AIP Conference Proceedings

A hierarchical Weibull Bayesian model for series and parallel systems

F. L. Bhering, A. Polpo, and C. A. De B. Pereira

Citation: AIP Conf. Proc. **1490**, 59 (2012); doi: 10.1063/1.4759589
View online: http://dx.doi.org/10.1063/1.4759589
View Table of Contents: http://proceedings.aip.org/dbt/dbt.jsp?KEY=APCPCS&Volume=1490&Issue=1
Published by the American Institute of Physics.

A Hierarchical Weibull Bayesian Model for Series and Parallel Systems

Bhering, F.L.[*], Polpo, A.[†] and Pereira, C.A. de B.[*]

[*]*University of São Paulo*
[†]*Federal University of São Carlos*

Abstract. In this paper we present a hierarchical Bayesian approach to the estimation of component's reliability in a series and parallel systems using the Weibull model. The reliability problem of a series system is similar to the survival problem of right-censored data, while the parallel system is related to left-censored data. We used the Weibull model for the reliability time and a gamma distribution for first level on hierarchy for both, scale and shape, parameters of the model. The estimation is done using Monte Carlo Markov Chain tools and Expectation-Maximization algorithm. To exemplify the efficiency of the model we presented an study with simulated data.

Keywords: Series system, parallel system, Weibull model, hierarchical model
PACS: 2.50.-r, 02.70.Rr

1. INTRODUCTION

For the problem of estimate the reliability of series system or competing risk, there are many works in literature. A very important paper for this problem is the Kaplan and Meier [2], with developed a non-parametric estimator using a frequentist approach, and Polpo and Sinha [5] is his Bayesian counterpart. For the case o parallel system, for the best of our knowledge, Polpo and Pereira [4] was the first to address the estimation problem, using a non-parametric estimator under Bayesian paradigm. For parametric models, Polpo et al. [6] presented the estimation of Weibull model under Bayesian approach with a non-informative prior. However, in our analysis this procedure showed to be hard to estimate. The convergence of Monte Carlo Markov Chain (MCMC) algorithms are not easy to be achieved, and thus the estimation problem is a very hard task. In this work we suggest the use of a non-informative prior too, but considering one hierarchical level, to have a more robust model and less problems in estimation than those find in the Polpo et al. [6]. Also, we provide an easy way to find credible bound for the reliability function.

A series system is a arrangement of components which works only if all components are active, that is, whenever one fails the system fails too. On the other hand, for the parallel system fail, all components must fail.

Rodrigues et al. [7] performed a simulation study of three different methods to estimate the reliability data. They compared the Kaplan-Meier estimator [2], maximum likelihood estimator (MLE) and the Bayesian plug-in estimator (BPE) for parametric Weibull model. Their results indicated that both MLE and BPE performed very similar and the Kaplan-Meier was the inferior estimator. However they did not addressed the question of credible bounds for the reliability function, which is one of our objective.

The estimation of reliability function was made as follows: (i) we draw a sample

XI Brazilian Meeting on Bayesian Statistics
AIP Conf. Proc. 1490, 59-66 (2012); doi: 10.1063/1.4759589

from the posterior distribution of the Weibull parameters; (ii) using the appropriate transformation, we evaluated a sample from the posterior of reliability function; (iii) for each reliability time, we evaluate the posterior mean, based in the posterior distribution of the reliability function, and use as point estimate. The high posteriori density (HPD) procedure was used to evaluate the credible region for the reliability function for each point. We call the reader attention that we are not using the plug-in estimator in this work, we are evaluating the posterior mean of the reliability function based in the Weibull model. Hence, this procedure is superior, in our opinion, than those used by Rodrigues et al. [7], and also simplifying the problem of estimating credible bounds. A comparative study of both procedures is a subject for a future work and will not be treat here.

This paper is organized as follows. In Section 2, we describe all functions involved in the estimation procedure. In Section 3, a numerical example is analyzed to illustrate the procedure presented in this paper. Some consideration and comments are given in Section 4.

