BAYESIAN RELIABILITY MODEL FOR BEAM PERMIT SYSTEM OF RHIC AT BNL*

P. Chitnis#, Stony Brook University, NY, USA
K. A. Brown, Brookhaven National Laboratory, NY, USA

Abstract

Bayesian Analysis provides a statistical framework for updating prior knowledge as observational evidence is acquired. It can handle complex and realistic models with flexibility. The Beam Permit System (BPS) of RHIC plays a key role in safeguarding against the faults occurring in the collider, hence directly impacts RHIC availability. Earlier a multistate reliability model [1] was developed to study the failure characteristics of the BPS that incorporated manufacturer and military handbook data. Over the course of its 15 years of operation, RHIC has brought forth operational failure data. This work aims towards the integration of earlier reliability calculations with operational failure data using Bayesian analysis. This paper discusses the Bayesian inference of the BPS reliability using a two-parameter Weibull survival model, with unknown scale and shape parameters. As the joint posterior distribution for Weibull with both parameters unknown is analytically intractable, the Markov Chain Monte Carlo methodology with Metropolis-Hastings algorithm is used to obtain the inference. Selection criteria for the Weibull distribution, prior density and hyperparameters are also discussed.

INTRODUCTION

The Beam Permit System (BPS) of Relativistic Heavy Ion Collider (RHIC) monitors the health of RHIC subsystems and takes active decisions regarding beam-abort and magnet power dump, upon a subsystem failure. The reliability of BPS thus directly impacts the RHIC downtime, and hence its availability. A Monte Carlo reliability model based on exponential competing risks was developed for BPS previously [1]. This model simulated the progression of basic component failures to system level catastrophic events. This work together with a quantitative fault tree analysis [2] helped characterize the failure rate and structural importance of each basic component of the BPS. RHIC has been operational for 15 years, and has gathered hardware failure data over this time period. This data represents the actual BPS failure attributes from a top level perspective. Bayesian analysis is a good candidate for combining these two information sources to get a combined inference about the system failure characteristics.

BAYESIAN PARADIGM

Bayesian statistics is branch of mathematics that deals with updating current knowledge about a system or process when new information is acquired. Statistical analysis follows two major approaches, namely frequentist and Bayesian. In the widely used frequentist approach, the probability distribution of an event is calculated by observing its occurrence over a large period of time, and the distribution parameters are assumed to be constant over time. In contrast, Bayesian approach keeps updating the probability distribution as new data arrives. The parameters of the distribution are treated as random variables that are modified according to the new information gathered. This becomes quite important when there are two sources of information about a system or process that indicate different results, and both sources hold significance to the inference.

The underlying framework for Bayesian analysis is Bayes theorem. Bayesian analysis involves the continuous form of the Bayes theorem [3], which is represented as

$$\pi(\theta|x) = \frac{L(\theta|x) \times \pi(\theta)}{f(x)}$$

The unknown parameter is θ, which defines the probability distribution of any process and is subject to change with the arrival of new information. Variable x is new source of information in the form of data observations. The term $\pi(\theta)$ is called as the prior distribution of θ, which can be elicited by using another parameter(s) called the *hyperparameter(s)*. $L(\theta|x)$ is the likelihood function for θ which is calculated by the gathering the new data. The term $\pi(\theta|x)$ is called the posterior distribution of θ, which is a combination of both prior and data likelihood function. $f(x)$ is the unconditional distribution of the variate x that acts as a normalizing factor in the equation. Because $f(x)$ is independent of θ:

$$\pi(\theta|x) \propto L(\theta|x) \times \pi(\theta) \qquad (1)$$

This equation forms the foundation for Bayesian analysis discussed in this paper. We will discuss the selection of the prior distribution and data distribution (likelihood function) in subsequent sections.

PRELIMINARY ANALYSIS

Preliminary analysis is needed to find the suitability of Bayesian analysis to our problem and for choosing the distribution appropriate to the information sources we have. We analyze two the sources of information, the results from a Monte Carlo (MC) model [1] and the historical failure data obtained from the RHIC hardware maintenance records.

