayesian PLP hierarchical model is a natural choice because it makes sense to think of these sub-computers as coming from some population for which we do not know the values of the parameters. To our knowledge, this type of model has never been introduced in the literature. In Section 2, we present a description of the supercomputer. In Section 3.1, we

review the literature for a Bayesian PLP model for a single system. In Section 3.2, we introduce our Bayesian hierarchical PLP model for multiple repairable systems. This model is defined in a way such that elicitation of the hyperparameters is simple–given some expert opinion. Section 4.1 shows summaries from the fit of the model described in Section 3.2. Also in Section 4.1, we determine the current reliability of the supercomputer. We define reliability to be the probability that a job of length l submitted to all 48 sub-computers at start time s will finish. We will use this metric to assess the reliability growth (or decay) of the supercomputer over time. In Section 4.2, to assess whether or not the hierarchy in the model is needed, simpler Bayesian models are also fit. Then, Bayes factors are used to compare the fits of these competing models.

2 The Blue Mountain Supercomputer

The basic repairable system that we study is a sub-computer. Assume that there are C sub-computers in all. Also, suppose that, $N_i(0,t)$, $N_i(t,2t),\ldots,N_i((M-1)t,Mt)$ is recorded for each computer $i=1,2,\ldots,C$, where $t>0$ is fixed and known. Note that the times at which data are collected are equally spaced. Figure 1 is a diagram of these times. To make latter notation more concise, let x_{ij} be the number of failures for the ith computer during the jth time interval. That is, let $x_{ij}=N_i((j-1)t,jt)$ for $i=1,2,\ldots,C$ and $j=1,2,\ldots,M$.

The Blue Mountain supercomputer consists of 48 "desktop" or sub-computers. There are 128 processors per sub-computer, and there is a complicated interconnect that links the sub-computers together. Periodically, a sub-computer will "fail". These failures are hardware related. For example, a memory module may need to be replaced. When one of the sub-computers fails, the sub-computer is repaired. Then, the sub-computer is restarted. If a job is submitted to a sub-group of the 48 sub-computers and one of the sub-computers in the sub-group fails, the

job will not finish and will need to be resubmitted. The number of failures per month for each sub-computer is recorded. These data for the first 9 months of operation are provided in Table 1. (Thus, in our case, $N = 48$ sub-computers, and $M = 9$ equally spaced data collection times that are $t = 1$ month apart.)

3 Bayesian PLP Models

3.1 One-System Models

Consider the data collection scheme described in Section 2, and consider modeling the ith sub-computer with a PLP with parameters ϕ and η. Let $\underline{x} = (x_{i1}, x_{i2}, \ldots, x_{iM})$ be the vector of failure counts for the ith sub-computer. Then, the sampling distribution for the failure counts for the ith sub-computer in time interval j is

$$x_{ij}|\phi,\eta \overset{indep}{\sim} Poisson\left(\left(\frac{tj}{\eta}\right)^{\phi} - \left(\frac{t(j-1)}{\eta}\right)^{\phi}\right) \text{ for } j = 1, \ldots, M,$$

and $\underline{x}$ has probability mass function

$$p(\underline{x}|\phi,\eta) = \prod_{j=1}^{M} \left[\frac{\left\{\left(\frac{tj}{\eta}\right)^{\phi} - \left(\frac{t(j-1)}{\eta}\right)^{\phi}\right\}^{x_{ij}}}{x_{ij}!}\right] exp\left\{-\left(\frac{Mt}{\eta}\right)^{\phi}\right\}.$$

As mentioned in Section 1, Guida *et al.* (1989) presented an easy to elicit informative prior on (ϕ, η). The suggested procedure places a gamma distribution on the expected number of failures up to some specified time T, $\mu(T)$. All the expert needs to do is provide a value for T and a mean μ and standard deviation σ for $\mu(T)$. After a change of variables, the prior density for $\eta|\phi$ is

$$p(\eta|\phi) \quad = \quad \phi\left(\frac{\mu}{\sigma^2}\right)^{\left(\frac{\mu}{\sigma}\right)^2} T^{\phi\left(\frac{\mu}{\sigma}\right)^2} \eta^{-\phi\left(\frac{\mu}{\sigma}\right)^2-1} exp\left[-\frac{\mu}{\sigma^2}(T/\eta)^{\phi}\right] / \Gamma\left(\left(\frac{\mu}{\sigma}\right)^2\right),$$

where μ, σ, and T are given by an expert. With the conditional prior for $\eta|\phi$ specified, all that is needed is a marginal prior distribution for ϕ. Guida *et al.* suggested a uniform distribution with end points

- $(0.3, 1.1)$ when there is a "strong conviction of a reliability growth, but no information on what the ϕ value (≤ 1) is"
- $(0.3, 3.0)$ when there is "weak information about the failure process"
- $(1.0, 5.0)$ when there is a "strong conviction of a degradation phenomena, but weak information on what the ϕ value (> 1) is".

Note that when $t = \eta$, $\mu(t) = 1$. Thus, η can be interpreted as the time in which you expect the first failure. If $\phi = 3$, you expect one failure in the first η time units, and you expect $2^3 - 1 = 8$ failures in the second η time units. So, in many applications, it seems reasonable that $\phi > 3$ would indicate more decay than expected and that an expert would be able to rule these values out. Thus, the Guida *et al.* choice of priors for ϕ seems to make sense.