2. THE MODEL

We will use the same notation as in Polpo et al. [6]. Consider a system of k components and let $(X_j)_{j\in\{1,\ldots k\}}$ the sequence of failures times of all components. We assume that this sequence is composed of independently random variables with Weibull distribution. Recall that we only observe a random vector of two variables, namely, (T,δ) with $T = \min(X_1,\ldots,X_k)$ for the series system and $T = max(X_1,\ldots,X_k)$ for the parallel system, and $\delta = j$ if $T = X_j$, for $j = 1,\ldots,k$. The δ quantity can be viewed as an indicator function of the component that caused the system failure.

Consider a sample (size n) of independent and identically distributed systems (all series or parallels). The observations are represented by $(T,\delta) = \{(T_i,\delta_i) : i = 1,\ldots,n\}$. Also, the reliability function of $j-th$ component is given by $R_j(t) = P(X_j > t)$, $j = 1,\ldots,k$. Therefore, for the whole series system the reliability function is $R(t) = \prod_{j=1}^k R_j(t)$ and for the parallel system is $R(t) = 1-\prod_{j=1}^k(1-R_j(t))$.

2.1. Likelihood, Priors, Posterioris

We define a random variable X with Weibull distribution and parameters $\theta = (\beta,\eta)$, that is,

$$P(X > x|\theta) = R(x|\theta) = \exp\left\{-\left(\frac{x}{\eta}\right)^{\beta}\right\} \tag{1}$$

for $x > 0$, $\beta > 0$ (shape) and $\eta > 0$ (scale).

Then, the likelihood function for the series system is given by

$$L(\boldsymbol{\theta}|t,\delta) \propto \prod_{j=1}^{k}\prod_{i=1}^{n}[f_j(t_i|\theta_j)]^{I_{\{\delta_i=j\}}}[R_j(t_i|\theta_j)]^{1-I_{\{\delta_i=j\}}} \tag{2}$$

and for the parallel system,

$$L(\boldsymbol{\theta}|t,\delta) \propto \prod_{j=1}^{k}\prod_{i=1}^{n}[f_j(t_i|\theta_j)]^{I_{\{\delta_i=j\}}}[1-R_j(t_i|\theta_j)]^{1-I_{\{\delta_i=j\}}} \tag{3}$$

where f is the density function of a random variable with Weibull distribution, $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_k)$, $\theta_j = (\beta_j,\eta_j)$, for $j=1,\ldots,k$ and I is the indicator function.

The priors distributions were considered independent with $\beta_j \sim$ gamma$(m_{\beta_j}, v_{\beta_j})$, $\eta_j \sim$ gamma(m_{η_j}, v_{η_j}), $\pi(m_{\beta_j}) \propto \pi(m_{\eta_j}) \propto 1$, and v_{β_j} and v_{η_j} are known constants, $j=1,\ldots,k$. Then

$$\begin{aligned}
\pi(\boldsymbol{\theta}) &\propto \pi(\boldsymbol{\theta}|\boldsymbol{m_\beta},\boldsymbol{v_\beta},\boldsymbol{m_\eta},\boldsymbol{v_\eta})\pi(\boldsymbol{m_\beta})\pi(\boldsymbol{v_\beta})\pi(\boldsymbol{m_\eta})\pi(\boldsymbol{v_\eta}) \\
&\propto \prod_{j=1}^{k}\pi(\theta_j|m_{\beta_j},v_{\beta_j},m_{\eta_j},v_{\eta_j})\pi(m_{\beta_j})\pi(v_{\beta_j})\pi(m_{\eta_j})\pi(v_{\eta_j}) \\
&\propto \prod_{j=1}^{k}\frac{\beta_j^{m^2_{\beta_j}/v_{\beta_j}-1}\exp\{-m_{\beta_j}\beta_j/v_{\beta_j}\}}{(v_{\beta_j}/m_{\beta_j})^{m^2_{\beta_j}/v_{\beta_j}}\Gamma(m^2_{\beta_j}/v_{\beta_j})}\,\frac{\eta_j^{m^2_{\eta_j}/v_{\eta_j}-1}\exp\{-m_{\eta_j}\eta_j/v_{\eta_j}\}}{(v_{\eta_j}/m_{\eta_j})^{m^2_{\eta_j}/v_{\eta_j}}\Gamma(m^2_{\eta_j}/v_{\eta_j})}
\end{aligned}$$

where $\boldsymbol{m_\beta} = (m_{\beta_1},\ldots,m_{\beta_k})$ and $\boldsymbol{m_\eta} = (m_{\eta_1},\ldots,m_{\eta_k})$ are the prior mean parameters, $\boldsymbol{v_\beta} = (v_{\beta_1},\ldots,v_{\beta_k})$ and $\boldsymbol{v_\eta} = (v_{\eta_1},\ldots,v_{\eta_k})$ are the variance (precision) prior parameters, and $m_{\beta_j}, m_{\eta_j}, v_{\beta_j}, v_{\eta_j} > 0$, $j=1,\ldots,k$.