*Work supported by Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy.
#prachi.chitnis@stonybrook.edu

ISBN 978-3-95450-148-9

Source 1: Monte Carlo Results

The MC model defines the propagation of component failures to a system level, depending upon the states of other components and the structure of the system. As the component failures were taken from MIL-HDBK 217F [4] and the manufacturers' data, they provided point estimate for the hazard rate λ, with exponential survival distribution. These exponential failures might evolve as a different distribution on the system level. We combine the four system level failures modes in the MC model to a single failure for applying Bayesian analysis. Next we analyze the failure rate pattern of this single failure of BPS.

First, we check if the failure rate follows the Non Homogenous Poisson Process (NHPP) [5], for which specialized Bayesian analysis is needed. We plot the interval failure rates to see if it is time varying, and whether it follows a typical distribution. The number of failures in an interval of 10^7 hours is calculated, and the failure rate is plotted for each interval in Fig. 1. The rate is noisy but has a constant mean over the time. Thus we deduce that the MC model failure is not a NHPP.

Figure 1: Detection for NHPP failures.

Next, we find a suitable failure distribution function for the MC model. We run the simulation for 1.3E9 iterations and the total failures are recorded as point processes. The times between failures are fitted with exponential, Weibull and gamma distributions with the forms specified in [6] using MATLAB® [7]. The goodness of fit is estimated by Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [8].

Table 1: MC Model Failure Distribution

Distribution	Para meter	Point estimate	AIC	BIC
Exponential	λ	8.831e-5	3141125.52	3141136.15
	α	1		
Weibull	λ	8.829e-5	3141127.47	3141148.72
	α	1.00046		
Gamma	λ	8.84e-5	3141127.44	3141148.66
	α	1.00106		

As seen in table 1, the AIC and BIC are smallest for the exponential distribution, which indicates that the exponential distribution is the best fit to overall system failure. The failure rate is λ and the shape parameter is α.

Although the military handbook is quite old (1995), its applicability to the BPS can justified as RHIC has been running since 1997, which was contemporary to the release of this version of military handbook

Source 2: Historical Failure Data

We analyze the past 15 years of hardware failure data of BPS, and select the system level failures that are similar to the ones analyzed by the MC model. We find overall 16 data points for the time between failures owing to high reliability of BPS. To analyze the distribution of this data, we fit Exponential, Weibull, Gamma and Lognormal distributions using MATLAB® and goodness of fit is estimated using AIC and BIC.

Table 2: Historical Failure Data Distribution

Distribution	Para meter	Point estimate	AIC	BIC
Exponential	λ	1.20E-04	322.9279	323.7005
	α	1		
Weibull	λ	0.000171	317.6438	319.189
	α	0.627457		
Gamma	λ	6.03E-05	318.1511	319.6963
	α	0.503074		
Lognormal	μ	7.7676	319.0121	320.5573
	σ	1.99141		

Looking at the table 2, we see that the AIC and BIC are now smallest for the Weibull distribution, asserting that the BPS has a Weibull survival distribution with decreasing failure rate. The historical data distribution represents the actual failure characteristics of the system, even if the MC model suggests that the system follows exponential survival distribution.

BAYESIAN RELIABILITY MODEL

The parameters of the posterior distribution reflect the tradeoff between the prior distribution and the data distribution (incorporated using likelihood function). This tradeoff level is determined by the relative strength of prior and data distributions. The influence of either can be changed by altering the hyperparameters. It is often desirable to choose a prior of a form such that the posterior distribution calculated is mathematically tractable. One of the techniques is to employ a "conjugate prior" that yields a posterior of the same form as the data distribution [3], but with different parameters. The posterior parameters specify the adjustment between prior and data.