A more flexible prior for ϕ was suggested by Kyparisis and Singpurwalla (1985). Kyparisis and Singpurwalla suggested a scaled beta distribution with density

$$p(\phi) = \frac{\Gamma(k_1 + k_2)}{\Gamma(k_1) + \Gamma(k_2)} \frac{(\phi - l)^{k_1 - 1}(u - \phi)^{k_2 - 1}}{(u - l)^{k_1 + k_2 - 1}},$$

where $0 \leq l < \phi < u$ and k_1, $k_2 > 0$. For a scaled beta distribution with mean $\mu \in (l, u)$, the variance σ^2 must be less than $(u - l)^2 \mu(1 - \mu)$. Beginning with a valid mean and variance for a scaled beta distribution, the parameters k_1 and k_2 are

$$\begin{aligned} k_1 &= \frac{\mu^2(1 - \mu)}{\sigma^2} - \mu \\ k_2 &= \frac{\mu(1 - \mu)^2}{\sigma^2} - (1 - \mu). \end{aligned}$$

3.2 A Multiple-System Hierarchical Model

Suppose the number of failures for the ith sub-computer follows a PLP with intensity parameters ϕ_i and η_i. Define $\underline{\phi} = (\phi_1, \phi_2, \ldots, \phi_C)$ and $\underline{\eta} = (\eta_1, \eta_2, \ldots, \eta_C)$. Furthermore, suppose that given $\underline{\phi}$ and $\underline{\eta}$, sub-computers fail independently. That is,

$$x_{ij}|\underline{\phi_i}, \underline{\eta_i} \overset{ind}{\sim} Poisson\left(\left(\frac{tj}{\eta_i}\right)^{\phi_i} - \left(\frac{t(j-1)}{\eta_i}\right)^{\phi_i}\right) \quad \text{for } i = 1, \ldots, N,\ j = 1, \ldots, M.$$

Thus, the sampling distribution for the data $X = [x_{ij}]$ (an $N \times M$ matrix) has probability mass function

$$p(X|\underline{\phi}, \underline{\eta}) = \prod_{i=1}^{C}\left[\prod_{j=1}^{M}\left[\frac{\left\{\left(\frac{tj}{\eta_i}\right)^{\phi_i} - \left(\frac{t(j-1)}{\eta_i}\right)^{\phi_i}\right\}^{x_{ij}}}{x_{ij}!}\right] exp\left\{-\left(\frac{Mt}{\eta_i}\right)^{\phi_i}\right\}\right].$$

Using priors for $\underline{\eta}$ as suggested by Guida *et. al*, let

$$\begin{aligned} p(\underline{\eta}|\underline{\phi}, \mu_T, \sigma_T) &= \prod_{i=1}^{C}\left[\phi_i\left(\frac{\mu_T}{\sigma_T^2}\right)^{\left(\frac{\mu_T}{\sigma_T}\right)^2} T^{\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2} \eta_i^{-\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2 - 1} exp\left[-\frac{\mu_T}{\sigma_T^2}(T/\eta_i)^{\phi_i}\right] / \Gamma\left(\left(\frac{\mu_T}{\sigma_T}\right)^2\right)\right] \\ &= \left(\frac{\left(\frac{\mu_T}{\sigma_T^2}\right)^{\left(\frac{\mu_T}{\sigma_T}\right)^2}}{\Gamma\left(\left(\frac{\mu_T}{\sigma_T}\right)^2\right)}\right)^N \prod_{i=1}^{C}\left[\phi_i T^{\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2} \eta_i^{-\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2 - 1} exp\left[-\frac{\mu_T}{\sigma_T^2}(T/\eta_i)^{\phi_i}\right]\right]. \end{aligned}$$

Next, we suggest a gamma prior distribution for $\underline{\phi}$ that is parameterized in terms of the

mean μ_ϕ and variance σ_ϕ. That is, we use density

$$
\begin{aligned}
p(\underline{\phi}|\mu_\phi,\sigma_\phi) &= \prod_{i=1}^{C}\left[\frac{\left(\frac{\mu_\phi}{\sigma_\phi^2}\right)^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2}}{\Gamma\left(\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2\right)}\phi_i^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2-1} exp\left(-\frac{\mu_\phi}{{\sigma_\phi}^2}\phi_i\right)\right] \\
&= \left(\frac{\left(\frac{\mu_\phi}{\sigma_\phi^2}\right)^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2}}{\Gamma\left(\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2\right)}\right)^C \prod_{i=1}^{C}\left[\phi_i^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2-1} exp\left(-\frac{\mu_\phi}{{\sigma_\phi}^2}\phi_i\right)\right].
\end{aligned}
$$

So, the distribution of $(\underline{\phi},\underline{\eta}|\mu_T,\sigma_T,\mu_\phi,\sigma_\phi)$ has density

$$
p(\underline{\eta},\underline{\phi}|\mu_T,\sigma_T,\mu_\phi,\sigma_\phi) = p(\underline{\eta}|\underline{\phi},\mu_T,\sigma_T)p(\underline{\phi}|\mu_\phi,\sigma_\phi).
$$

Finally, let

$$
\begin{aligned}
\mu_T &\sim Weibull(a_{\mu_T},b_{\mu_T}) \\
\sigma_T &\sim Weibull(a_{\sigma_T},b_{\sigma_T}) \\
\mu_\phi &\sim Weibull(a_{\mu_\phi},b_{\mu_\phi}) \\
\sigma_\phi &\sim Weibull(a_{\sigma_\phi},b_{\sigma_\phi})
\end{aligned}
$$

where the $Weibull(a,b)$ distribution has density

$$
p(x) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} exp\left[-\left(\frac{x}{b}\right)^a\right].
$$

Thus, $(\mu_T,\sigma_T,\mu_\phi,\sigma_\phi)$ has density

$$
p(\mu_T,\sigma_T,\mu_\phi,\sigma_\phi) \propto \frac{{\mu_T}^{a_{\mu_T}-1}{\sigma_T}^{a_{\sigma_T}-1}{\mu_\phi}^{a_{\mu_\phi}-1}{\sigma_\phi}^{a_{\sigma_\phi}-1}}{exp\left[\left(\frac{\mu_T}{b_{\mu_T}}\right)^{a_{\mu_T}}+\left(\frac{\sigma_T}{b_{\sigma_T}}\right)^{a_{\sigma_T}}+\left(\frac{\mu_\phi}{b_{\mu_\phi}}\right)^{a_{\mu_\phi}}+\left(\frac{\sigma_\phi}{b_{\sigma_\phi}}\right)^{a_{\sigma_\phi}}\right]}.
$$

