UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

BEATRIZ SALES DA CUNHA

A TWO-STAGE BAYESIAN PROCEDURE FOR FAILURE COUNT DATA ASSUMING A NON-CONSTANT FAILURE INTENSITY FUNCTION

Recife

2025

BEATRIZ SALES DA CUNHA

A TWO-STAGE BAYESIAN PROCEDURE FOR FAILURE COUNT DATA ASSUMING A NON-CONSTANT FAILURE INTENSITY FUNCTION

> Doctoral thesis presented to the *Programa de Pós-Graduação em Engenharia de Produção* of *Universidade Federal de Pernambuco* as part of the requirements for the attainment of the doctorate degree in *Engenharia de Produção*. Concentration area: Operations Research.

Advisor: Prof. Dr. Márcio José das Chagas Moura

Recife

2025

BEATRIZ SALES DA CUNHA

A TWO-STAGE BAYESIAN PROCEDURE FOR FAILURE COUNT DATA ASSUMING NON-CONSTANT FAILURE INTENSITY

Doctoral thesis presented to the Programa de Pós-Graduação em Engenharia de Produção of Universidade Federal de Pernambuco as part of the requirements for the attainment of the doctorate degree in Engenharia de Produção. Concentration area: Operations Research.

Aprovado em: 10/03/2025

BANCA EXAMINADORA

_____
Prof. Dr. Márcio José das Chagas Moura (Orientador)
Universidade Federal de Pernambuco - UFPE

_____
Profª. Drª. Isis Didier Lins (Examinador Interno)
Universidade Federal de Pernambuco - UFPE

_____
Prof. Dr. Cristiano Cavalcante (Examinador Interno)
Universidade Federal de Pernambuco – UFPE

_____
Prof. Dr. Enrique López Droguett (Examinador Externo)
University of California - UCLA

_____
Prof. Dr. Danilo Colombo (Examinador Externo)
Petrobras

ACKNOWLEDGMENTS

Completing this doctorate marks a significant milestone in my professional life. Achieving this after the COVID-19 pandemic and personal challenges was not easy, but it was far from a solo journey. Henry Ford's words resonate deeply: "Whether you think you can or can't, you're right". Professor Marcio, thank you for consistently reminding me there is always a way to make things work. I will be forever grateful for your trust, guidance, and all the opportunities you have been giving me since you started advising me in 2016. Gratitude also goes to Professor Isis Lins, whose advisory role, though less frequent, has been no less pivotal. I extend my thanks to my colleagues at CEERMA-UFPE, especially those directly involved in this project: João Mateus, Caio, Rafael, Thais, Paulo, Ana Claudia and Léo. Thank you for the long meetings and your hard work.

To my mother, Solange, your strength and resilience are a constant source of inspiration, reminding me to be a better person and embrace responsibility. To my father, Carlos, you taught me the importance of learning and showed me that it can be not only fun but also a lifelong pursuit. For you both, thank you for encouraging me to spread my wings. I thank my brother, Carlinhos, for the mood-boosting playlists, conversations, and hugs, and for teaching me about selfless love. To my husband, João, who teaches me through our differences and shows me daily the importance of love and attention, thank you for sharing in my dreams. My dear friends, thank you all for putting up with me. I thank specifically for supporting me throughout this journey: Gabi, Macedo, Aninha, Vivi, Gabi K, Cristian, and Tayla.

ABSTRACT

Reliability analysis is essential in high-risk industries like Oil and Gas (O&G) for predicting equipment lifespan, anticipating costs, planning maintenance, and estimating system availability. However, failure data is often limited due to proprietary restrictions, high acquisition costs, or challenges in data collection. Bayesian inference addresses this limitation by enabling the integration of generic data, which can be updated with new specific data to generate a posterior distribution. The Offshore & Onshore Reliability Data (OREDA) provides a valuable source of generic data, created through collaboration among O&G companies to share information on equipment operation and maintenance. Traditional analysis of OREDA data often assumes a constant failure intensity function, which is not always accurate. This work generalizes the analysis by incorporating a non-constant failure intensity function using the Weibull distribution. Given the lack of a conjugate prior to this model, posterior estimates are obtained via Markov Chain Monte Carlo (MCMC) sampling. The model was validated using simulated data, demonstrating robust performance across various test sets, particularly in terms of relevant performance metrics, despite some variability in the prior distribution estimation stage. Following this validation, the model was applied to a real-world industrial case involving booster pumps, extending traditional reliability methodologies by integrating non-constant failure intensity into the analysis. This model was incorporated into the Petrobayes software, which is presented in this work and enables streamlined execution to enhance accessibility and practical application.

Keywords: OREDA, Bayesian inference, Weibull distribution, Population Variability Distribution, Markov Chain Monte Carlo.

RESUMO

A análise de confiabilidade é essencial em indústrias de alto risco, como a de Óleo e Gás, pois permite prever a vida útil dos equipamentos, antecipar custos, planejar estratégias de manutenção e estimar a disponibilidade dos sistemas. No entanto, a disponibilidade de dados é frequentemente limitada devido a restrições de propriedade, altos custos de aquisição ou dificuldades na coleta de dados. A inferência Bayesiana aborda essa limitação ao permitir a integração de dados genéricos, que podem ser atualizados com novos dados específicos para gerar uma distribuição a posteriori. O *Offshore & Onshore Reliability Data* (OREDA) fornece uma valiosa fonte de dados genéricos, criada através da colaboração entre empresas de Óleo e Gás para compartilhar informações sobre a operação e manutenção de equipamentos. A análise tradicional dos dados OREDA frequentemente assume uma taxa de falha constante, o que nem sempre é preciso. Este trabalho generaliza a análise ao incorporar taxas de falha não constantes usando a distribuição Weibull. Dada a ausência de uma priori conjugada para este modelo, as estimativas a posteriori são obtidas por meio de amostragem de Monte Carlo via Cadeia de Markov. O modelo foi validado utilizando dados simulados, demonstrando desempenho robusto em diversos conjuntos de testes, especialmente em termos de métricas de desempenho relevantes, apesar de apresentar variabilidade na etapa de estimativa da distribuição a priori. Após essa validação, o modelo foi aplicado a um caso industrial real envolvendo bombas de reforço, ampliando as metodologias tradicionais de confiabilidade ao integrar taxas de falha não constantes na análise. Este modelo foi incorporado ao software Petrobayes, apresentado neste trabalho, que permite uma execução simplificada para melhorar a acessibilidade e a aplicação prática.

Palavras-chave: OREDA, Inferência Bayesiana, distribuição Weibull, Distribuição de Variabilidade Populacional, Monte Carlo via Cadeia de Markov.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

O&G   Oil & Gas

MCMC   Markov Chain Monte Carlo

OREDA   Offshore & Onshore Reliability Data

PVD   Population Variability Distribution

ISO   International Organization for Standardization

MLE   Maximum Likelihood Estimate

LHS   Latin Hypercube Sampling

CDF   Cumulative Distribution Function

NHPP   Non-homogeneous Poisson Process

HPP   Homogeneous Poisson Process

PSO   Particle Swarm Optimization

NRMSE   Normalized Root Mean Squared Error

HMC   Hamiltonian Monte Carlo

NUTS   No-U-turn sampling

KS   Kolmogorov-Smirnov

LIST OF SYMBOLS

θ	Vector of unknown parameters
S	Specific data
$\pi_0(\theta)$	Prior probability density function of θ
$L(S|\theta)$	The likelihood function of S given θ
$\pi_1(\theta|S)$	The posterior probability density function of θ given S
ϕ	Vector of hyperparameters
$\rho\left(\theta\middle|\underline{\phi}\right)$	PVD of θ given ϕ
E	Generic data, where $E = \{E_1, \dots, E_{NP}\}$
NP	Number of subpopulations
i	Subpopulation index, with $i = 1 \dots NP$
NS_i	Number of systems per subpopulation i
j	System index, with $j = 1 \dots NU_i$
λ	The failure rate parameter of the Exponential distribution
α	The scale parameter of the Weibull distribution
β	The shape parameter of the Weibull distribution
a_θ	The shape parameter of the Gamma distribution of θ, addressed as a hyperparameter
b_θ	The rate parameter of the Gamma distribution of θ, addressed as a hyperparameter
$\hat{\mu}_\lambda$	Failure rate mean obtained directly from the data.
$\hat{\sigma}_\lambda^2$	Failure rate variance, obtained directly from the data
k_{ij}	Number of failures for system j of subpopulation i
t_{ij}	Observation time for system j of subpopulation i
$h(t)$	The failure intensity function of the Weibull distribution over time t

$P(E_i\|\alpha,\beta)$	Probability of observing E_i for specific α and β values (Weibull counting process)
n	Number of samples generated with LHS
l	Samples index, with $l = 1, \dots, n$
N	Number of disjoint partitions of LHS
$E[.]$	Expected value
$S[.]$	Standard deviation
it	PSO's number of iterations
x	Decision variable of PSO
lb_x	Lower bound of x for PSO search
ub_x	Upper bound of x for PSO search
$E\big[N(t_g)\big]$	Expected number of failures for time t_g estimated via Monte Carlo simulation
G	Number of equipment groups per subpopulation
g	Group index, with $g = 1 \dots G$
cs_{min}	Constant in the $lb_{S(\alpha)}$ calculation
cs_{max}	Constant in the $ub_{S(\alpha)}$ calculation
z	$z = \dfrac{\sum_i \sum_j t_{ij}}{\sum_i \sum_j k_{ij}}$
np	PSO's number of particles
c_1	PSO's cognitive parameter
c_2	PSO's social parameter
w	PSO's inertia weight

TABLE OF CONTENTS

1   INTRODUCTION

Every year, the Oil & Gas (O&G) industry allocates billions of dollars to petroleum exploration (Ren et al., 2023), navigating a complex landscape marked by technological advancements, aging workforces, environmental concerns, rising energy demands, and the constant challenge of managing emergencies (Mahmoudi, 2021). Amidst these challenges, accurately estimating the probability of equipment failure has become crucial, especially when the economic consequences, such as downtime and repair costs, are significant (Sandtorv & Thompson, 1996).