In this case, we have that the posterior distributions of series and parallel systems are, respectively,

$$\pi(\boldsymbol{\theta}|\boldsymbol{t},\boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta})\prod_{i=1}^{n}\left[\frac{t_i^{\beta_j-1}\beta_j}{\eta_j}\exp\left\{-\left(\frac{t_i}{\eta_j}\right)^{\beta_j}\right\}\right]^{I_{\{\delta_i=j\}}}\left[\exp\left\{-\left(\frac{t_i}{\eta_j}\right)^{\beta_j}\right\}\right]^{1-I_{\{\delta_i=j\}}},$$

and

$$\pi(\boldsymbol{\theta}|\boldsymbol{t},\boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta})\prod_{i=1}^{n}\left[\frac{t_i^{\beta_j-1}\beta_j}{\eta_j}\exp\left\{-\left(\frac{t_i}{\eta_j}\right)^{\beta_j}\right\}\right]^{I_{\{\delta_i=j\}}}\left[1-\exp\left\{-\left(\frac{t_i}{\eta_j}\right)^{\beta_j}\right\}\right]^{1-I_{\{\delta_i=j\}}},$$

where $\boldsymbol{t} = (t_1,\ldots,t_n)$ are the observed failure time of the system, $\boldsymbol{\delta} = (\delta_1,\ldots,\delta_n)$ are the indicators which component fail, and the others quantities are as defined before.

2.2. Estimation

Arguably, we can use the mode of the integrated likelihood of $(\boldsymbol{m_\beta},\boldsymbol{m_\eta})$ to determine a prior distribution [1]. The EM algorithm [3] can be used to obtain the maximum a posterior estimates of $\boldsymbol{m_\beta}$ and $\boldsymbol{m_\beta}$. The MCMC procedure is used to generate a sample from the posterior distribution of $\boldsymbol{\theta}$. We are omitting the details about these computational tools because they are already well known tools and are not the main subject of the present paper.

3. NUMERICAL EXAMPLE

Consider that X_1 has Weibull distribution with mean 2 and variance 2, X_2 has gamma distribution with mean 2 and variance 0.816, and X_3 has log-normal distribution with mean 2.014 and variance 2.639. We generated a sample (with size $n = 100$) of a series system with these three components and another sample (with size $n = 100$) of a parallel system with this same three components. The components was choose to have similar means but different variances and, moreover, different distributions. We use the same theoretical components in both simulation (series and parallel systems) to verify in each situation the difference of the proposed model and the available data. Note that, our interest is in the estimation of the reliability function of each component, and with our simulated example we have a huge amount of censored data. Then, this is a challenge example.

For the estimation procedures we used a MCMC tool and we the convergence of the algorithm was very "fast". We only needed to discard the first 10,000 samples from the posterior to achieve the stationary measure and then generate a sample from the posterior. To perform the estimation of the reliability functions and the credible intervals we used a sample with size $1,000$ from the posterior. To verify is the estimation was reasonable, we compared the "true" reliability of each component with the estimated reliability function. Table 1 present the posterior mean and the posterior standard deviation of each parameter involved in the model, for both systems. We notice that the standard deviations are relatively low, indicating the estimates are good.

TABLE 1. Summary of estimated parameters

	Series system		Parallel system	
	Mean	Std. Deviation	Mean	Std. Deviation
β_1	1.4646	0.1591	1.3554	0.1383
η_1	2.0999	0.2322	2.7416	0.2448
β_2	3.0913	0.3795	2.4724	0.2898
η_2	2.1567	0.1766	2.2634	0.1409
β_3	1.4777	0.1597	0.9485	0.1136
η_3	1.8199	0.1851	2.1766	0.2990

In figures 1–3, we can see by the 95% credible bound that "true" reliability of each component was good estimated. For the component X_1 in the parallel system (figures 4–6) the 95% credible bound does not contains the "true" reliability function. but the other two components has good estimation. Considering our challenging example we understand that this "problem" in the estimation it is normal, and when the estimate was not good, does not mean that was horrible.