3.3 Hyperparameter Specification

For the model defined in Section 3.2, a_{μ_T}, b_{μ_T}, a_{σ_T}, b_{σ_T}, a_{μ_ϕ}, b_{μ_ϕ}, a_{σ_ϕ}, and b_{σ_ϕ} are hyperparameters that need to be specified. To choose a_{μ_T} and b_{μ_T}, suppose that an expert provides T and two μ_T, quantile pairs (μ_1, p_1) and (μ_2, p_2). Then, taking

$$a_{\mu_T} = \frac{log\left[\frac{log(1-p_2)}{log(1-p_1)}\right]}{log\left(\frac{\mu_2}{\mu_1}\right)} \quad (1)$$

$$b_{\mu_T} = \frac{\mu_1}{[-log(1-p_1)]^{\frac{1}{a_\mu}}} \quad (2)$$

gives a prior with the desired properties. In our example, the computer expert believes that, for $T = 1$, μ_T is in the interval $(0.5, 15.0)$. Taking these values as the 0.05 and 0.95 quantiles of the prior distribution for μ_T, Equations 1 and 2 imply that $a_{\mu_T} = 1.20$ and $b_{\mu_T} = 5.99$. Similarly, the expert believes that, for $T = 1$, σ_T is in the interval $(0.01, 5.0)$, and, thus, Equations 1 and 2 imply that $a_{\sigma_T} = 0.654$ and $b_{\sigma_T} = 0.935$.

Since Guida *et al.* suggest that $(0.3, 3)$ is a non-informative range for a ϕ parameter, suppose that ϕ_i are in $(0.3, 3)$ for $i = 1, 2, \ldots, C$. Then, μ_ϕ would also be in this interval. Taking 0.3 and 3 as the 0.05 and 0.95 quantiles for μ_ϕ implies $a_{\mu_\phi} = 4.07$ and $b_{\mu_\phi} = 0.623$. Also, σ_ϕ would be at most the standard deviation of a population with half of the ϕ_is at 0.3 and the other half at 3. That is,

$$\begin{aligned} \sigma_\phi &\leq \sqrt{\frac{(.3 - 1.65)^2 + (3 - 1.65)^2}{2}} \\ &= 1.35. \end{aligned}$$

Thus, with 0.01 and 1.35 as the 0.05 and 0.95 quantiles for σ_ϕ implies $a_{\sigma_\phi} = 0.829$ and $b_{\sigma_\phi} = 0.359$.

4 Results

4.1 Fitting the Hierarchical PLP Model

To simulate draws from the posterior distribution of the hierarchical model described in Section 3.2 calculations were done using MCMC methods described in Section A.3. Basically, we used a forward substitution Markov Chain Monte Carlo Algorithm (Gilks *et. al*, 1996). To determine if the Markov Chain had mixed, time series plots of the parameters and the methods introduced by Raftery and Lewis (1996) were used. Summaries of the posterior distribution presented in this section are based on a sample size of 10,000. To obtain this sample, one chain of 1,010,000 iterations was generated. The first 10,000 iterations were discarded. Every 100th iteration was kept thereafter.

A way of assessing the fit of a PLP model is to plot nonparametric estimates and PLP estimates of the mean cumulative function $\mu(t)$ on the same plot. A nonparametric estimate for the ith computer's mean cumulative function at time j is simply the number of failures up to times j. That is, the nonparametric estimate is simply $N_i(j)$ for $j = 1, \ldots, M$. For hierarchical PLP estimates of the mean cumulative function for the ith computer, we use

$$\hat{\mu_i}(t) = \left(\frac{t}{\bar{\mu}_{\eta_i}}\right)^{\bar{\mu}_{\phi_i}}$$

where $\bar{\mu}_{\eta_i}$ and $\bar{\mu}_{\phi_i}$ are the posterior means of μ_i and ϕ_i from the posterior sample of size 10,000. Figure 2 is a plot of nonparametric and hierarchical PLP based estimates of the mean cumulative function for all 48 computers. Note that the hierarchical PLP estimates "shrink" toward the center of the nonparametric estimates. This is known as "borrowing of strength." Since the 48 computers are thought to have come from an underlying population, each computer provides information about the other 47. Thus, the estimates are "pulled" toward the center. On the

other hand, the nonparametric estimates are computed separately for each computer.

A $100(1-\alpha)\%$ highest posterior density region (HPDR) for a parameter θ with posterior density $p(\theta|data)$ is $\{\theta : p(\theta|data) > X_{1-\alpha}\}$, where $X_{1-\alpha}$ is such that

$$\int_{\{\theta:p(\theta|data)>X_{1-\alpha}\}} p(\theta|data)d\theta \quad = \quad 1-\alpha$$

In the case when an HPDR is an interval, it is the shortest interval with a posterior probability of $(1-\alpha)$. From the posterior sample of size 10,000, an analog of a 90% HPD interval can be calculated by

(1) ordering the sample $\theta^1, \theta^2, \ldots, \theta^{10{,}000}$

(2) finding i^* such that $\theta^{i^*+9{,}000} - \theta^{i^*} = \min_{i\in\{1,\ldots,1{,}000\}} \theta^{i+9{,}000} - \theta^i$

Table 2 contains numerical summaries of the posterior sample of size 10,000 for some of the parameters. As described above, $(\theta^{i^*}, \theta^{i^*+9{,}000})$ is reported as the 90% HPD interval. Since the 90% HPD interval for ϕ_1 is below 1, this indicates that sub-computer 1 seems to be undergoing reliability growth. Also note that the 90% HPD interval for μ_ϕ is below 1. This indicates that, on average, the population of sub-computers will undergo reliability growth. Since $T = 1$ was used in specifying the model, the fact that the posterior mean for μ_T is about 3.5 indicates that we expect a new sub-computer to fail about 3.5 times in the first month of use.