This estimation is crucial for forecasting the system's expected availability and optimizing the maintenance planning process, helping decision-makers in the face of uncertainty (Hartley & French, 2021). Given that equipment is often engineered to endure extreme environments and operating conditions (Macedo et al., 2023), quantitative models play a key role in guiding decisions about safety policies and operational procedures. These models help reduce both the likelihood and severity of failures (das Chagas Moura et al., 2016).

Failure is defined as the loss of a component's intended functionality, while the failure mode refers to the specific way in which this failure occurs (Rausand & Oien, 1996). Ensuring the accuracy of failure estimates typically requires a substantial amount of historical data on failure frequency (Sandtorv & Thompson, 1996), with the quality of the analyses highly dependent upon the quality of the data being used.

In this context, the Offshore & Onshore Reliability Data (OREDA) project arose as a collaboration between several companies in the industry. It comprises a database of reliability data collected on Topside and Subsea equipment from offshore and onshore operations. The objective is to promote safer operations, increased production availability, and optimized maintenance (Langseth et al., 1998), by sharing information regarding the failure and maintenance of equipment (SINTEF & NTNU, 2015). The challenge lies in determining what can be inferred about reliability measures, and with what level of confidence, by drawing on failure data from similar machines in generic databases such as OREDA ("generic data"), and failure data from the specific machine(s) under analysis ("specific data") (Kaplan, 1983).

In this context, Bayesian inference allows the construction of prior knowledge from generic data (first stage), which can be updated, as new specific data is acquired,

using Bayes' theorem (second stage). This two-stage approach enables data integration from various sources through a mathematically consistent procedure for equipment in operation that faces a shortage in historical data (Maior et al., 2022). Inference, as defined by Kelly & Smith (2009), is the process of drawing conclusions based on accumulated knowledge. Gribok et al. (2020) emphasize that the Bayesian approach enhances the frequentist method by incorporating prior information into the likelihood function.

However, two important aspects need to be taken into consideration regarding the OREDA database: (i) it brings information about the performance of equipment in different industrial plants, produced by different suppliers, and, usually, subject to different operating conditions, and (ii) it contains failure count data. Considering (i), it may be unrealistic to consider similar systems exposed to different conditions to have the same behavior reliability-wise, which means that the distinct systems should be treated as non-homogeneous. This hypothesis entails a variability of failure intensity among a population of systems. A representation of this variability, in the form of a probability distribution, is referred to as the Population Variability Distribution (PVD) (Droguett et al., 2004). Hence, the assessment of PVD with generic data offers an alternative to incorporate relevant information for prior distribution estimation in Bayesian inference.

Regarding (ii), failure count data represents the number of failures that occur over a period of time (Hamada et al., 2008). The current methodology for processing failure count non-homogeneous data proposed in SINTEF & NTNU (2015) and ISO 14224 (ISO, 2016) considers a constant failure intensity function and will be detailed in Section 2.2. However, assuming that a component is subject to a constant failure intensity function is not always appropriate, especially due to environmental factors, degradation, and maintenance strategy (Kumari et al., 2022; Q. Yang et al., 2012).

To address this issue, this work focuses on a time-dependent failure intensity function that follows a two-parameter (scale α and shape β) Weibull distribution. This approach can model a variety of behaviors including the case of constant failure intensity function, when the shape parameter equals 1. The idea is develop a two-stage Bayesian approach suitable for application to OREDA's data. In this context, an appropriate Weibull counting process is required to deal with failure count data and

perform the estimation of the Weibull parameters' prior distributions via Empirical Bayes (Shultis et al., 1981). In addition, the Bayesian inference approach is implemented to estimate the posterior distribution, by updating prior beliefs using specific data in a Weibull likelihood. In this case, the posterior distribution cannot be obtained analytically, and Markov Chain Monte Carlo (MCMC) will be applied for a numerical solution (Bolstad, 2009). The MCMC method is widely used because it enables direct sampling from the posterior distribution, without the need to solve potentially high-dimensional integrals (Straub & Papaioannou, 2015).

The proposed approach will be tested using simulated data of different sizes to assess the model's ability to reproduce the original known distribution that generated the data. Sensitivity analyses and performance metrics such as p-value, Normalized Root Mean Squared Error (NRMSE), and log-likelihood were employed to assess the model's robustness and adaptability, providing valuable insights into its applicability across various scenarios. The validation is followed by an application to a real-world industrial case from OREDA involving booster pumps. To enhance accessibility and usability, the methodology was encapsulated in a web application, called Petrobayes, enabling users to conduct reliability analysis efficiently.

Hence, the main contribution of this work is twofold: the use of prior information in the form of failure count data within a non-constant failure intensity model based on the Weibull distribution, and the development of a standalone module within a web application, providing analysts with an user-friendly interface to apply this model. Indeed, these contributions are particularly significant as they overcome inherent limitations with using the OREDA database, especially concerning the challenging use of non-constant failure intensity function. Therefore, this work overcomes limitations in risk and reliability analyses in the O&G sector and supports effective risk management for equipment in operation with a shortage of historical data. The proposed model has been published in (Sales da Cunha et al., 2024) and the web application was described in (Santana et al., 2023) and presented at the 32nd European Safety and Reliability Conference (ESREL 2022).

1.1 Motivation

O&G assets are complex, multi-component structures designed to operate for approximately 20 to 30 years under challenging environmental conditions (Maior et al., 2022). Due to these extended lifecycles and harsh conditions, gathering

comprehensive and reliable failure data over time is a resource-intensive and lengthy process, often requiring years of observation and monitoring to gather sufficient information for accurate analysis and decision-making. In fact, Sandtorv & Thompson (1996) pointed out that pooling data collection efforts across multiple companies offers significant cost-benefit advantages compared to individual company initiatives.

The OREDA database exemplifies such collaborative efforts. Initially developed to inform decisions during the design and engineering phases (Sandtorv & Thompson, 1996), this cross-company database can also support decision-making during the operational phase, where this work is focused on. Evaluating the reliability of equipment under conditions of limited failure data is a common challenge in O&G operations. These systems operate in extreme environments and are subject to high costs and operational risks, making the scarcity of failure data a significant challenge for reliability analysis.

To address this challenge, this work proposes a Bayesian inference model that integrates generic data sources like OREDA to generate updated reliability estimates for equipment of interest. The proposed model is specifically tailored to handle failure count data, which is the type of data commonly found in the OREDA database. This model accounts for time-varying failure intensity function and system aging, common assumptions in O&G environments. It also allows for iterative updates as new specific data becomes available, creating a dynamic learning curve that enhances prediction accuracy and decision-making over time.

Hence, the model incorporates two key assumptions: (i) non-constant failure intensity and (ii) minimal repair conditions. Unlike current methodologies that often assume constant failure intensity function, (i) accounts for factors such as environmental degradation, maintenance practices, and repair events (Kumari et al., 2022; Q. Yang et al., 2012). The minimal repair condition, where a system is restored to its state just before failure, is modeled through a non-homogeneous Poisson process (NHPP) (das Chagas Moura et al., 2014). In the O&G industry, maintenance often addresses only the immediate cause of failure, leaving the equipment's overall wear level unchanged, as full replacements or major repairs are costly and time-consuming due to offshore or remote installation locations (Osheyor Gidiagba et al., 2023). In fact, O&G assets operate under extreme conditions such as high pressure, corrosion, and mechanical stress. These harsh environments accelerate equipment

degradation, making minimal repairs more feasible and economical, balancing operational continuity with maintenance costs.

This work tackles the challenge of estimating reliability in scenarios with limited data, leveraging the OREDA database as a foundation. By considering the nature of OREDA's failure count data and adopting realistic assumptions for the Oil & Gas sector—such as time-varying failure intensity and minimal repair—it makes important methodological advances. A key contribution is the adaptation of the traditional constant failure intensity model to accommodate non-constant failure rates, enabling its practical application to real-world data. This required revising the original formulation and developing tailored methods to address the specific nuances of the dataset. The use of the Weibull distribution, known for its flexibility in modeling different failure behaviors, in combination with the NHPP to reflect minimal repair scenarios, represents a novel integration of techniques. This approach fills a gap in the literature and provides a more accurate and practical framework for reliability estimation in the O&G industry.

Hence, the impacts of this work span economic, environmental, and financial dimensions. Economically, the proposed model enhances maintenance strategies development for equipment in operation, improving resource allocation and operational efficiency. By reducing unplanned downtime and extending equipment lifespan, it lowers overall operational costs and improves productivity. Environmentally, it enhances maintenance planning, which minimizes the need for emergency repairs that could lead to oil spills or unanticipated emissions. This fosters safer and more sustainable operations, reducing risks to the environment. Financially, the improved accuracy of failure predictions through Bayesian techniques enables better budgeting for maintenance expenditures. By optimizing investments in preventive maintenance and reducing unnecessary corrective costs, the model strengthens financial planning, enhances cash flow stability, and improves long-term profitability.

