Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
©2023 ESREL2023 Organizers.
Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P421-cd

A Bayesian Population Variability-based Methodology for Reliability Assessment in the Oil and Gas Industry

Beatriz Cunha, Rafael Azevedo, Marcio Moura, João Mateus Santana, Caio Souto Maior, Isis Lins, Paulo Gabriel Siqueira, Thais Lucas

Center for Risk Analysis, Reliability Engineering and Environmental Modeling (CEERMA),

Industrial Engineering Department, Universidade Federal de Pernambuco, Brazil.

E-mail: paulo.gabriel@ufpe.br, thais.lucas@ufpe.br, isis.lins@ufpe.br, marcio.cmoura@ufpe.br, beatriz.cunha@ufpe.br, joao.mateus@ufpe.br, azevedo144@hotmail.com, caio.maior@ufpe.br

Everton Lima

Petrobras S.A., Brazil. E-mail:evertonlima@petrobras.com.br

Renato Mendes

ABRISCO, E-mail: renatomendes@abrisco.com.br

Scarcity of historical failure data is very common in many situations, especially in the Oil and Gas (O&G) industry. In this context, the Bayesian analysis is paramount to obtain reliable estimates for the system of interest. To perform this analysis, we propose using the Bayesian population variability analysis in a two-step approach. Such an approach allows the assessment of the variability of reliability measures among a similar population of systems. The first step is based on the prior estimation, and it involves gathering available data from similar systems (generic data) and constructing the prior distributions, that represents the population variability. This prior information consists of data of systems that exhibit similar, yet different reliability behavior. In the second step, one can proceed to posterior estimation, where the prior distribution is updated with the available evidence from the system of interest. To obtain the posterior estimates, Markov Chain Monte Carlo-based methods are required. In this work, we illustrate this approach assuming systems with non-constant failure rates, and the model was validated using synthetic data and the results indicate the usefulness of this approach in the O&G industries to better address the reliability measures of its systems.
Keywords: Bayesian analysis; Particle swarm optimization; Markov Chain Monte Carlo; Oil & Gas Industry.

1. Introduction

High quality reliability estimates commonly requires a considerable amount of data, what can be scarce, expensive, or even unfeasible to collect. To face this situation, we proposed a Bayesian approach to use generic data that considers non-constant failure rates that can be modelled by a power function of time, such as a Weibull distribution. Thus, an appropriate Weibull counting process is required to perform the estimation of the Weibull parameters of the prior distributions, via Empirical Bayes (Figure 1), using the Maximum Likelihood Estimation (MLE) (Shultis et al., 1981). The Bayesian approach is implemented to estimate the posterior distribution, by updating prior beliefs using system-specific failure data (censored or not) as a Weibull likelihood. In this case, the posterior distribution cannot be obtained analytically, and Markov Chain Monte Carlo (MCMC) (Bolstad, 2010) was applied for a numerical solution.

2. Methodology applied

We focus on generic data given as paired entries (k_{ij}, t_{ij}) of the number of failures over an observation time for equipment j of subpopulation i. The estimation of the prior distribution is obtained via Empirical Bayes in terms of the hyperparameters ($h = a_\alpha,\ b_\alpha,\ a_\beta,\ b_\beta$) and Particle Swarm Optimization (PSO) (Bratton & Kennedy, 2007) and MLE are applied to find the set that maximizes the likelihood function of the prior information Eq.(1). Figure 1 presents the methodology that incorporates non-homogeneous generic data as basis for the posterior distribution estimation with specific data.

$$\sum_{i=1}^{NP} \log P(E_i | a_\alpha, b_\alpha, a_\beta, b_\beta) \tag{1}$$

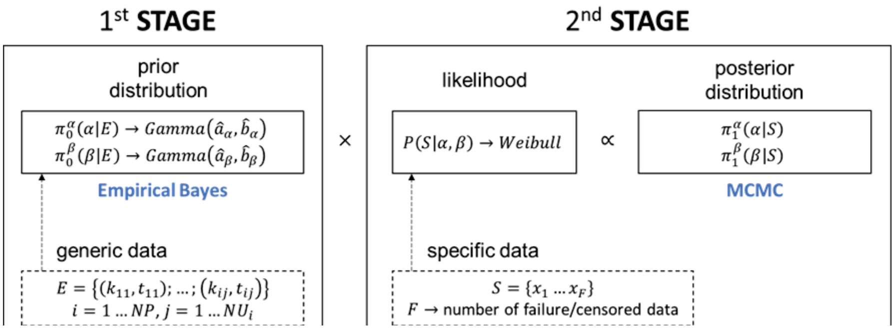

Figure 1. Two-Stages Bayesian approach to non-constant failure rate.

For each case, we ran PSO several times to better assess the variability inherent to the method. NRMSE (Normalized Root Mean Squared Error) is the metric we use to evaluate the results of the prior distribution estimation, since the PSO's results can be feasible (i.e., an optimum solution satisfying the restrictions), although they may not properly represent the input data.

The non-parametric posterior distribution is obtained via MCMC and fit tests are performed to identify the parametric probability distribution that best represents the sampled data. The maximum likelihood hyperparameters are estimated and the quality of fit of each of the resulting distributions is assessed using a Kolmogorov-Smirnov test (Corder & Foreman, 2014).

3. Results

The proposed methodology was run 30 times using synthetic data comprised of 5 equipment and 5 subpopulations for the prior estimation $\{a_\alpha, b_\alpha, a_\beta, b_\beta = 8, 0.0022, 13, 8.5\}$ and a simulation of failure times considering a Weibull (with α=3566.61 and β=1.34) for the posterior. Fig. 2 brings the results of this test. The posterior distributions seem to deviate from the prior in the sense that it is more skewed toward data.

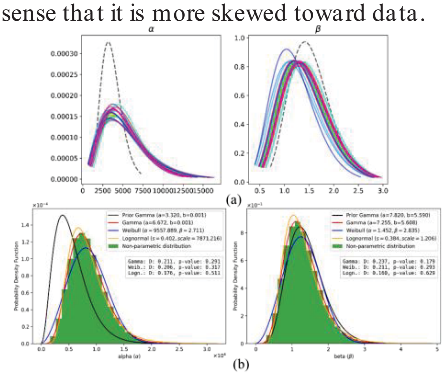

Figure 2. Prior (a) and posterior (b) estimation results.

Aside from the non- parametric distribution results, we also fit these results to three distributions: Gamma, Weibull, and Lognormal. Generally, the Gamma distribution provides a good fit in terms of the p-value based on the Kolmogorov-Smirnov test (always above 0.05).

4. Conclusion

The proposed methodology provides a solution to enable reliability estimation when we face the following problems: scarce specific data, generic data in the form of (k,t) and non-constant failure rates. As future work, we intend to investigate the variability in the prior estimation by using different heuristic methods for the MLE.

Acknowledgement

The authors thank Petrobras S.A., CNPq, FACEPE, and PRH 38.1, managed by ANP and FINEP, for the financial support through research grants. This study was financed in part by CAPES – Finance Code 001.

References

Bolstad, W. (2010). *Understanding Computational Bayesian Statistics* (Vol. 4, Issue 1). Wiley.

Bratton, D., & Kennedy, J. (2007). *Defining a standard for particle swarm optimization*. Proceedings of the 2007 IEEE Swarm Intelligence Symposium, SIS 2007, Sis, 120–127.

Corder, G., & Foreman, D. (2014). *Nonparametric Statistics: A Step-by-Step Approach*. Wiley.

Shultis, J., Johnson, D., Miliken, G., & Eckhoff, N. (1981). *Use of nonconjugate prior distributions in compound failure models*.