For some unobserved quantity of interest, its posterior predictive distribution is its conditional marginal distribution given the data. Under a Bayesian hierarchical model, population parameters are random variables. Thus, it is possible to simulate from the posterior predictive distribution for the "next" sub-computer–sub-computer 49. Figure 3 shows the posterior predictive distribution of ϕ_{49} and η_{49}. Since there is little posterior density for ϕ_{49} greater than 1, this suggests that the "next" sub-computer will undergo reliability growth in the early stages

of its implementation. Also, since the posterior mode for η_{49} is about 0.2, expect the "next" sub-computer to fail for the first time in just under a week.

The goal of this modeling effort was to determine the reliability of the Blue mountain supercomputer. For the ith sub-computer, $N_i(a,b)$ is a Poisson random variable with mean $\mu_i(a,b)$. Thus, the probability of no failures for the ith sub-computer in (a,b) is

$$\begin{aligned} P(N_i(a,b)=0|\phi_i,\eta_i) &= 1-exp(-\mu_i(a,b)) \\ &= 1-exp\left(\left(\frac{a}{\eta_i}\right)^{\phi_i}-\left(\frac{b}{\eta_i}\right)^{\phi_i}\right) \end{aligned}$$

With our definition of reliability as the probability that a job of length l and starting time s finishes, since the Blue mountain is a series system in the 48 sub-computers, reliability $R(l,s)$ is

$$\begin{aligned} R(l,s) &= \prod_{i=1}^{48} P(N_i(s,s+l)=0|\phi_i,\eta_i) \\ &= \prod_{i=1}^{48}\left[1-exp\left(\left(\frac{s}{\eta_i}\right)^{\phi_i}-\left(\frac{s+l}{\eta_i}\right)^{\phi_i}\right)\right]. \end{aligned}$$

Figure 4 is a plot of reliability for 6 hour jobs with different starting times. These times are at the end of months 1, 5, and 10. As the starting time is increased, reliability increases–further evidence of reliability growth.

4.2 A Comparison of Models

Suppose that we have Bayesian models for data Y. In other words, for a given model, we have a joint distribution for the data and parameters. For the ith Bayesian model with parameters $\underline{\theta_i}$, let $(Y,\underline{\theta_i})$ have joint density $p_i(Y,\underline{\theta_i})$. Then, if data $Y=y$ are observed

$$b_i = \int p_i(y,\underline{\theta_i})d\underline{\theta_i}$$

is a measure of how likely the data are under model i. The *Bayes factor* for comparing model i to model j is

$$B_{ij} = \frac{b_i}{b_j}.$$

Bayes factors can be used to compare the fit of two models, where a "large" value of B_{ij} suggests that model i provides a better fit to the data than model j. A nice property of Bayes factors is that they can be used to compare two entirely different models. Getting a corresponding frequentist measure to compare models can be difficult when the models are not nested. Interpretation of a Bayes factor is nevertheless somewhat problematic (i.e., how large is a "large" Bayes factor). Kass and Raftery (1994) provided some guidelines for how large is "large". For example, they argued that a Bayes factor greater than 10 is strong evidence that one model is better than another.

Computing a Bayes factor in closed-form is not always possible. DiCiccio *et al.* (1997) provided ways of approximating Bayes factors and discussed the asymptotic properties of these approximations. To approximate a Bayes factor, we use the Laplace approximation

$$b_i = (2\pi)^{(p/2)}|\hat{\Sigma}|^{\frac{1}{2}}h(\hat{\Theta})$$

where p is the number of parameters in the model, $\hat{\Sigma}$ is the posterior variance-covariance matrix of the parameters, $h(\cdot)$ is the (possibly unnormalized) posterior and $\hat{\Theta}$ is the posterior mean.

Table 3 is a list of Bayesian PLP models. Model 1 is the hierarchical PLP model described in Section 3.2. Model 2 does not allow the shape parameter ϕ to vary over the 48 computers. Model 3 does not allow the scale parameter η to vary over the 48 computers. Model 4 allows neither the scale nor shape parameters to vary over the 48 computers. Model 5 is an HPP that allows the scale parameter to vary over the 48 computers. Model 6 is an HPP that does not

allow the scale parameter to vary over the 48 computers. Table 4 has the Bayes factors for Models 1-6. For example, the Bayes factor to compare the model 1 to model 4 is 17.62. This suggests the need for the hierarchical PLP model over a common ϕ and η model.

5 Conclusions

This paper presented a PLP in a Bayesian hierarchical framework. Prior distributions were defined to facilitate the transfer of qualitative prior information to quantitative prior distributions. The model was used to access the reliability of the Blue mountain supercomputer. The data indicate that the Blue mountain is undergoing reliability growth. Also, since reliability is increasing at a decreasing rate, the supercomputer may be entering a "flat spot". Thus, the supercomputer may be attaining an asymptotic reliability. Models that estimate this limiting reliability is a topic for future research. The Bayesian hierarchical model presented allowed for the estimation of quantities that are important to a computer scientist. Bayes factors were used to show the need for a hierarchical model as opposed to simpler, non-hierarchical Bayesian models.

References

Aven, T. (1989), "Some Tests for Comparing Reliability Growth/Deterioration Rates of Repairable Systems," *IEEE Transactions On Reliability*, 38, 440-443.

Bar-Lev, S.K. Lavi, I. and Reiser, B. (1992), "Bayesian Inference for the Power Law Process," *Annals of the Institute of Statistical Mathematics*, 44, 623-639.

Berman, M. (1981), "Inhomogeneous and Modulated Gamma Processes," *Biometrika*, 68, 143-

152.

Black, S.E. and Rigdon, S.E. (1996), "Statistical Inference for a Modulated Power Law Process," *Journal of Quality Technology*, 28, 81-90.

Calabria, R. Guida, M. and Pulcini, G. (1990), "Bayes Estimation of Prediction Intervals for a Power Law Process," *Communications in Statistics–Theory and Methods*, 19, 3023-3035.

Crow, L.J. (1974), "Reliability Analysis for Complex, Repairable Systems," *Reliability and Biometry*, (eds Proschan, F. and Serfling, R.F.), Philadelphia: Society of Industrial and Applied Mathematics. pp. 379-410.