In conclusion, this work addresses a critical problem in the O&G industry by advancing reliability analysis methodologies. It offers significant economic, environmental, and financial benefits, contributing to safer, more efficient, and sustainable operations in high-risk industrial environments.

1.2   Research Objectives

1.2.1 General Objective

The primary objective of this work is to develop a Bayesian inference model leveraging failure count data, which is the type of data found in OREDA, as prior knowledge. This model operates under the assumption that failure times follow a Weibull distribution with minimal repair conditions.

1.2.2 Specific Objectives

To successfully achieve the goal of this work, the following specific objectives have been identified and outlined:

- To perform a literature review: to conduct a thorough study to review the current state of the art on Bayesian inference, considering both constant and non-constant failure intensity function hypotheses.

- To develop a Bayesian framework: to design a Bayesian inference model that integrates prior knowledge in the form of failure count data, which is the type of data found in the OREDA database. The model should be adaptable to real-world scenarios, assuming failure times follow a Weibull distribution for accurate reliability assessment.

- To investigate the minimal repair assumption: to investigate situations where systems undergo minimal repairs, where the system is restored to the state it was in just before the failure occurred, and develop the model based on this assumption.

- To perform prior distribution analysis: to develop methods and solutions to estimate prior distributions using OREDA data, ensuring that the estimation process is effective and reliable for real-world applications.

- To perform posterior distribution analysis: to generate and evaluate posterior distributions for failure data, updating prior information from OREDA with new equipment-specific data via MCMC.

- To evaluate model performance: to conduct simulations and sensitivity analysis to validate the model's robustness, particularly in scenarios with limited amounts of data.

- To apply the model to a real case study: to apply the Bayesian model to actual failure data from booster pumps obtained from OREDA and use specific data to demonstrate its practical application.

- To develop an easy-to-use interface for model execution: to develop a standalone module inside the web app called Petrobayes, for risk analysts to execute the model with ease and enhance reliability analysis.

1.3   Research Methodology

The scientific methodology adopted in this work follows a structured approach to ensure analytical rigor and validity. In terms of its nature, this is applied research aimed at solving a specific problem related to reliability analysis in the O&G industry. Regarding the approach to the problem, the research is quantitative, relying on statistical models, simulations, and numerical evaluations. From the perspective of its objectives, the study is explanatory, as it seeks to understand and explain the reliability behavior of equipment in the O&G industry. This is achieved through an innovative two-stage Bayesian inference model that integrates both generic and specific data sources. Concerning technical procedures, the work is experimental. Simulations are conducted to generate synthetic datasets, which are used to validate the proposed model. This experimental approach supports a robust evaluation of the model's performance, ensuring that its applicability extends to real-world industrial contexts.

1.4   Thesis Structure Outline

This thesis is organized into chapters that aim to guide the reader through the progression of the research, from the initial concepts and methodology to the detailed analysis, findings, and conclusions drawn from the study. The connection between the chapters is presented in Figure 1.

Figure 1 - Thesis outline.
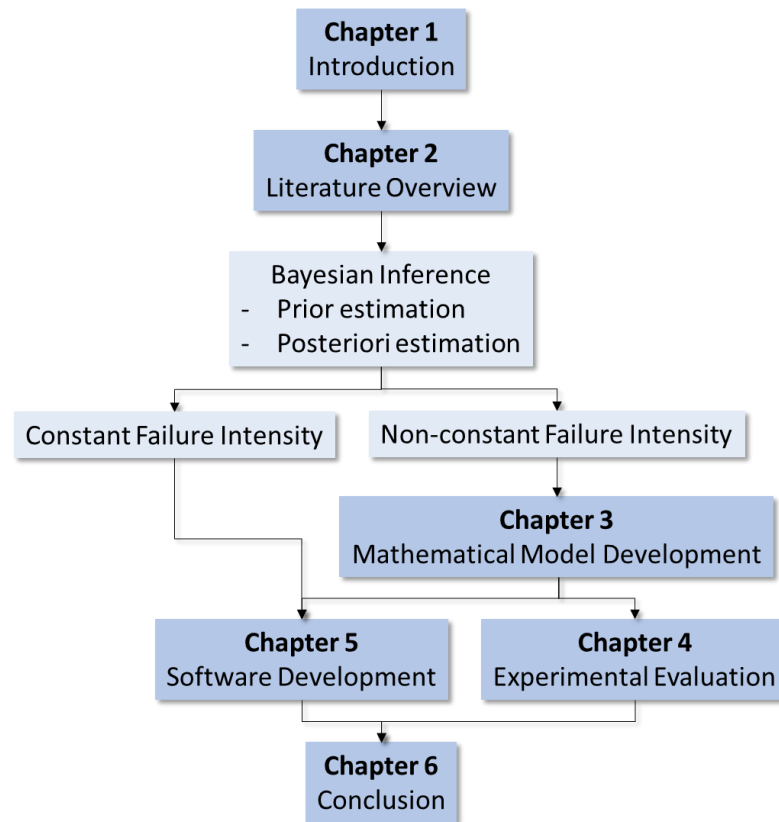
Source: The author (2025).

A detailed description regarding the chapters' content is provided as follows:

- Chapter 2 explores the foundational concepts of Bayesian inference, more specifically on the two-stage Bayesian approach based on prior estimates (first stage) and posterior estimates (second stage). It focuses particularly on the commonly adopted case of constant failure intensity function and the methods involved in adapting this formulation to the non-constant failure intensity case. It also provides more details regarding the OREDA database creation, format, and application.

- Chapter 3 outlines the Bayesian inference model for the non-constant failure intensity case under the assumption of minimal repair, detailing the methods used for estimating both prior and posterior distributions.

- Chapter 4 presents the results for both prior and posterior estimations, including a detailed description of the data simulation process used for validation and the parameters employed for testing. It also provides an industrial application example, with data from booster pumps gathered from OREDA.

- Chapter 5 presents a comprehensive overview of the web application Petrobayes, focusing on the standalone module designed to execute the proposed model. It details the module's implementation within the platform and its integration with other functionalities.
- Chapter 6 summarizes the key findings and discusses potential directions for future research.

2   LITERATURE REVIEW

Reliability analysis plays a crucial role in predicting equipment lifespan, enabling the anticipation of costs, the planning of maintenance strategies, and the estimation of system availability. Understanding the probabilistic behavior of failures is essential for informed decision-making throughout the equipment's entire life cycle. According to the International Organization for Standardization (ISO) (ISO, 2020), reliability is defined as "the degree to which a system, product or component performs specified functions under specified conditions for a specified period of time". The expectation regarding the system's ability to perform its intended function is frequently expressed in terms of the failure intensity (Droguett et al., 2004).

2.1   Bayesian Inference

Ensuring the accuracy of failure estimates typically requires a substantial amount of data. However, in industries with highly reliable equipment, such as the O&G, data availability is often limited due to proprietary restrictions, high acquisition costs, or the impracticality of data collection (René Van Dorp & Mazzuchi, 2004). To address this challenge, in the context of Bayesian inference, Kaplan (1983) introduced a two-stage Bayesian procedure that leverages industry-wide, partially relevant data, referred to as "generic data". Although this data may not be directly tied to the system of interest, it is collected from similar systems or applications, which are assumed to exhibit comparable reliability patterns despite differences in design and operating conditions (Droguett et al., 2004).

The two-stage Bayesian procedure is a specific instance within the broader framework of hierarchical Bayes (Kelly & Smith, 2009). In the first stage, a prior distribution is built around the variability observed in the failure behavior across similar systems with generic data (Kodoth et al., 2020). In the second stage, the prior probability distribution is updated based on the likelihood of evidence specific to the system under analysis ("specific data") (Lin, 2002). This updating process, based on Bayes' Theorem, results in a posterior distribution that reflects the revised uncertainty regarding the failure behavior (Sedehi et al., 2019).

Bayes' Theorem, shown in Equation (1) indicates that the posterior distribution is proportional to the product of the prior distribution and the likelihood, as these two components are independent (Bolstad, 2007). The posterior $\pi_1(\theta|S)$ represents the

probability distribution of the unknown parameter θ given the specific data S. The prior probability density function $\pi_0(\theta)$ reflects the uncertainty about θ, while $L(S|\theta)$ represents the likelihood of the specific data S given θ. Depending on the chosen prior distribution $\pi_0(\theta)$ and the likelihood function $L(S|\theta)$, numerical integration may be required.

$$\pi_1(\theta|S) = \frac{L(S|\theta)\pi_0(\theta)}{\int L(S|\theta)\pi_0(\theta)d\theta} \tag{1}$$

Thus, the statistical model comprises two key components: the prior distribution $\pi_0(\theta)$ and the likelihood function $L(S|\theta)$. The parameter θ in the likelihood function is treated as an unknown quantity, and the prior distribution is used to express the uncertainty about this parameter. The following sections provide a detailed explanation of these two components.

Bayesian analysis differs from classical frequency-based analysis in its treatment of probability. While classical methods tie probability to the frequency of events, Bayesian analysis interprets probability as a subjective measure of belief or knowledge about model parameters (Bolstad, 2007). This does not imply that model parameters are random but rather that their values are uncertain, and this uncertainty is expressed through a probability density function (Hamada et al., 2008).