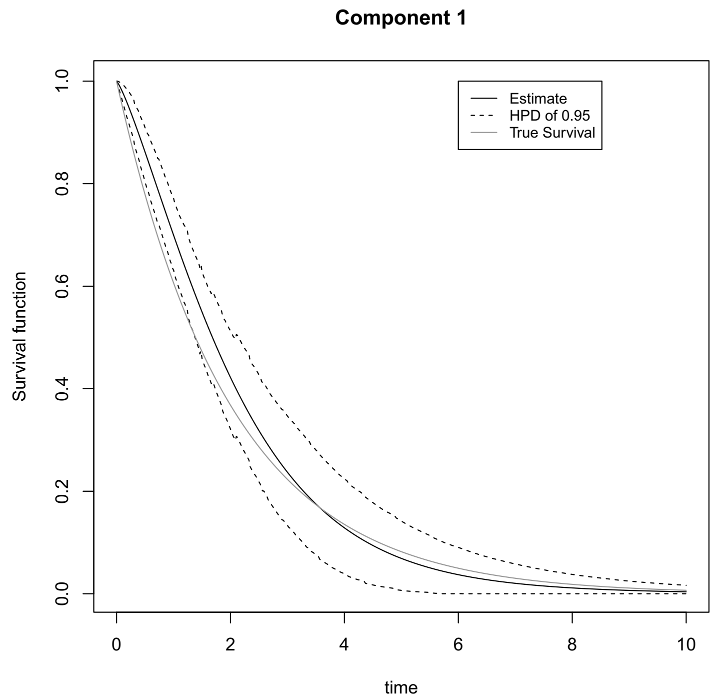

FIGURE 1. Series system example: component 1.

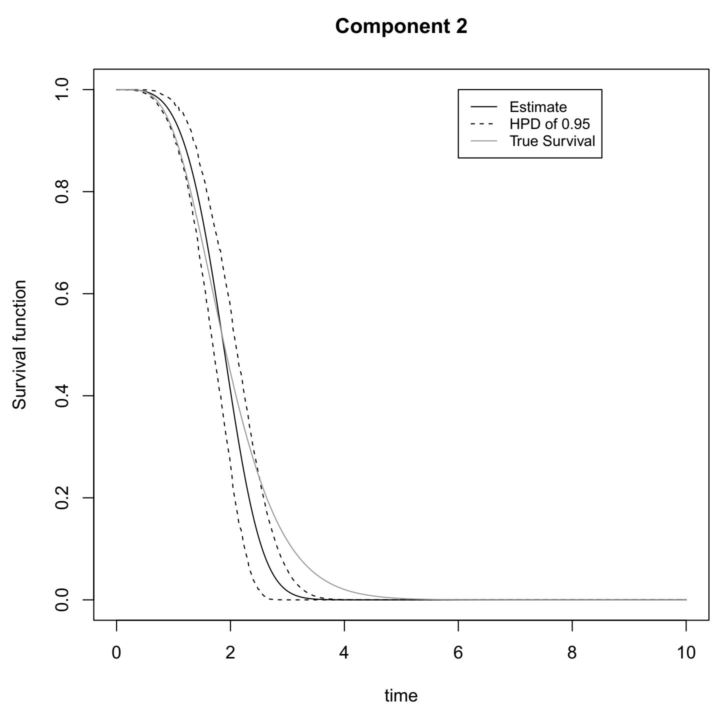

FIGURE 2. Series system example: component 2.