DiCiccio, T.J. Kass, R.E. Raftery, A. and Wasserman, L. (1997), "Computing Bayes Factors By Combining Simulation and Asymptotic Approximations," *Journal of the American Statistical Association*, 92, 903-915.

Duane, J.T. (1964), "Learning Curve Approach to Reliability Monitoring," *IEEE Transactions on Aerospace*, 2, 563-566.

Finkelstein, J.M. (1976), "Confidence Bounds on the Parameters of the Weibull Process," *Technometrics*, 18, 115-117.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996), *Markov Chain Monte Carlo in Practice*, London: Chapman & Hall.

Guida, M. Calabria R. and Pulcini G. (1989), "Bayes Inference for a Non-Homogeneous Poisson Process with Power Intensity Law," *IEEE Transactions on Reliability*, 38, 603-609.

Hastings, W.K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 57, 97-109.

Higgins, J.J. and Tsokos, C.P. (1981), "A Quasi-Bayes Estimate of the Failure Intensity of a Reliability-Growth Model," *IEEE Transactions on Reliability*, 30, 471-475.

Kass, K.E. and Raftery, A.E. (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90, 773-795.

Kyparisis, J. and Singpurwalla, N.D. (1985), "Bayesian Inference for the Weibull Process with Applications to Assessing Software Reliability Growth and Predicting Software Failures," *Computer Science and Statistics:The Interface*, 57-64.

Lee, L. and Lee, S.K. (1978), "Some Results on Inference for the Weibull Process," *Technometrics*, 20, 41-45.

Littlewood, B. and Verrall, J.L. (1989), "A Bayesian Reliability Growth Model for Computer Software," *Applied Statistics*, 22, 332-346.

Meeker, W.Q. and Escobar, L.A. (1998), *Statistical Methods for Reliability Data*, pp. 393-426. New York: John Wiley & Sons, Inc.

Raftery, A.E. and Lewis, S.M. (1996), "Implementing MCMC," *Markov Chain Monte Carlo In Practice* (eds Gilks, W.R. Richardson, S. and Spiegelhalter, D.J.), New York: Chapman & Hall. pp. 115-130.

Sen, A. (1998), "Estimation of Current Reliability in a Duane-Based Reliability Growth Model," *Technometrics*, 40, 334-344.

Silverman, B.W. (1980), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

A Posterior Simulation Methods

A.1 Markov Chain Monte Carlo (MCMC)

Suppose we are interested in making statistical inference about a parameter (possibly vector valued) Θ. We may have some information (or lack of information) about the distribution of Θ which we will call $\pi(\Theta)$ (prior distribution). Data are collected and represented by the likelihood or $f(\mathbf{x}|\Theta)$. In any Bayesian analysis, inference on the parameters is carried out by calculating the posterior distribution

$$\pi(\Theta|\mathbf{x}) = \frac{\pi(\Theta)f(\mathbf{x}|\Theta)}{\int_{\Theta}\pi(\Theta)f(\mathbf{x}|\Theta)d\Theta}. \tag{3}$$

In many situations, the denominator of (3) is not a well known integral and must be calculated numerically. In cases where the denominator cannot be calculated explicitly, a technique known as Markov Chain Monte Carlo (MCMC) can often be employed. The technique proceeds by letting $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$ be an k dimensional vector, and Θ_{-v} be Θ with the vth element removed. A successive substitution implementation of the MCMC algorithm proceeds as follows:

(1) Initialize $\Theta^{(0)}$ and set $t = 1$.

(2) Set $v = 1$.

(3) Generate an observation $\theta_v^{(t)}$ from the distribution of $[\theta_v|\Theta_{-v}^{(t-1)}]$, replacing recently generated elements of $\Theta_{-v}^{(t-1)}$ with elements of $\Theta_{-v}^{(t)}$ if they have been generated.

(4) Increment v by 1 and go to (3) until $v = k$.

(5) If $v = k$ increment t by 1 go to (2).

As $t \to \infty$ and under conditions outlined in Hastings (1970), the distribution of $\{\theta_1^{(t)}, \dots, \theta_k^{(t)}\}$ tends to the joint posterior distribution of Θ, as desired.

Typical implementation of the algorithm generates an initial "large" number of iterations (called the *burn-in*) until the behavior of the algorithm has stabilized. The burn-in samples are discarded, and the observations generated thereafter are used as observations from the posterior distribution of Θ. Nonparametric density estimators (Silverman 1980) can then be used to approximate the posterior distribution.

A.2 Metropolis-Hastings

Some complete conditional distributions may not be available in closed form. That is, it may be difficult to sample from $[\theta_v|\Theta_{-v}^{(t-1)}] \propto g(\theta_v)$. Obtaining observations from such distributions is facilitated by implementing a *Metropolis-Hastings* step (Hastings 1970) for step (3) in the algorithm above. This is difficult because the distribution is only known up to a constant.

(1) Initialize $\theta_{v_{old}}^{(0)}$ and set $j = 0$.

(2) Generate an observation $\theta_{v_{new}}^{(j)}$ from a *candidate* distribution $q(\theta_{v_{old}}^{(j)}, \theta_{v_{new}}^{(j)})$, where $q(x, y)$ is a probability density in y with mean x.

(3) Generate a uniform (0,1) observation u.

(4) Let

$$\theta_{v_{new}}^{(j+1)} = \begin{cases} \theta_{v_{new}}^{(j)}, & \text{if } u \leq \alpha(\theta_{v_{old}}^{(j)}, \theta_{v_{new}}^{(j)}) \\ \theta_{v_{old}}^{(j)}, & \text{otherwise,} \end{cases}$$

where $\alpha(x, y) = \min\left\{\frac{g(y)q(y,x)}{g(x)q(x,y)}, 1\right\}$.

(5) Increment j and go to (2).