For our Bayesian model, the prior information is an exponential distribution and the data is a Weibull distribution with shape parameter less than 1. We thus assume that the prior information is also a Weibull distribution with shape parameter equal to 1. Also scale parameters of both the information sources is different. Thus we need to choose a Bayesian model for Weibull distribution with both shape and scale parameters unknown. To implement this, we first need to choose a conjugate prior distribution that is suitable to the Weibull distribution of the data. There is no best way to define a prior distribution for Bayesian analysis. The following

ISBN 978-3-95450-148-9

sections outine the development of the Bayesian model step by step.

Data Distribution

Vidakovic [9] suggests a model for the Bayesian inference of Weibull distribution that is unknown in both shape and scale parameters. We follow this model throughout our analysis. Following is the Weibull distribution used by Vidakovic, with α as the shape parameter and $\eta^{-1/\alpha}$ as the scale parameter. Variable x is the times between failures data in years. This form does not have explicit posteriors for α and η. We will draw Bayesian inference for these two parameters.

$$f(x|\alpha,\eta) = \alpha\eta x^{\alpha-1}e^{-\eta x^{\alpha}}$$

The likelihood function is then equal to

$$L(\alpha,\eta|x) = \prod_{i}^{k} \alpha\eta x_i^{\alpha-1}e^{-\eta x_i^{\alpha}}$$

The α and η are treated as variables in the likelihood function, and have *[0, ∞)* support. The most probable values are given by

$$\alpha = 0.6275, \eta = 1.2904 \tag{2}$$

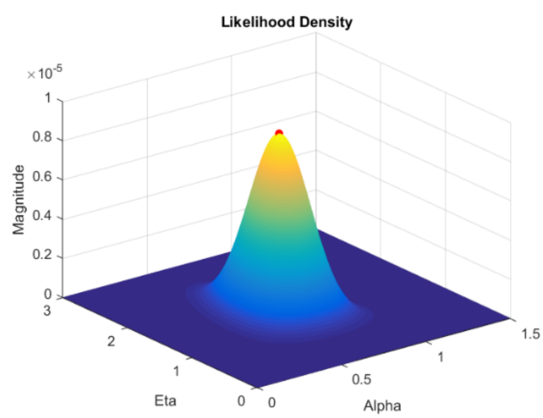

Figure 2: Likelihood function plot with α, η.

These values are calculated by transforming the parameters from the Weibull distribution of data from table 2. Fig. 2 plots the 3D likelihood function with α and η as variables. Note that α and η values from Eq. 2 correspond to the maximum likelihood point.

Conjugate Prior Distribution

A conjugate prior distribution is proposed for the Weibull distribution in [9]. This is a joint distribution for α and η, with a hyperparameter β.

$$p(\alpha,\eta) \propto e^{-\alpha}\eta^{\beta-1}e^{-\eta\beta} \tag{3}$$

Note that we only use the kernel (omitting the proportionality constant) of the prior distribution. This is explained later in the posterior inference. The prior parameter λ (and α=1) representing the exponential distribution (Weibull with shape as 1) in table 1 are converted to α and η as

$$\alpha = 1, \eta = 0.7741 \tag{4}$$

These are the point estimates. We need to define the joint distribution of α and η so that it best represents our beliefs about the prior information. We choose the hyperparameter β equal to 3. The reason for choosing β is explained by the figures below. The 3D prior density from Eq. 3 is plotted with α and η as variables in Fig. 3. Note that the magnitude on the plot is unnormalized. The point estimates from Eq. 4 are plotted as a red dot on the same figure. Note that this point does not correspond to the peak of magnitude in the plot.

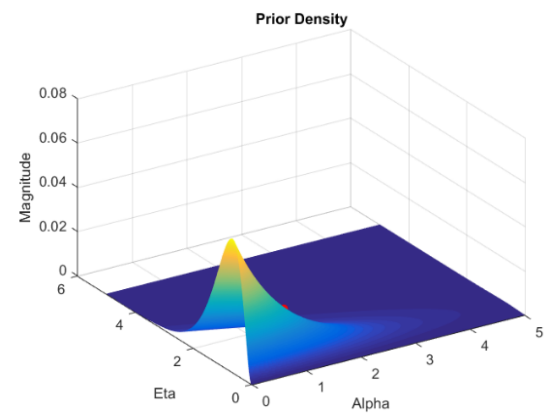

Figure 3: Prior density plot with α, η.