As a direct consequence of its interpretation of subjective probability, one significant advantage of the Bayesian approach, as highlighted by Hamada et al. (2008), is its ability to incorporate information from expert opinions and domain expertise. Expert opinions are often used to estimate prior distributions or specific reliability metrics, such as failure rates, reliability, or percentile values (Guo et al., 2018; Siqueira et al., 2022). However, eliciting expert opinions requires expertise in formulating questions and managing human resources, and there is the problem of bias in risk estimates, which may necessitate correction (Mosleh et al., 1988; Zhou et al., 2019). Although expert opinion is not the focus of this work, it remains a crucial aspect of Bayesian inference addressed in various studies (Droguett et al., 2004; Hartley & French, 2021; Macedo et al., 2023; Maior et al., 2022).

Another benefit of the Bayesian approach is its ability to reuse posterior probabilities as priors for new specific data sets. This reduces the need for re-analyzing the entire data when new information becomes available, as using the original and the

new data sets separately leads to the same posterior probabilities (Bolstad, 2007). By updating the current knowledge, the ability to predict uncertain events is improved (Porn, 1996).

Bayesian inference has been applied in many disciplines, such as behavioral science, finance, human health, process control, and ecological risk assessment, especially with the advent of MCMC that made intractable problems to be solved easily (Kelly & Smith, 2009). According to (Gelman et al., 2013), Bayesian inference involves three main steps: (1) constructing a full probability model that defines the joint distribution for both observable and unobservable quantities based on the scientific problem and data collection, (2) conditioning on the observed data to calculate and interpret the posterior distribution, which represents the probability of unobserved quantities given the data, and (3) evaluating the model's fit and the implications of the posterior distribution to determine if adjustments are needed. This process may be repeated to improve the model based on its fit and sensitivity to assumptions.

2.1.1 Prior distribution

In Bayesian inference, prior distributions are classified as either informative or non-informative (Hamada et al., 2008). Non-informative priors are designed to exert minimal influence on the likelihood, reflecting the fact that little is known about the parameter, allowing the specific data to drive the results. An example of a weak prior application is provided by Kelly & Smith (2009), using a Jeffreys prior because the goal was to have results driven by the specific data. The Jeffreys prior implies a state of ignorance about the parameter under analysis given the characteristics of equidistribution on a log scale (Erto & Giorgio, 2013; Frohner, 1985).

In contrast, informative priors incorporate additional knowledge and may be concentrated in a particular region of the parameter space. When data is scarce, relying on a non-informative prior can result in significant uncertainty in the posterior estimation. Thus, it is crucial to incorporate as much relevant information as possible when constructing the prior distribution (L. Yang et al., 2019), as they improve reliability estimates (Gelman et al., 2017; Wilson & Fronczyk, 2016). In fact, from a classical frequentist point of view, inappropriate priors may introduce bias (Gribok et al., 2020).

In the context of reliability analysis, there are six main sources of information for constructing informative prior distributions: physical/chemical theory, computational

analysis, previous engineering and qualification test results from a process development program, industry-wide generic reliability data (such as OREDA), past experience with similar devices, and expert opinion (Hamada et al., 2008). When dealing with industry-wide generic reliability data, or simply generic data, given that one can associate failure behavior with individual systems, one can also consider the variability of failure behavior among a population of systems, introducing non-homogeneity. A representation of this variability, in the form of a probability distribution, is referred to as the PVD (Droguett et al., 2004). Hence, the assessment of PVD with generic data offers an alternative to incorporate relevant information for prior distribution estimation in Bayesian inference. Assuming PVD is given by a parametric distribution, if enough generic data is available, it is possible to derive the distribution from the dataset (das Chagas Moura et al., 2016).

The PVD of the unknown parameter θ is denoted as $\rho\left(\theta\middle|\underline{\phi}\right)$, with $\underline{\phi}$ as the hyperparameter vector. Defining the PVD involves estimating $\underline{\phi}$ assuming $\rho\left(\theta\middle|\underline{\phi}\right)$ is given by a parametric distribution. $\underline{\phi}$ is called a hyperparameter because it is a parameter that does not appear in the likelihood function (Hamada et al., 2008). In the Empirical Bayes approach (Siu & Kelly, 1998), PVD hyperparameters are estimated from the generic data E. Extending the Bayesian inference to include the treatment of the hyperparameters is also referred to as Bayes Empirical Bayes (Porn, 1996).

Prior distributions that take the same functional form as the posterior distribution are called conjugate prior distributions (Porn, 1996). In simple problems, conjugate prior distributions can make posterior analysis easy because they eliminate the need to determine normalizing constants numerically (Hamada et al., 2008). But this should not be the driver for model formulation (Gelman et al., 2017).

2.1.2 Likelihood

The nature of the reliability measure guides the choice of the parametric distribution that will represent the PVD. For instance, Gamma or Lognormal distributions are appropriate choices if θ represents the constant failure intensity because they are defined only for positive real values and are suited to model a wide range of skewed behaviors. Conversely, if θ represents the probability of failure, the Beta distribution is more suitable because its support is the interval [0, 1] (Droguett et

al., 2004). The likelihood function, however, depends on the type of evidence available and the assumptions. For example, if a constant failure intensity function is assumed and if dealing with failure count data—representing the number of failures over a period of time—the Poisson distribution is typically used. On the other hand, if dealing with failure times, the Exponential distribution would be employed since it models the time between events in a Poisson process.

In assessing the variability of the failure behavior, various formulations have been proposed (Frohner, 1985; Grabski & Sarhan, 1996; ISO, 2016; Kaplan, 1983; Langseth & Lindqvist, 2006; Porn, 1996; Vaurio, 1987), with the likelihood function typically based on a Poisson model, which focuses exclusively on failure count data. In fact, many failure data banks are built on the assumption that component failures follow a Poisson process, and thus, systems are often assumed to have a constant failure intensity function (Porn, 1996).

Generic reliability databases, such as OREDA, primarily contain failure count data. The estimation of prior distributions for failure count data is based on three key assumptions: (i) homogeneity within each subpopulation, (ii) heterogeneity between subpopulations, and (iii) a constant failure intensity function (Droguett et al., 2004). The first two assumptions stem from the observation that systems within the same subpopulation operate under similar conditions and share similar design features, which may differ across subpopulations.

The third assumption, a constant failure intensity function, is widely adopted due to its analytical simplicity when using the Poisson distribution for failure count data. While this assumption has been applied in OREDA (SINTEF & NTNU, 2015) and is recommended in ISO 14224 (ISO, 2016), it is somewhat restrictive. The following sections will analyze the constant failure intensity function model and explore potential modifications to generalize this approach.

2.2 The constant failure intensity function case

Assume that both generic (E) and specific (S) data are provided as failure count data, and that the systems operate under a constant failure intensity function, denoted by λ, or simply failure rate. Under these conditions, the Bayesian procedure can be described by Equation (1), with θ representing λ. To apply this procedure, we must define the prior distribution $\pi_0(\lambda)$ and the likelihood function $L(S|\lambda)$. Since we have

failure count data and assume a constant failure intensity function, the failure distribution follows a Homogeneous Poisson process (HPP). The likelihood function, $L(S|\lambda)$, is therefore given by Equation (2), where the specific data S consists of the number of failures K over an observation time T.

$$L(S|\lambda) = \frac{(\lambda T)^K e^{-\lambda T}}{\Gamma(K+1)} \tag{2}$$

Now, we need to determine the prior distribution $\pi_0(\lambda)$. It is common practice to assume a Gamma distribution for the PVD of λ, as shown in Equation (3), due to its conjugate relationship with the Poisson likelihood. This conjugacy simplifies the Bayesian updating process and ensures analytical tractability. Here, a_λ and b_λ represent the hyperparameters of the Gamma prior.

$$\pi_0(\lambda) = \rho(\lambda|a_\lambda, b_\lambda) = \frac{b_\lambda^{a_\lambda} \lambda^{a_\lambda - 1} e^{-b_\lambda \lambda}}{\Gamma(a_\lambda)} \tag{3}$$

Choosing a Gamma distribution for the prior allows for analytical solutions to Equation (1), resulting in the posterior distribution presented in Equation (4). As a result of this choice, the posterior also follows a Gamma distribution, maintaining the conjugate structure (Siu & Kelly, 1998).

$$\pi_1(\lambda|S) = \frac{(b_\lambda + T)^{a_\lambda} \lambda^{(a_\lambda + K) - 1} e^{-(b_\lambda + T)\lambda}}{\Gamma(a_\lambda + K)} \tag{4}$$

Given this Bayesian model, the mean failure rate estimate and percentiles (p) are given by Equations (5) and (6) (Siu & Kelly, 1998). A credibility interval can be constructed, for example, using the 5% and 95% percentiles. The interpretation of such an interval is that 90% of the λ's falls within the credibility interval. Since there is no closed-form solution for calculating the percentiles, a common approach is to use an approximation based on the Chi-squared distribution (a special case of the Gamma distribution), where $\chi^2_{p\%, v}$ is the percentile p of the Chi-squared distribution with v degrees of freedom. The Bayesian credible intervals are analogous to confidence intervals and are called posterior probability intervals or posterior credible intervals (Bolstad, 2007; Hamada et al., 2008).

$$\hat{\lambda} = E[\lambda] = \int_0^\infty \lambda \times \pi_1(\lambda|S) d\lambda = \frac{a_\lambda + K}{b_\lambda + T} \tag{5}$$

$$\hat{\lambda}_{p\%} = \frac{1}{2(b_\lambda + T)} \chi^2_{p\%, 2(a_\lambda + K)} \tag{6}$$

The two-stage Bayes approach for the constant failure intensity case is illustrated in Figure 2. As mentioned before, the second stage has an analytical solution due to the conjugate relationship between the Gamma prior and the Poisson likelihood. The problem now relies on the first stage, to determine a_λ and b_λ.