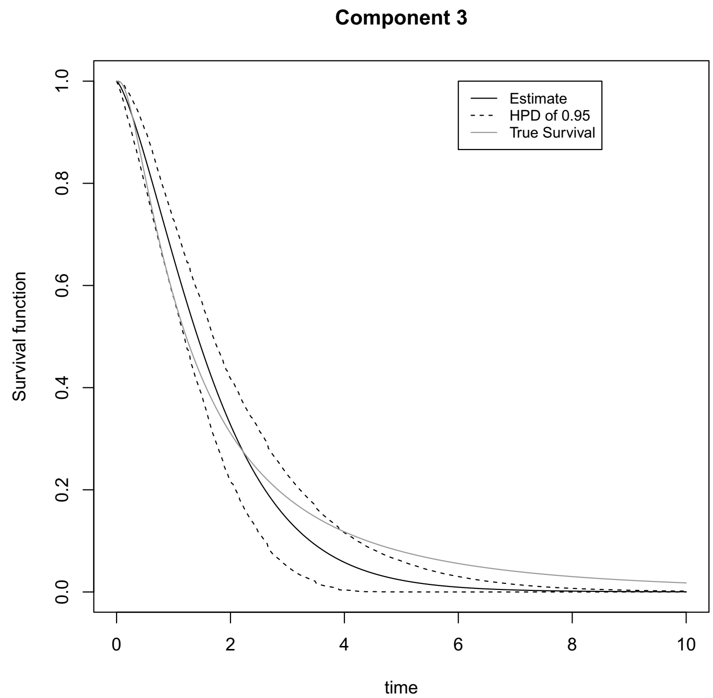

FIGURE 3. Series system example: component 3.

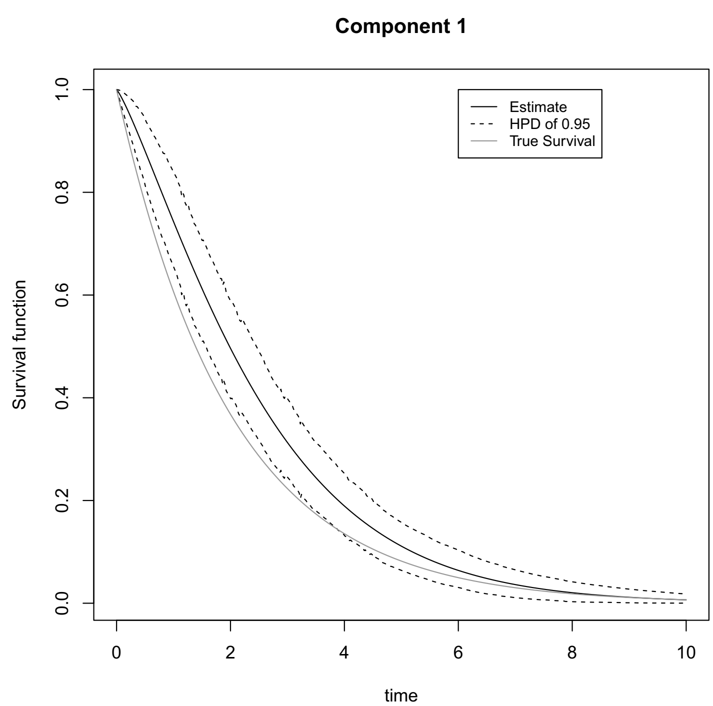

FIGURE 4. Parallel system example: component 1.

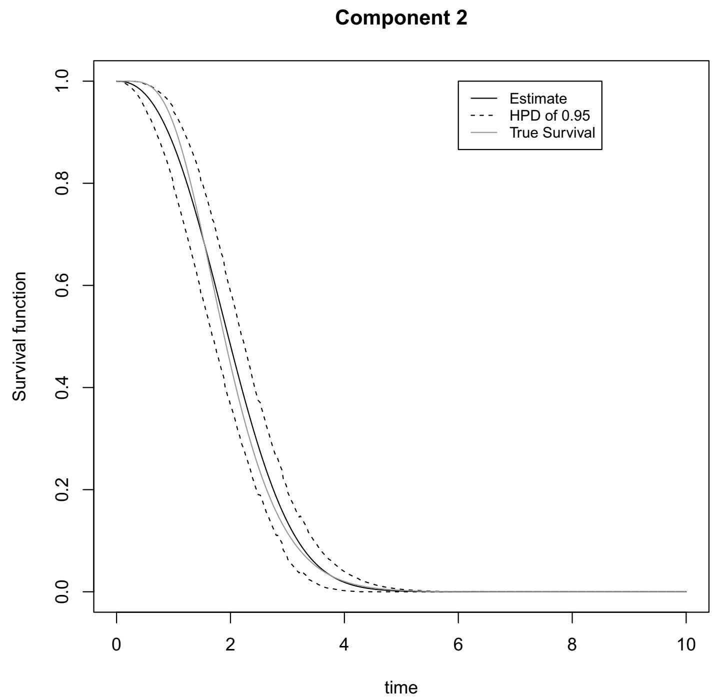

FIGURE 5. Parallel system example: component 2.