To get a clear picture we plot the 2D prior density for η with constant $\alpha = 1$ in Fig. 4a and the 2D prior density for α with constant $\eta = 0.7741$ in Fig. 4b.

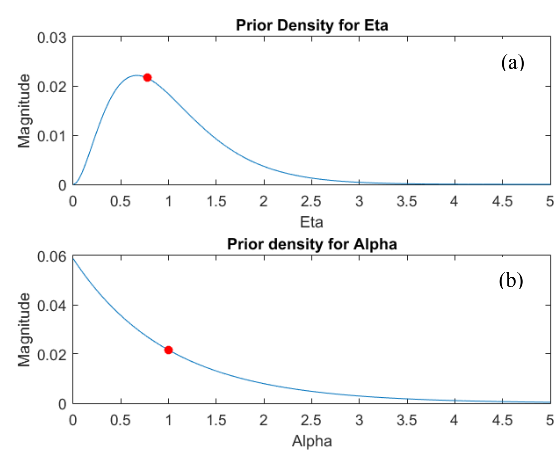

Figure 4: 2D Prior densities for α, η.

In Fig. 4a, the point estimate lies almost on the peak of the density distribution which signifies our confidence in the value of scale parameter of the prior obtained from the MC model. But if we look at Fig. 4b, the point estimate

ISBN 978-3-95450-148-9

does not lie on the peak, rather much lower than the highest magnitude point. This is chosen because the shape parameter is a system characteristic, which is less likely to change. For the data distribution we saw that α is much less than one in Eq. 2. So we express more confidence in the lower values of α than the one obtained from the MC model. Thus the prior density is higher for smaller values of α in Fig. 4b.

Posterior Inference

The posterior distribution is a fusion of the prior and data distributions that contains all the information of the parameters of the system, in our case α and η. From Eq. 1 we get the proportionality equation (or kernel) for posterior as:

$$p(\alpha,\eta|x) \propto \alpha^k \eta^{k+\beta-1}\left(\prod_i^k x_i\right)^{\alpha-1} e^{(-\eta \sum x_i^\alpha - \alpha - \eta\beta)}$$

Here k is the total number of data points. As this is a complicated joint distribution of α and η, it is not possible to obtain independent and identically distributed samples directly from this unnormalized kernel. We use the Metropolis Hastings (MH) algorithm which is a type of iterative Markov Chain Monte Carlo (MCMC) technique [3]. The parameters α and η are calculated as sample averages of realizations of Markov chains, so one has to ensure that the Markov chain has converged before drawing the samples. To generate the random samples of α and η, a "proposal density" is used, given the samples from previous iteration. Following proposal density is suggested in [9].

$$q(\alpha',\eta'|\alpha,\eta) = \frac{1}{\alpha\eta} e^{\left(\frac{\alpha'}{\alpha} - \frac{\eta'}{\eta}\right)}$$

Here α',η' are the new samples and α,η are the previous samples. In MH algorithm first we draw samples from the proposal density. This sample is then accepted or rejected as per the acceptance probability given by:

$$a\big((\alpha',\eta'),(\alpha,\eta)\big) = min\left\{1, \frac{p(\alpha',\eta')/q(\alpha',\eta'|\alpha,\eta)}{p(\alpha,\eta)/q(\alpha,\eta|\alpha',\eta')}\right\}$$

There is a typical advantage of MH algorithm that we need not consider the full conditionals because the normalizing factors cancel in the ratio of acceptance probability equation. For more details on the MH algorithm please refer to [3] and [9].