Figure 2 – Classical two-stage Bayesian approach for constant failure intensity function.
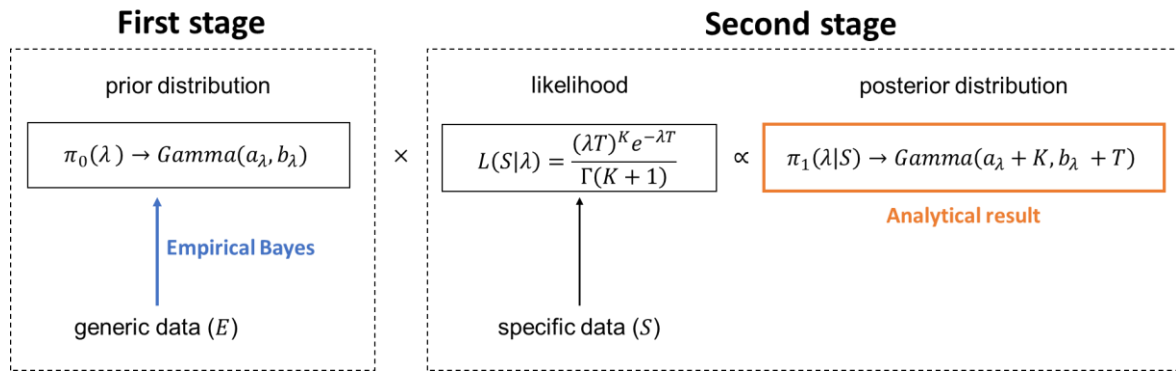
Source: adapted from (Sales da Cunha et al., 2024).

In the Empirical Bayes approach, a_λ and b_λ estimates can be obtained from the Method of Moments by solving Equations (7) and (8), where $\hat{\mu}_\lambda$ and $\hat{\sigma}^2_\lambda$ are the mean and variance of the failure rate λ, obtained directly from the generic data (SINTEF & NTNU, 2015).

$$\frac{a_\lambda}{b_\lambda} = \hat{\mu}_\lambda \tag{7}$$

$$\frac{a_\lambda}{b_\lambda^2} = \hat{\sigma}^2_\lambda \tag{8}$$

Assuming the data can be categorized into NP subpopulations and that for each subpopulation i ($i = 1, \ldots, NP$) there is a record (k_i, t_i), meaning k_i failures within a time interval t_i, a valid estimator for λ_i (failure rate of subpopulation i) is $\hat{\lambda}_i = k_i/t_i$. So, the most commonly used estimators for μ_λ and σ^2_λ are the mean and variance sample of λ_i, in Equation (9) and (10), respectively.

$$\hat{\mu}_\lambda = \bar{\lambda} = \frac{\sum_{i=1}^{NP} \hat{\lambda}_i}{NP} = \frac{\sum_{i=1}^{NP} k_i/t_i}{NP} \tag{9}$$

$$\hat{\sigma}_\lambda^2 = S_\lambda^2 = \sum_{i=1}^{NP} \frac{\left(\hat{\lambda}_i - \hat{\mu}_\lambda\right)^2}{NP - 1} = \sum_{i=1}^{NP} \frac{\left(k_i/t_i - \frac{\sum_{i=1}^{NP} k_i/t_i}{NP}\right)^2}{NP - 1} \tag{10}$$

However, especially in cases involving identical (or very close) $\hat{\lambda}_i$, the sample variance may underestimate the true variance of the failure rate. Similarly, having many cases with $k_i = 0$ may overestimate the mean (Vaurio, 1987). To prevent such issues, some authors have incorporated special terms in the process of estimating the mean and variance of λ. The "special" estimation procedure for μ_λ e σ_λ^2 used in the OREDA database (SINTEF & NTNU, 2015) is described below.

1. Calculate an initial estimate for the mean (Equation (11)), by grouping the data.

$$\hat{\mu}_\lambda^{(1)} = \frac{\sum_{i=1}^{NP} k_i}{\sum_{i=1}^{NP} t_i} \tag{11}$$

2. Calculate S_1 (Equation (12)), S_2 (Equation (13)), V (Equation (14)).

$$S_1 = \sum_{i=1}^{NP} t_i \tag{12}$$

$$S_2 = \sum_{i=1}^{NP} t_i^2 \tag{13}$$

$$V = \sum_{i=1}^{NP} \frac{\left(k_i - \hat{\mu}_\lambda^{(1)} \times t_i\right)^2}{t_i} \tag{14}$$

3. Calculate the variance estimate (Equation (16)).

$$V^* = \left(V - (NP - 1) \times \hat{\mu}_\lambda^{(1)}\right) \times \frac{S_1}{S_1^2 - S_2} \tag{15}$$

$$\hat{\sigma}_\lambda^2 = \begin{cases} V^*, & V^* > 0 \\ S_\lambda^2, & V^* \leq 0 \end{cases} \tag{16}$$

4. Update the mean estimate (Equation (17)).

$$\hat{\mu}_\lambda = \frac{1}{\sum_{i=1}^{NP} \frac{1}{\frac{\hat{\mu}_\lambda^{(1)}}{t_i} + \hat{\sigma}_\lambda^2}} \times \sum_{i=1}^{NP} \left(\frac{1}{\frac{\hat{\mu}_\lambda^{(1)}}{t_i} + \hat{\sigma}_\lambda^2} \times \frac{k_i}{t_i} \right) \tag{17}$$

ISO 14224 (ISO, 2016) also refers to this procedure but proposes a different decision rule for the variance estimation. Kelly & Siu (1998) state that the estimators obtained through Method of Moments are easy to calculate, not requiring the execution of optimization techniques, which motivates its adoption by OREDA (SINTEF & NTNU, 2015) and recommendation in ISO 14224 (ISO, 2016).

While the two-stage Bayesian approach in Figure 2 is widely adopted, its limitation lies in the assumption of constant failure intensity function, which may not be true for systems exposed to degrading factors, e.g., high temperatures, pressure, vibration, corrosive fluids (Taofeek Popoola et al., 2013), as seen in the O&G industry. The next section introduces a likelihood function for a non-constant failure intensity hypothesis.

2.3 The non-constant failure intensity function case

Consider the generic data (E) consisting of data from a population divided into NP subpopulations. Systems within the same subpopulation are homogeneous, while systems across different subpopulations are non-homogeneous. Now, assume that the failure intensity of systems in each subpopulation is modeled by a two-parameter Weibull distribution, with scale parameter α and shape parameter β, as shown in Equation (18), where t denotes the time. The Weibull is flexible, accommodating various failure intensity patterns observed throughout a system's lifecycle as seen in the bathtub curve (Smith, 2005), including the constant rate case ($\beta = 1$).

$$h(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha} \right)^{\beta-1} \tag{18}$$

Since α and β may vary across the subpopulations, it is necessary to model their variability with a PVD. The prior distribution captures the uncertainty in α and β across subpopulations. According to (Gouet et al., 2019) there are three potential approaches for modeling the prior distributions of the Weibull parameters:

1. **Known β ($\beta = \beta_0$) and unknown α:** α is typically modeled by a Gamma distribution. In this case, the likelihood is given by the Weibull model with $\beta = \beta_0$ and the posterior distribution depends solely on α, resulting in an analytically derived posterior distribution.

2. **Discrete β ($\pi_0(\beta_l) = p_l$) and α conditioned on β:** in this case, α is modeled by a Gamma distribution. Making a reparameterization, with $\lambda = \alpha^{-\beta}$, allows for the conjugation of the prior distribution with the likelihood, resulting in an analytically derived posterior distribution, which will be also a Gamma (Soland, 1969).

3. **Independent α and β distributions:** Both α and β are assumed to follow independent Gamma distributions. In this approach, the posterior distribution is not solved analytically, requiring numerical methods (Kundu, 2008).

The first approach is not applicable since the shape parameter β is frequently unknown. The second approach has faced criticism from Kaminskiy & Krivtsov (2005) due to challenges in acquiring the necessary prior information and its practical implementation. Therefore, the third approach proposed by Kundu (2008) is adopted here. While it requires more computational effort, it offers a more feasible and pragmatic solution (Lee et al., 2014).