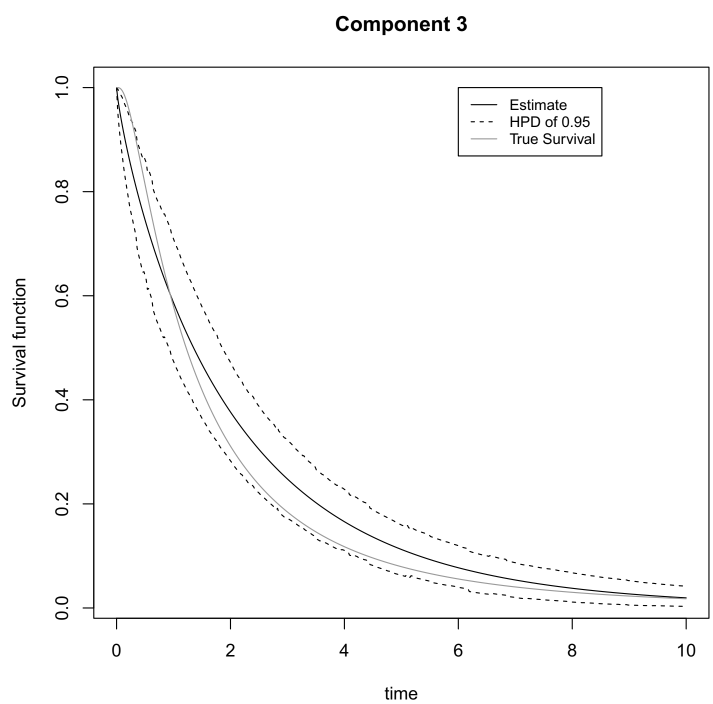

FIGURE 6. Parallel system example: component 3.

4. FINAL REMARKS

We presented a Bayesian statistical analysis with a hierarchical model for the problem of estimating the reliability function and credible bounds. This model shows a good convergence of MCMC tools, making easier the inference process. Also, we used in our formulation an "non-informative" scheme to define the prior hyper-parameters, which implies that the procedure can be used for any type of data. The simulated example showed that the model performed well, however for component X_1 in the parallel system we found the "worst" estimates for reliability function. We believe with an improvement in our algorithms can give us better estimates. This will be addressed in future works. Another important aspect is the ease of obtain the credible bounds for the reliability function, which is not a simple task when one use a plug-in estimator for the reliability function. Some questions that should be addressed in future works are: (i) the hypothesis test of the components, for example, one can be interested in test the hypotheses of equal means of all components (or a subset of components), and (ii) other types of systems, extend this work to coherent system perhaps.

REFERENCES

1. B. P. Carlin and T. A. Louis. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall/CRC, 2nd edition, 2000.
2. E.L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–481, 1958.
3. G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. John Wiley & Sons, 2008.
4. A. Polpo and C.A.B. Pereira. Reliability nonparametric Bayesian estimation in parallel systems. *IEEE Transactions on Reliability*, 58(2):364–373, June 2009.
5. A. Polpo and D. Sinha. Correction in Bayesian nonparametric estimation in a series system or a competing-risks model. *Statistics & Probability Letters*, 81(12):1756–1759, December 2011. doi: 10.1016/j.spl.2011.07.023.
6. A. Polpo, M. Coque, and C.A.B. Pereira. Statistical analysis for weibull distributions in presence of right and left censoring. In *The Proceedings of 8th International Conference on Reliability, Maintainability and Safety (ICRMS)*, volume 1, pages 219–223, Chengdu, 2009. IEEE. doi: 10.1109/ICRMS.2009.5270204.
7. A.S. Rodrigues, T.C.M. Dias, M. Lauretto, and A. Polpo. Reliability analysis in series systems: An empirical comparison between bayesian and classical estimators. In *31st International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Waterloo, AIP Conference Proceedings*, volume 1443, pages 214–221, Melville, 2011. American Institute of Physics. doi: 10.1063/1.3703638.