RESULTS

After running 25K iterations of the MH algorithm on the posterior density, we reject the initial 5000 samples to allow the Markov chain convergence. Next we plot the remaining samples to get a 3D plot of the posterior. Fig. 5 shows the samples (green dots) obtained from the MH algorithm, and a connecting surface is plotted.

To obtain the values of α and η from the posterior, we look at the samples obtained from the MH algorithm. Fig. 6 and Fig. 7 show the samples of α and η and their histograms.

As seen, the subsequent sample generation looks stationary [5], thus we can say that the Markov chains have converged. Also looking at the histograms in Fig. 7, they resemble the normal distribution, thus the means of the samples represent the meaningful inference for α and η parameters.

We obtain the following values of α and η for the posterior:

$$\alpha = 0.6327, \eta = 1.2225 \tag{5}$$

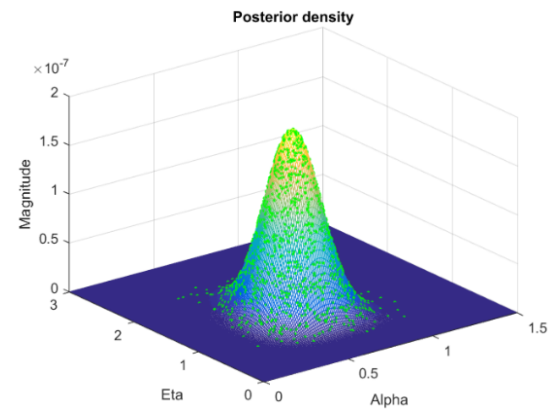

Figure 5: Posterior density for α, η.

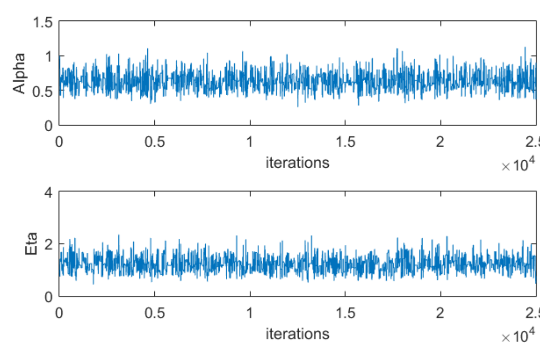

Figure 6: α,η samples from MH algorithm.

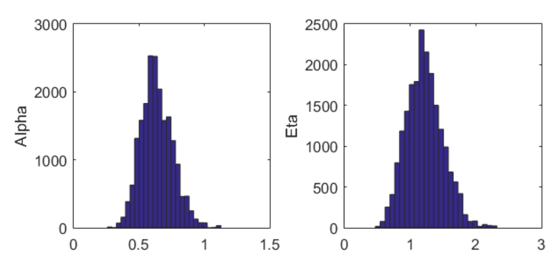

Figure 7: α,η samples' histogram from MH algorithm.

ISBN 978-3-95450-148-9

DISCUSSION

Fig. 8 shows the cumulative failure distribution function of the prior, data and posterior distributions for Weibull using the parameters in Eq. 2, Eq. 4 and Eq. 5. According to our discussion on choosing the hyperparameter β, we expressed low confidence in the value of alpha being 1. This can be seen in Fig. 8 where the posterior shape is more like the data distribution, i.e. the relative strength of the data distribution is much more than the prior distribution.

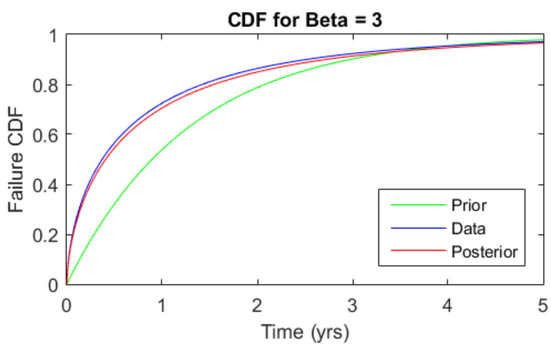

Figure 8: Cumulative failure distributions for β=3.