Now, we will adapt the procedure illustrated in Figure 2 to incorporate a Weibull-based likelihood. Assume that the prior distribution for each Weibull parameter is given by an independent Gamma distribution, as shown in Equation (19). Here, $\pi_0(\theta)$ denotes the Gamma prior distribution with a_θ and b_θ as the shape and scale parameters, respectively, hereafter referred to as hyperparameters. Note that θ can represent either α or β.

$$\pi_0(\theta) = \rho(\theta|a_\theta, b_\theta) = \frac{b_\theta^{a_\theta} \theta^{a_\theta-1} e^{-b_\theta\theta}}{\Gamma(a_\theta)} \tag{19}$$

The first stage involves determining the hyperparameters using generic data, which can be effectively achieved through the application of Empirical Bayes via the Maximum Likelihood Estimate (MLE) method, as outlined by Shultis et al. (1981). Following this, the next step is to establish the Weibull likelihood function, denoted as $L(S|\alpha, \beta)$, where S represents specific data in the form of failure times. It is defined in Equation (20), where t^m represents the m-th failure time.

$$L(S|\alpha, \beta) = \prod_m \left[\frac{\beta}{\alpha} \left(\frac{t^m}{\alpha} \right)^{\beta-1} e^{-\left(\frac{t^m}{\alpha} \right)^{\beta}} \right] \tag{20}$$

Finally, the joint posterior distribution is given by the Bayes' theorem, as shown in Equation (21). The posterior distribution represents the updated knowledge about α and β, the parameters governing the system's failure probability model. This update occurs through the application of the likelihood function of the two-parameter Weibull model, denoted as $L(S|\alpha, \beta)$.

$$\pi_1(\alpha, \beta|S) = \frac{L(S|\alpha, \beta)\pi_0(\alpha)\pi_0(\beta)}{\iint L(S|\alpha, \beta)\pi_0(\alpha)\pi_0(\beta) \, d\beta \, d\alpha} \tag{21}$$

If the posterior distribution was in the same family as the prior distributions, Equation (21) could be solved analytically using the characteristics of the conjugate family. Examples of works that take advantage in conjugate distributions are seen in (Jackson & Mosleh, 2016; Siu & Kelly, 1998; Vaurio, 1987).

However, since this is not the case, analytical approximations and numerical simulations are two primary computational approaches for addressing this problem (Guan et al., 2012). Such approaches include the Laplace approximation (Jerez et al., 2022), highest posterior density estimation (Wu et al., 2006), variational inference (Bhattacharyya, 2021; Ling et al., 2024), Importance Sampling (Jia & Guo, 2022), and MCMC (Erto & Giorgio, 2013; Lee et al., 2014). Among the simulation-based methods, MCMC stands out due to its flexibility and effectiveness in sampling from complex posterior distributions using various sampling strategies (Straub & Papaioannou, 2015). Its versatility and robustness have led to its widespread application in Bayesian inference as seen in (Craiu & Rosenthal, 2014; Karandikar et al., 2014; Kelly & Smith, 2009; René Van Dorp & Mazzuchi, 2004; L. Wang et al., 2017).

Hence, the two-stage Bayes approach for the non-constant failure intensity case is illustrated in Figure 3. Section 3 will detail the methods used for estimating both prior and posterior distributions for this case. Much of the previous research on failure modeling has focused on the assumption that repairs restore a component to a "same as new" condition, treating the process as a renewal process with independently and identically distributed inter-arrival times (BahooToroody et al., 2020). However, less attention has been given to the more realistic assumption that repairs only restore the component to a "same as old" condition (Kelly & Smith, 2009), which will be the focus of this work.

Figure 3 –The two-stage Bayesian approach for non-constant failure intensity.
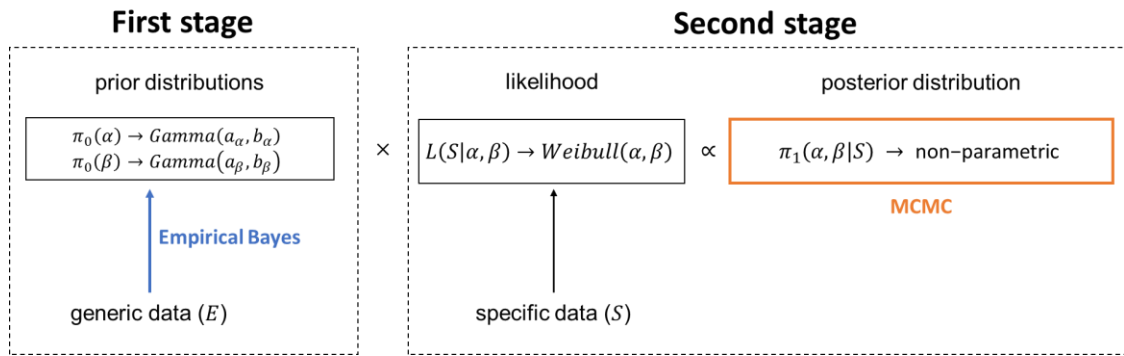
Source: adapted from (Sales da Cunha et al., 2024).

2.3.1 Markov Chain Monte Carlo (MCMC)

MCMC methods have been used by researchers in many scientific fields, including biology, chemistry, computer sciences, economics, engineering, material sciences, physics, and statistics, particularly to solve statistical computation problems related to Bayesian inference (Craiu & Rosenthal, 2014; Liu, 2004). They aim to mimic the sampling process of a probability distribution that takes place in two stages: (i) generation of proposals and (ii) decision for acceptance or rejection of proposals. The primary goal is to ensure that accepted proposals accurately represent the target distribution (Betancourt, 2017).

As the name suggests, MCMC methods possess a Markovian property, which is a fundamental characteristic of the proposal generation process. This begins with selecting an initial point, representing the chain's starting state (Liu, 2004). From this state, a new proposal is generated, depending only on the current state (Robert & Casella, 2004). Then, a decision is made whether to accept the proposal. If accepted, the proposal becomes a sample and the new state of the chain. This process repeats,

with each new proposal generated based on the latest state and evaluated for acceptance. This continues until the desired number of iterations is reached. Unlike the traditional Monte Carlo methods, in which the samples are independent, MCMC samplers yield dependent draws (Craiu & Rosenthal, 2014).

The ideal behavior of the MCMC algorithm can be divided into three phases (Betancourt, 2017):

1. **Search Phase (Figure 4a):** Initially, the algorithm globally searches for the *typical set* — the most representative region of the target probability distribution. During this phase, estimators are highly biased, as the MCMC process has not yet found the main region of interest for sampling.

2. **Discovery Phase (Figure 4b):** Once the typical set is found, the algorithm enters an exploration phase where it begins to gather more representative samples. In this phase, the bias of the initial samples decreases, and the accuracy of the MCMC estimators improves quickly.

3. **Permanence Phase (Figure 4c):** In this final stage, the MCMC process thoroughly explores the typical set. At this point, the improvement in the precision of the estimators continues, but at a slower rate as the algorithm refines its sampling from the target distribution.

These phases describe how MCMC algorithms progress from initial exploration to full convergence, ensuring that the typical set is adequately sampled for reliable parameter estimation.

Figure 4 – The three phases of ideal MCMC behavior, namely: (a) search, (b) discovery, and (c) permanence.

(a)

(b)

(c)
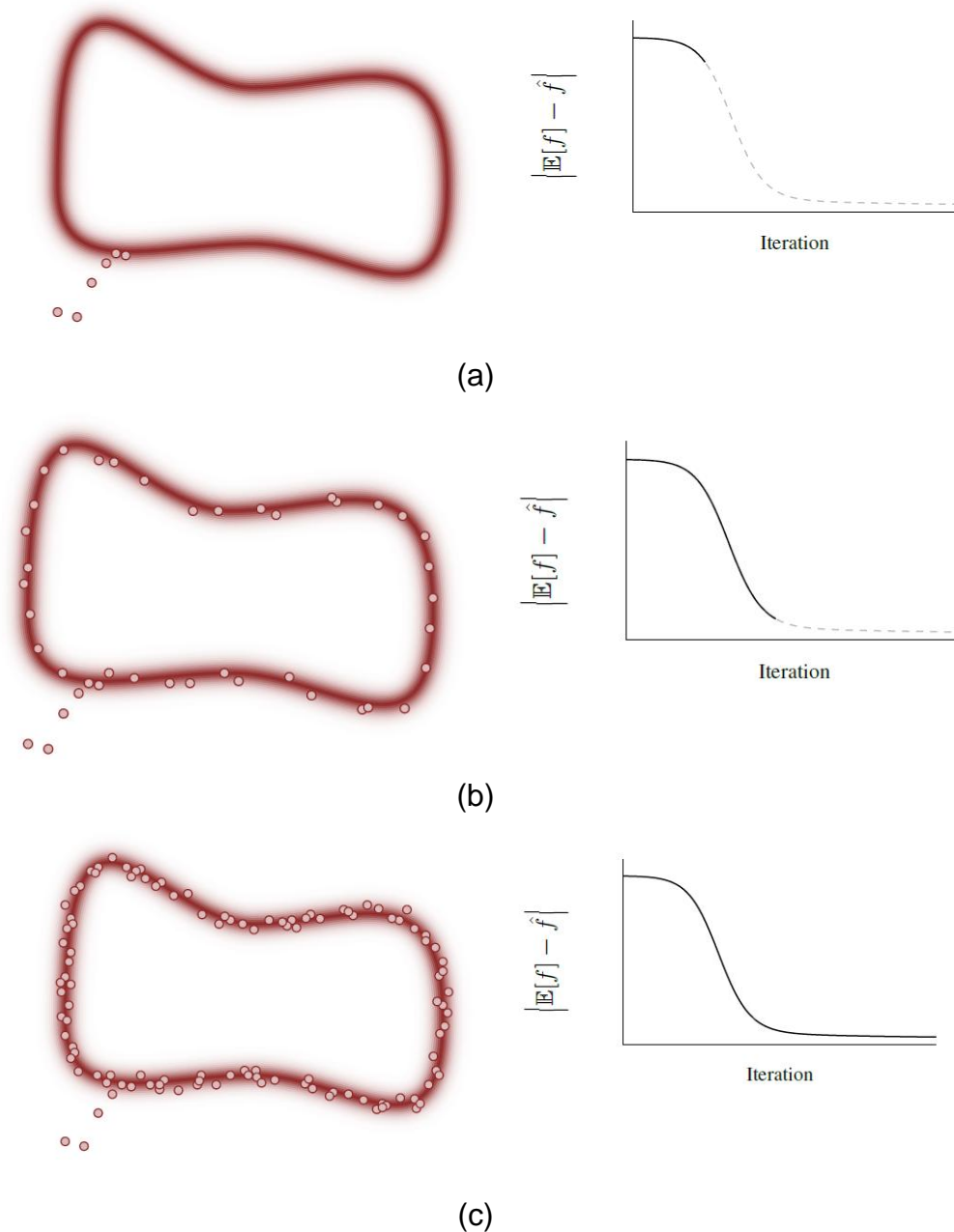
Source: (Betancourt, 2017).