The candidate distribution can be almost any distribution (Gilks et al. 1996), although a symmetric distribution such as the normal results in a simplification of the algorithm, and is called

a *Metropolis step* (as opposed to a Metropolis-Hastings step). A common choice for $q(x, y)$ is a normal distribution with mean x and some variance which allows the random deviates to be a representative sample from the entire complete conditional distribution.

A.3 Implementing MCMC Methods for the Computer Problem

Define $\underline{\theta} = (\mu_T, \sigma_T, \mu_\phi, \sigma_\phi)$. The posterior distribution of the parameters $(\underline{\phi}, \underline{\eta}, \underline{\theta}|X)$ has density

$$\begin{aligned} p(\underline{\phi}, \underline{\eta}, \underline{\theta}|X) &\propto p(X|\underline{\phi}, \underline{\eta}, \underline{\theta})p(\underline{\phi}, \underline{\eta}|\underline{\theta})p(\underline{\theta}) \\ &= p(X|\underline{\phi}, \underline{\eta})p(\underline{\phi}, \underline{\eta}|\underline{\theta})p(\underline{\theta}), \end{aligned}$$

where the final equality follows from the hierarchical model assumption that

$$p(X|\underline{\phi}, \underline{\eta}, \underline{\theta}) = p(X|\underline{\phi}, \underline{\eta}).$$

The conditional distributions needed to implement the MCMC algorithm are

$$p(\phi_i|X,\underline{\phi}_{-i},\underline{\eta},\underline{\theta}) \propto \frac{\prod_{j=1}^{M}\left[\left\{\left(\frac{tj}{\eta_i}\right)^{\phi_i}-\left(\frac{t(j-1)}{\eta_i}\right)^{\phi_i}\right\}^{x_{ij}}\right]\left(\frac{T}{\eta_i}\right)^{\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2}\phi_i^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2}}{exp\left[\left(\frac{Mt}{\eta_i}\right)^{\phi_i}+\frac{\mu_T}{\sigma_T{}^2}\left(\frac{T}{\eta_i}\right)^{\phi_i}+\frac{\mu_\phi}{\sigma_\phi{}^2}\phi_i\right]}$$

$$p(\eta_i|X,\underline{\phi},\underline{\eta}_{-i},\underline{\theta}) \propto \frac{\prod_{j=1}^{M}\left[\left\{\left(\frac{tj}{\eta_i}\right)^{\phi_i}-\left(\frac{t(j-1)}{\eta_i}\right)^{\phi_i}\right\}^{x_{ij}}\right]}{exp\left[\left(\frac{Mt}{\eta_i}\right)^{\phi_i}\right]\eta_i^{\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2+1}exp\left[\frac{\mu_T}{\sigma_T{}^2}\left(\frac{T}{\eta_i}\right)^{\phi_i}\right]}$$

$$p(\mu_T|X,\underline{\phi},\underline{\eta},\underline{\theta}_{-1}) \propto \left[\frac{\left(\frac{\mu_T}{\sigma_T{}^2}\right)^{\left(\frac{\mu_T}{\sigma_T}\right)^2}}{\Gamma\left(\left(\frac{\mu_T}{\sigma_T}\right)^2\right)}\right]^C \frac{T^{\left(\frac{\mu_T}{\sigma_T}\right)^2\sum_{i=1}^{C}\phi_i}\mu_T{}^{a_{\mu_T}-1}}{\prod_{i=1}^{C}\left[\eta_i{}^{\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2}exp\left[\frac{\mu_T}{\sigma_T{}^2}\left(\frac{T}{\eta_i}\right)^{\phi_i}\right]\right]exp\left[\left(\frac{\mu_T}{b_{\mu_T}}\right)^{a_{\mu_T}}\right]}$$

$$p(\sigma_T|X,\underline{\phi},\underline{\eta},\underline{\theta}_{-2}) \propto \left[\frac{\left(\frac{\mu_T}{\sigma_T{}^2}\right)^{\left(\frac{\mu_T}{\sigma_T}\right)^2}}{\Gamma\left(\left(\frac{\mu_T}{\sigma_T}\right)^2\right)}\right]^C \frac{T^{\left(\frac{\mu_T}{\sigma_T}\right)^2\sum_{i=1}^{C}\phi_i}\sigma_T{}^{a_{\sigma_T}-1}}{\prod_{i=1}^{C}\left[\eta_i{}^{\phi_i\left(\frac{\mu_T}{\sigma_T}\right)^2}exp\left[\frac{\mu_T}{\sigma_T{}^2}\left(\frac{T}{\eta_i}\right)^{\phi_i}\right]\right]exp\left[\left(\frac{\sigma_T}{b_{\sigma_T}}\right)^{a_{\sigma_T}}\right]}$$

$$p(\mu_\phi|X,\underline{\phi},\underline{\eta},\underline{\theta}_{-3}) \propto \left[\frac{\left(\frac{\mu_\phi}{\sigma_\phi{}^2}\right)^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2}}{\Gamma\left(\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2\right)}\right]^C \prod_{i=1}^{C}\left[\frac{\phi_i{}^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2}}{exp\left\{\frac{\mu_\phi}{\sigma_\phi{}^2}\phi_i\right\}}\right]\frac{\mu_\phi{}^{a_{\mu_\phi}-1}}{exp\left\{\left(\frac{\mu_\phi}{b_{\mu_\phi}}\right)^{a_{\mu_\phi}}\right\}}$$

$$p(\sigma_\phi|X,\underline{\phi},\underline{\eta},\underline{\theta}_{-4}) \propto \left[\frac{\left(\frac{\mu_\phi}{\sigma_\phi{}^2}\right)^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2}}{\Gamma\left(\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2\right)}\right]^C \prod_{i=1}^{C}\left[\frac{\phi_i{}^{\left(\frac{\mu_\phi}{\sigma_\phi}\right)^2}}{exp\left\{\frac{\mu_\phi}{\sigma_\phi{}^2}\phi_i\right\}}\right]\frac{\sigma_\phi{}^{a_{\sigma_\phi}-1}}{exp\left\{\left(\frac{\sigma_\phi}{b_{\sigma_\phi}}\right)^{a_{\sigma_\phi}}\right\}}.$$

Since none of these conditional distributions have known normalizing constants, parameters can be updated in the MCMC algorithm with one iteration of a Metropolis-Hastings algorithm as described in Sections A.1 and A.2.