To illustrate the concept of relative strength of prior and data, we increase the confidence in prior by increasing the hyperparameter value. Fig. 9 shows the prior, data and posterior distribution for β equal to 15.

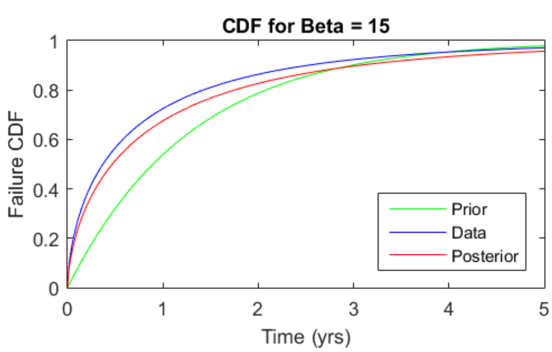

Figure 9: Cumulative failure distributions for β=15.

The prior strength is now increased that is apparent on the posterior, which is now closer to the prior as compared to Fig. 8. The posterior parameters for $\beta = 15$ are a higher value of $\alpha = 0.6404$ and lower value of $\eta = 1.1249$. For our analysis, we uphold the value of $\beta = 3$, because it represents our high confidence in the actual machine failure data, with a mild influence of the MC model results.

CONCLUSION

RHIC beam permit system has been extensively studied for its reliability characteristics. The MC model provides many insights to the reliability performance of the BPS. This includes the marginal probability values of various system level catastrophic events, marginal probabilities of failure modes of individual modules, importance of each component with respect to its failure rate and structural placement, paths of failure propagation and bottlenecks in the system. This helped understand very fine failure dynamics of the beam permit system. However it uses the military handbook which is quite conservative in its approach.

On the other hand, the historical failure data of BPS provides us with the actual failure aspects of the system. This helps to quantify the overall system failure distribution that emerged as a Weibull failure distribution with decreasing failure function. It represents the real survival behavior of the BPS. However due to high the reliability of BPS, we have a small data of only 16 failures points till date, which does not allow us to take a profound look into the system.

Thus it is necessary to emphasize the importance of both the information sources. Bayesian paradigm facilitates an excellent way to coalesce these two to furnish the most informed inference [10] about the BPS reliability, with flexibility to regulate the influence of either of the information sources according our confidence in them.

ACKNOWLEDGEMENT

This work would not have been possible without the guidance of S. T. Rachev, who helped built strong fundamentals in Bayesian statistical framework. We thank him for his continued support. We also thank T. G. Robertazzi for building basics in probability and statistics. Additionally we thank R. Schoenfeld and C. Theisen for help in mining the historical failure data.

REFERENCES

[1] P. Chitnis et al., MOPPC075, ICALEPCS 2013, San Francisco.
[2] P. Chitnis et al., MOPPC076, ICALEPCS 2013, San Francisco.
[3] S. T. Rachev et al., *Bayesian Methods in Finance*, 2008, Wiley.
[4] MIL-HDBK-217F, *Military Handbook-Reliability Prediction of Electronic Equipment*, Department of Defense, 1995.
[5] S. M. Ross, *Introduction to Probability Models*, 10th ed., 2010, Elsevier.
[6] L. M. Leemis, *Reliability, Probabilistic Models and Statistical Methods*, 2nd edition, 2009.
[7] MATLAB® Release 2015a, The MathWorks Inc., Natick, MA, 2015.
[8] K.P. Burnham et al., *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*, 2nd ed., 2002, Springer.
[9] B. Vidakovic, "MCMC methodology", http:// www2.isye.gatech.edu/~brani/isyebayes/bank/handout10.pdf
[10] K.A. Brown et al., WEPGF134, these proceedings, ICALEPCS'15, Melbourne, Australia(2015).

ISBN 978-3-95450-148-9