As the states generated during the initial convergence phase of a Markov chain are biased and can negatively affect MCMC estimators, the accuracy of these estimators can be significantly improved by using only samples generated once the chain starts exploring the typical set. This leads to the common practice of introducing a "burn-in" period, during which the initial samples are discarded before calculating MCMC estimators because they are excessively biased toward the (arbitrary) initial value (Craiu & Rosenthal, 2014).

Formally, a sufficient condition to ensure the optimal behavior of MCMC estimators is known as geometric ergodicity (Roberts & Rosenthal, 2004). However, verifying geometric ergodicity theoretically for most problems is extremely difficult. As a result, empirical diagnostics are necessary, the most important of which is the split-$\hat{R}$ statistic (Hoffman & Gelman, 2014). This statistic measures the variation across a set of Markov chains initialized from different points in the parameter space. Issues arising from a lack of geometric ergodicity cause inconsistencies between individual chains, leading to large values of split-$\hat{R}$. Consequently, if the split-$\hat{R}$ deviates significantly from the nominal value of 1, it indicates that geometric ergodicity is not satisfied, and the resulting estimators are likely of poor quality.

The $\hat{R}$ statistic (Vehtari et al., 2021) is computed based on the between-sample variance (Equation (22)) and the within-sample variance (Equation (25)). In this context, M represents the number of Markov chains, and N is the number of samples. A chain m is represented by θ_m, containing $\theta_m^{(n)}$ samples.

$$B = \frac{N}{M-1} \sum_{m=1}^{M} \left(\bar{\theta}_m^{(\cdot)} - \bar{\theta}_{\cdot}^{(\cdot)} \right)^2 \tag{22}$$

$$\bar{\theta}_m^{(\cdot)} = \frac{1}{N} \sum_{n=1}^{M} \bar{\theta}_m^{(n)} \tag{23}$$

$$\bar{\theta}_{\cdot}^{(\cdot)} = \frac{1}{M} \sum_{n=1}^{M} \bar{\theta}_m^{(\cdot)} \tag{24}$$

$$W = \frac{1}{M} \sum_{m=1}^{M} s_m^2 \tag{25}$$

$$s_m^2 = \frac{1}{N-1} \sum_{n=1}^{N} \left(\theta_m^{(n)} - \theta_m^{(\cdot)} \right)^2 \tag{26}$$

The variance of the $\hat{R}$ estimator is a combination of the inter-chain and intra-chain variances (Equation (27)). The $\hat{R}$ statistic is computed through Equation (28). The split-$\hat{R}$ is estimated by calculating the $\hat{R}$ statistic for each half of every chain separately. This process also provides information on the stationarity of the method. By splitting each chain and evaluating the convergence within these halves, this

method ensures that not only is the variation between chains minimized but that the chains themselves have reached a stable distribution, helping to confirm whether the MCMC process has truly converged.

$$\widehat{\text{var}}^{+}(\theta|y) = \frac{N-1}{N}W + \frac{1}{N}B \tag{27}$$

$$\hat{R} = \sqrt{\frac{\widehat{\text{var}}^{+}(\theta|y)}{W}} \tag{28}$$

The result of MCMC simulations is typically a non-parametric posterior distribution, summarized by key statistics such as the mean, variance, and credible intervals (percentiles). These summary statistics capture essential aspects of the posterior distribution, offering insights into the parameters being estimated. For more comprehensive understanding, graphical representations such as density plots or histograms are often preferable, as they present the entire posterior distribution in a visual format, highlighting its shape and any potential skewness or multimodality.

In some cases, parametric estimators for marginal posterior densities can be employed, based on the samples generated from MCMC (Chen et al., 2000). These parametric marginal posterior estimators provide a way to summarize the posterior distribution using a known parametric form, which can sometimes offer computational advantages or better interpretability in complex models.

2.4   The OREDA database

The OREDA project (https://oreda.com/) has been running for more than 43 years, with about 11 O&G companies as members. It is comprised of a database of reliability data collected on Topside and Subsea equipment from offshore and onshore operations. For non-members, part of the data is available in handbooks, with the latest version published in 2015. OREDA database has been used by researchers, practitioners, scholars and other stakeholders in the O&G industry (Animah & Shafiee, 2020), to support safety and reliability analyses during development of new oil fields and improving existing facility operation (Langseth et al., 1998).

The database for any given equipment is structured into three parts: Inventory, Failure, and Maintenance (Mahmoudi, 2021). The Inventory part stores detailed descriptions of equipment units, including technical and environmental data. The

Failure part logs all failure events for each unit. The Maintenance part records both preventive and corrective actions, with corrective maintenance linked to specific failure events, sometimes requiring multiple actions for a single failure (Sandtorv & Thompson, 1996), as exemplified in Figure 5.

Figure 5 –The OREDA database structure.
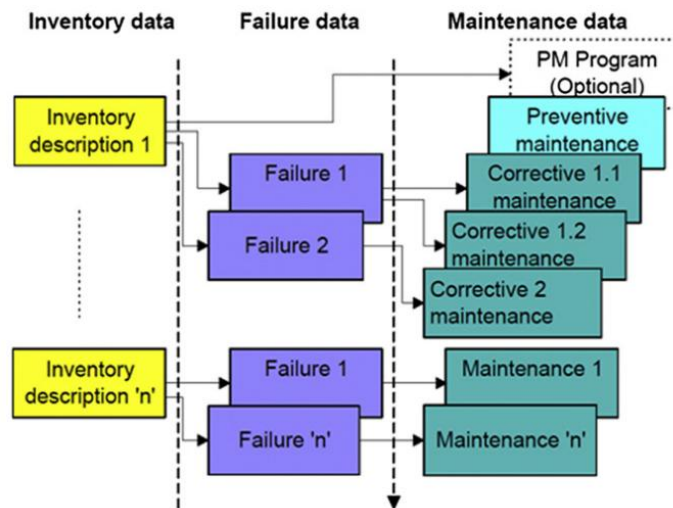
Source: (Hameed et al., 2011).

Specifically, in the Failure part, failures are classified into four groups based on the number and nature of failures, namely: critical, degraded, incipient, and unknown failure (Mahmoudi, 2021; Rausand & Oien, 1996). Failure rates for the different criticality groups can be calculated, based on calendar time or operating time. These analyses can be applied to all systems within a specific category (e.g., all gas turbines) or refined to focus on specific groups (e.g., gas turbines >5 MW driving compressors). (Sandtorv & Thompson, 1996). It is commonly assumed that all components within a group share the same failure intensity. However, despite being grouped by type, size, and operating mode, environmental and maintenance conditions introduce variability, making complete homogeneity in reliability across the group unrealistic (Porn, 1996).

The OREDA project has led to key developments, such as creating standards and guidelines for reliability data collection and analysis, including contributing to an ISO standard. They have also produced specialized software, enhanced understanding of reliability data needs, and developed expertise in data collection processes. Additionally, the project has fostered collaboration and knowledge-sharing between companies and countries, benefiting the broader industry (Sandtorv & Thompson, 1996). In fact, two oil companies reported total savings of USD 70 million

by opting for alternative designs over the original ones, with reliability data playing a key role in informing these decisions (Hameed et al., 2011).

OREDA provides information about the intermediate failure times, the failure mechanism, and the maintenance activity (Langseth & Lindqvist, 2006). Traditional OREDA database analysis typically assumes a constant failure intensity function (SINTEF & NTNU, 2015), due to the nature of the data, which reflects the number of failures occurring over a given time period. This assumption entails the use of the model described in Section 2.2. This work seeks to adapt the model outlined in Section 2.3 for non-constant failure intensity, making it suitable for application to OREDA's failure count data. This required not only revising the original formulation, but also developing tailored methods to effectively address the specific characteristics and challenges of the data and the more realistic assumptions adopted here to customize it to the O&G context.

3 MATHEMATICAL MODEL DEVELOPMENT

In the two-stage Bayesian procedure proposed by Kaplan (1983), data provided by generic databases ("generic data", E) can be used as prior knowledge and updated by incorporating system-specific failure data ("specific data", S), yielding a posterior distribution (Kaplan, 1983; Kodoth et al., 2020; Lin, 2002). This approach is crucial for predicting the reliability performance of systems with limited data, as it leverages the data from similar systems or applications, as thoroughly discussed in Section 2.1. This chapter is based on the article (Sales da Cunha et al., 2024).