B Tables and Figures

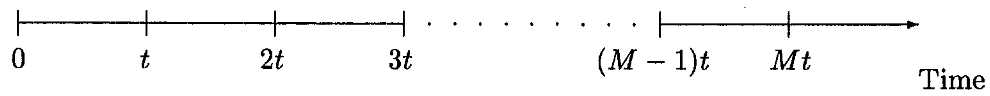

Figure 1: Data Collection Times for the ith Sub-Computer.

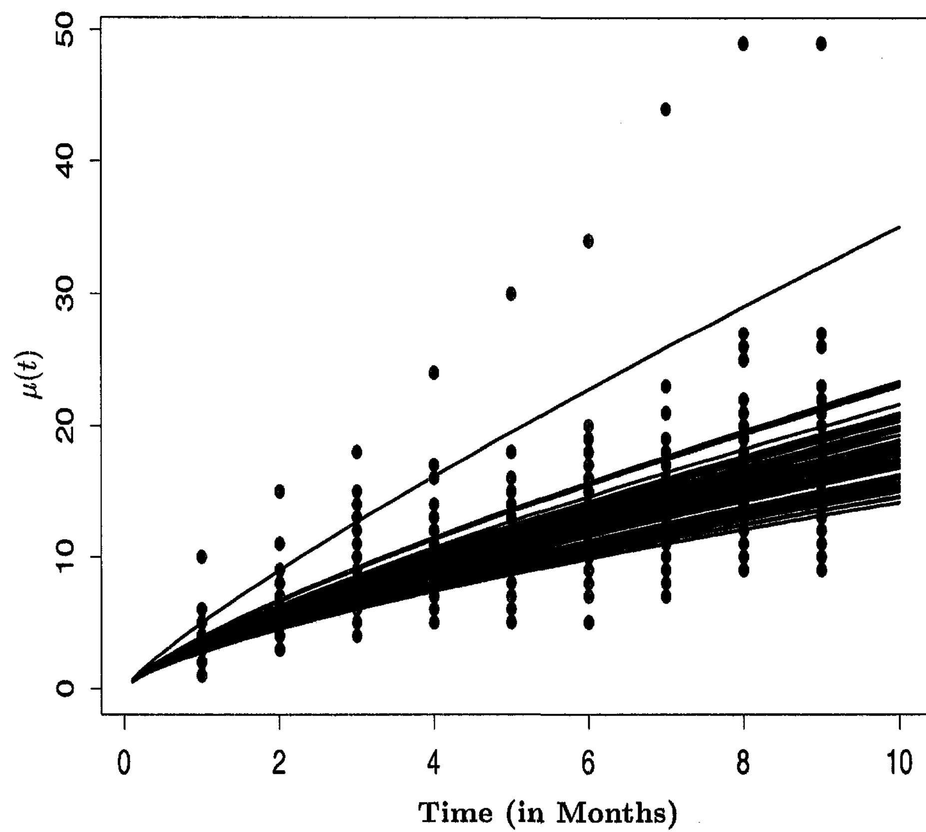

Figure 2: Empirical and Estimated Mean Cumulative Functions.

Table 1: Number of Failures Per Month for Each Sub-computer

	Month								
Sub-computer	1	2	3	4	5	6	7	8	9
1	1	5	2	1	0	1	1	2	0
2	5	4	6	1	2	1	4	4	0
3	1	5	3	0	0	0	0	2	2
4	4	3	2	1	2	1	1	3	0
5	2	2	2	0	1	1	1	4	0
6	3	4	3	1	1	4	1	4	0
7	1	7	3	6	1	2	3	3	0
8	3	3	3	1	0	3	2	4	0
9	6	2	3	1	0	1	0	5	0
10	4	4	5	1	1	1	2	7	1
11	4	7	3	2	0	0	0	4	0
12	4	4	3	3	0	0	1	4	0
13	2	4	3	1	0	0	0	3	0
14	3	4	3	2	0	3	1	2	0
15	2	5	3	3	1	0	1	5	0
16	2	2	3	1	2	1	1	4	1
17	4	3	3	2	0	1	2	3	0
18	5	4	3	0	1	4	1	2	0
19	5	2	3	3	2	4	2	5	0
20	2	3	1	1	0	3	0	1	0
21	10	5	3	6	6	4	10	5	0
22	5	2	5	1	1	1	1	2	0
23	3	3	2	1	1	1	2	5	0
24	2	3	2	1	2	3	0	3	0
25	2	3	2	0	1	1	1	3	1
26	3	2	3	0	2	5	4	1	0
27	1	2	1	2	2	4	3	1	0
28	1	2	2	1	2	2	0	3	1
29	2	3	1	2	2	3	0	1	1
20	5	4	1	1	1	4	0	1	2
31	1	5	1	3	1	4	3	3	0
32	1	3	4	1	3	4	0	2	1
33	4	3	4	0	1	1	1	3	6
34	1	2	1	2	0	1	0	2	0
35	1	3	2	1	2	1	0	2	0
36	2	6	3	1	0	2	0	2	1
37	1	3	1	2	0	0	1	2	0
38	3	2	1	1	2	0	1	2	0
39	2	2	1	0	0	0	2	3	0
40	5	3	2	0	2	2	2	2	0
41	3	3	3	3	4	2	0	4	0
42	5	4	2	2	5	0	0	3	0
43	2	4	3	0	4	2	0	3	3
44	5	2	3	2	3	0	1	3	0
45	1	3	1	1	5	1	0	3	0
46	1	2	2	0	2	1	0	3	0
47	5	2	2	1	1	1	0	2	0
48	2	2	1	0	2	0	1	2	0

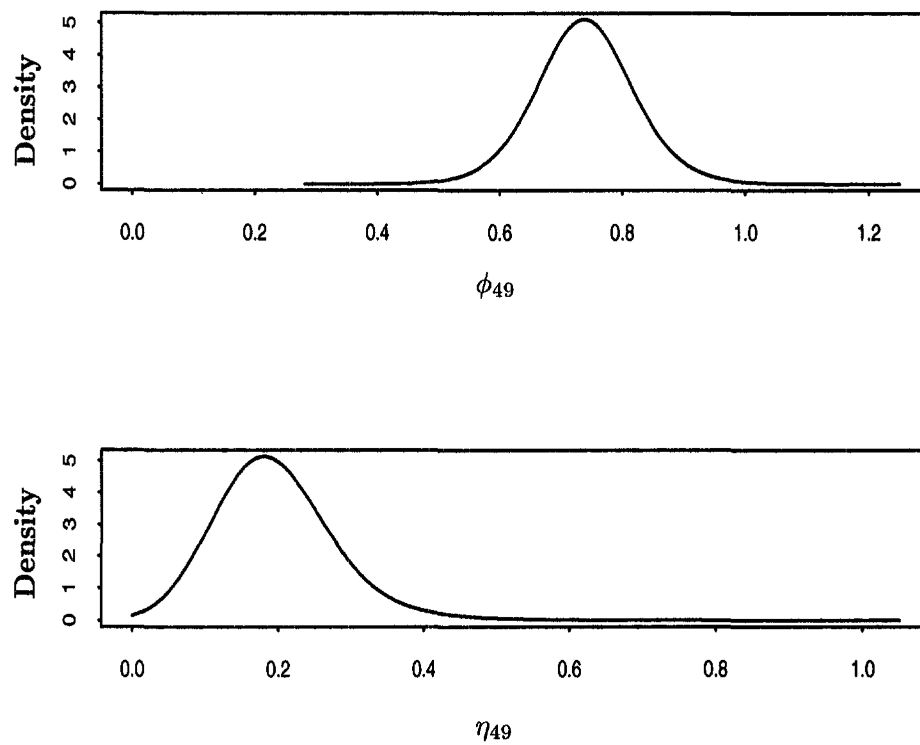

Figure 3: Posterior Predictive Distributions for Computer 49.

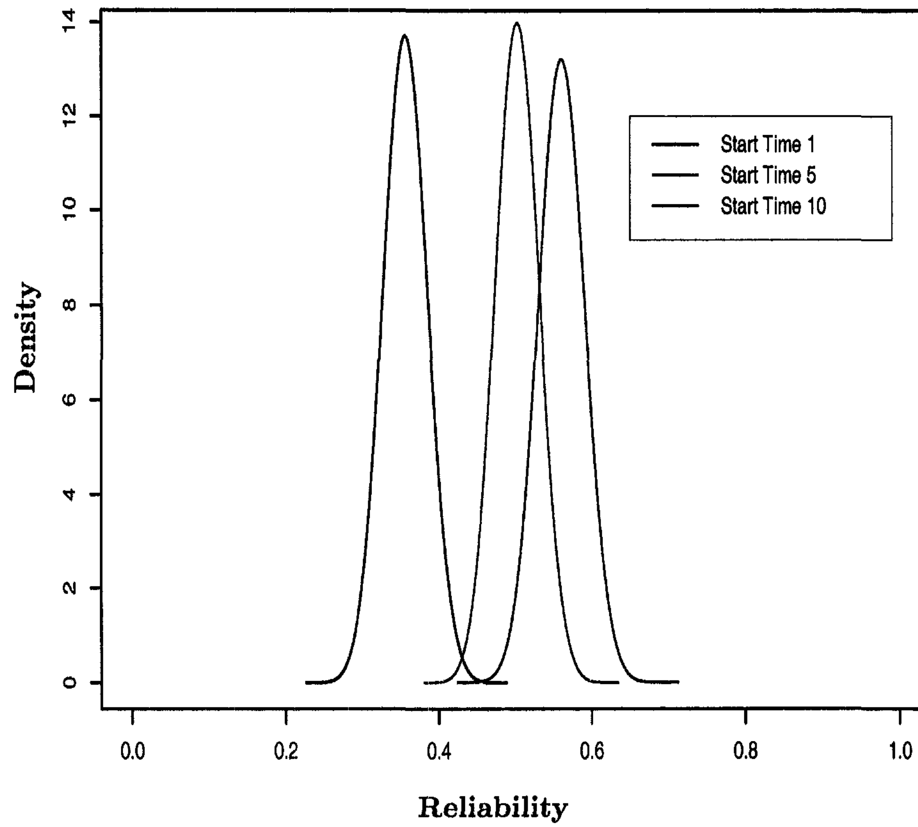

Figure 4: Reliability Posterior Predictive Distributions for 6 Hour Jobs.

Table 2: Hierarchical Bayes Parameter Estimates

Parameter	Mean	Median	Standard Error	90% HPD Interval
ϕ_1	0.721	0.724	0.057	$(0.628, 0.814)$
η_1	0.211	0.201	0.070	$(0.106, 0.321)$
μ_T	3.509	3.499	0.259	$(3.088, 3.931)$
σ_T	0.647	0.663	0.247	$(0.229, 1.059)$
μ_ϕ	0.739	0.740	0.029	$(0.691, 0.786)$
σ_ϕ	0.055	0.054	0.026	$(0.011, 0.094)$

Table 3: A List of Bayesian PLP Models

Model Number	ϕ_i	η_i
1	ϕ_i	η_i
2	ϕ	η_i
3	ϕ_i	η
4	ϕ	η
5	1	η_i
6	1	η

Table 4: Bayes Factors for the Bayesian PLP Models

		j					
		1	2	3	4	5	6
	1	1.00	NA	NA	17.62	6.07e+12	1.31e+15
	2		1.00	NA	NA	NA	NA
i	3			1.00	NA	NA	NA
	4				1.00	3.44e+11	7.46e+13
	5					1.00	216.72
	6						1.00