Consider a generic data consisting of a population divided into NP subpopulations. Each subpopulation i, where $i = 1 \dots NP$, consists of NS_i systems that share the same probabilistic failure behavior. In other words, systems within the same subpopulation are homogeneous, while systems across different subpopulations are non-homogeneous.

Generic reliability databases, such as the OREDA database (SINTEF & NTNU, 2015), typically contains failure count data. Therefore, the evidence from each subpopulation, denoted by E_i, with $E = \{E_1, \dots, E_{NP}\}$, consists of sets of failure count data pairs $\{(k_{ij}, t_{ij})\}$, where k_{ij} is the number of failures over an observation time t_{ij}, for system j of subpopulation i with $j = 1 \dots NS_i$. Figure 6 illustrates how the failure count data is collected, as demonstrated in the case of 1 subpopulation with 3 individual systems, where tf represents failure times and tc are censored times. However, the OREDA database does not provide the information regarding the failure times, only the sets of failure count data pairs $\{(k_{ij}, t_{ij})\}$.

Figure 6 – A timeline illustrating the failures of 3 hypothetical systems in the same subpopulation 1, where t_{ij}^m represents the mth time of failure of system j from subpopulation i.
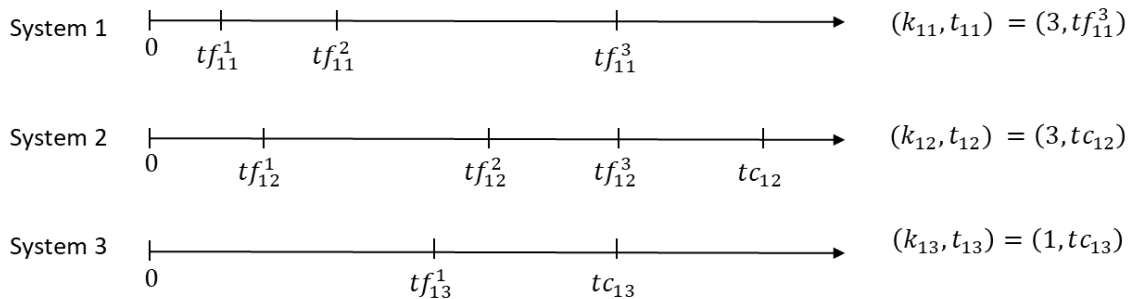
Source: adapted from (Sales da Cunha et al., 2024).

Failure time data provide a significantly richer source of information compared to failure count data. While failure counts indicate how many times a system failed over a period, they do not capture *when* those failures occurred. This distinction is clearly illustrated by comparing System 1 and System 2 in Figure 6: although both systems experienced the same number of failures (three), the timing of those events reveals distinct reliability behaviors. By preserving the temporal aspect of failure events, time-to-failure data allows for a more accurate and nuanced understanding of system performance and degradation over time.

In this context, the prior distribution estimation for failure count data, based on the PVD, relies on three fundamental assumptions: (i) homogeneity within each subpopulation, (ii) heterogeneity among subpopulations, and (iii) a constant failure intensity function (Droguett et al., 2004). The first two assumptions arise from the observation that systems within the same subpopulation operate under similar conditions and share similar design features, which may differ from those in other subpopulations. The third assumption, a constant failure intensity function, on the other hand, while commonly adopted due to the analytical convenience it offers with the Poisson distribution for failure count data, is somewhat limiting. This assumption has been adopted by OREDA (SINTEF & NTNU, 2015) and recommended in ISO 14224 (ISO, 2016) for its analytical tractability, as described in Section 2.2.

However, this study aims at generalizing the third assumption to account for cases with non-constant failure intensity by using the Weibull distribution, as described in Section 2.3 and illustrated in Figure 3. This introduces a challenge, as the data format differs from what a standard Weibull model typically handles—specifically, time-to-failure data. Consequently, appropriate methods are developed to address this situation effectively and they will be detailed below. Section 3.1 presents the methods involved in estimating the prior distribution, followed by Section 3.2 where the posterior distributions estimation will be detailed. The entire implementation of the methodology was conducted in Python.

The main contribution of this work lies in the prior distribution estimation, which required revising the original formulation of constant failure intensity and developing tailored methods to address the specific nuances of the dataset. The use of the Weibull distribution, known for its flexibility in modeling different failure behaviors, in

combination with the NHPP to reflect minimal repair scenarios, represents a novel integration of techniques in the context of Bayesian Inference.

3.1 Prior Distribution Estimation

Empirical Bayes methods can be used to estimate the hyperparameters that define the independent prior distributions of α and β, given the evidence from generic data (E). The challenge here is that the evidence of each subpopulation E_i is given in the form of $\{(k_{ij}, t_{ij})\}$ pairs, where k_{ij} is the number of failures over an observation time t_{ij}, for system j of subpopulation i with $j = 1 \dots NS_i$. This data format deviates from what a typical Weibull model takes as input, i.e., time to failure.

To apply Empirical Bayes, the maximum marginal likelihood can be approximated using the Maximum Likelihood Estimate (MLE) (Shultis et al., 1981). Assume $P(E_i|\alpha, \beta)$ represents the probability of observing E_i for specific α and β values. Since α and β are uncertain, the likelihood is obtained by integrating over all possible values of α and β, making it unconditional with respect to the Weibull parameters. Hence, the likelihood of the evidence of each subpopulation E_i given the hyperparameters is presented in Equation (29).

$$L(E_i|a_\alpha, b_\alpha, a_\beta, b_\beta) = \iint P(E_i|\alpha, \beta)\pi_0(\alpha)\pi_0(\beta)\, d\alpha\, d\beta \tag{29}$$

Equation (29) is an indefinite integral with no closed solution. However, it can be approximated via Importance Sampling (Barbu & Zhu, 2020). By generating n samples of α and β through $\pi_0(\alpha)$ and $\pi_0(\beta)$, respectively, Equation (30) provides an approximation of $L(E_i|a_\alpha, b_\alpha, a_\beta, b_\beta)$.

$$\hat{L}(E_i|a_\alpha, b_\alpha, a_\beta, b_\beta) \cong \frac{1}{n}\sum_{l=1}^{n} P(E_i|\alpha_l, \beta_l) \leftrightarrow \alpha_l \sim \rho(\alpha|a_\alpha, b_\alpha), \beta_l \sim \rho(\beta|a_\beta, b_\beta) \tag{30}$$

To avoid variability of this estimation, Latin Hypercube Sampling (LHS) will be used as the sampling strategy for α and β. LHS ensures samples cover the entire range of possible values for α and β by dividing their range into N disjoint partitions and randomly generating n/N samples in each of them (Sheikholeslami & Razavi, 2017). In practice, as shown in Figure 7, LHS samples the Cumulative Distribution Function (CDF) of θ, where $\theta = \alpha$ or $\theta = \beta$, with a range between 0 and 1. These sampled CDF values can then be transformed into θ when applying the quantile

function of the Gamma distribution (Martinez et al., 2013). For reference, in the example of Figure 7 we have $N = 8$ and $n = 8$.

Figure 7 – LHS sampling strategy
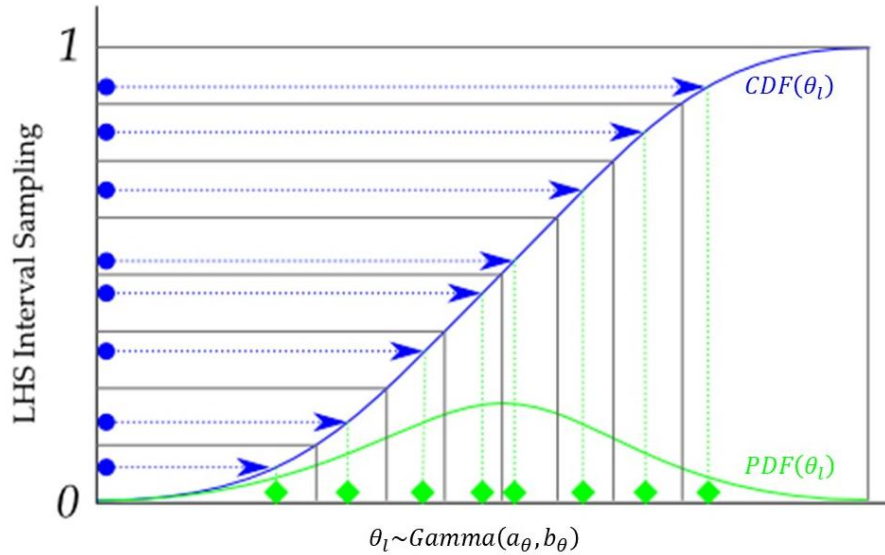
Source: adapted from (Martinez et al., 2013).

Figure 8 compares the sampling of the Gamma distribution with parameters $a_\theta = 7.93$ and $b_\theta = 0.0021$ using LHS versus random number generation, considering $n = 100$ and $N = 50$. The comparison is based on the p-value from the Kolmogorov-Smirnov test (Corder & Foreman, 2014), which evaluates the empirical distributions produced by both methods against the original Gamma distribution. The sampling was executed 4 independent times with the same parameters. Visually, the empirical distributions generated by LHS more closely resemble the original Gamma distribution. This is further supported by the p-values consistently equal to 1 for LHS and varying between approximately 0.2 to 0.9 for the random samples. This result corroborates the implementation of LHS for consistency in the results, even with a small sample.

Figure 8 – LHS versus random sampling of Gamma distribution
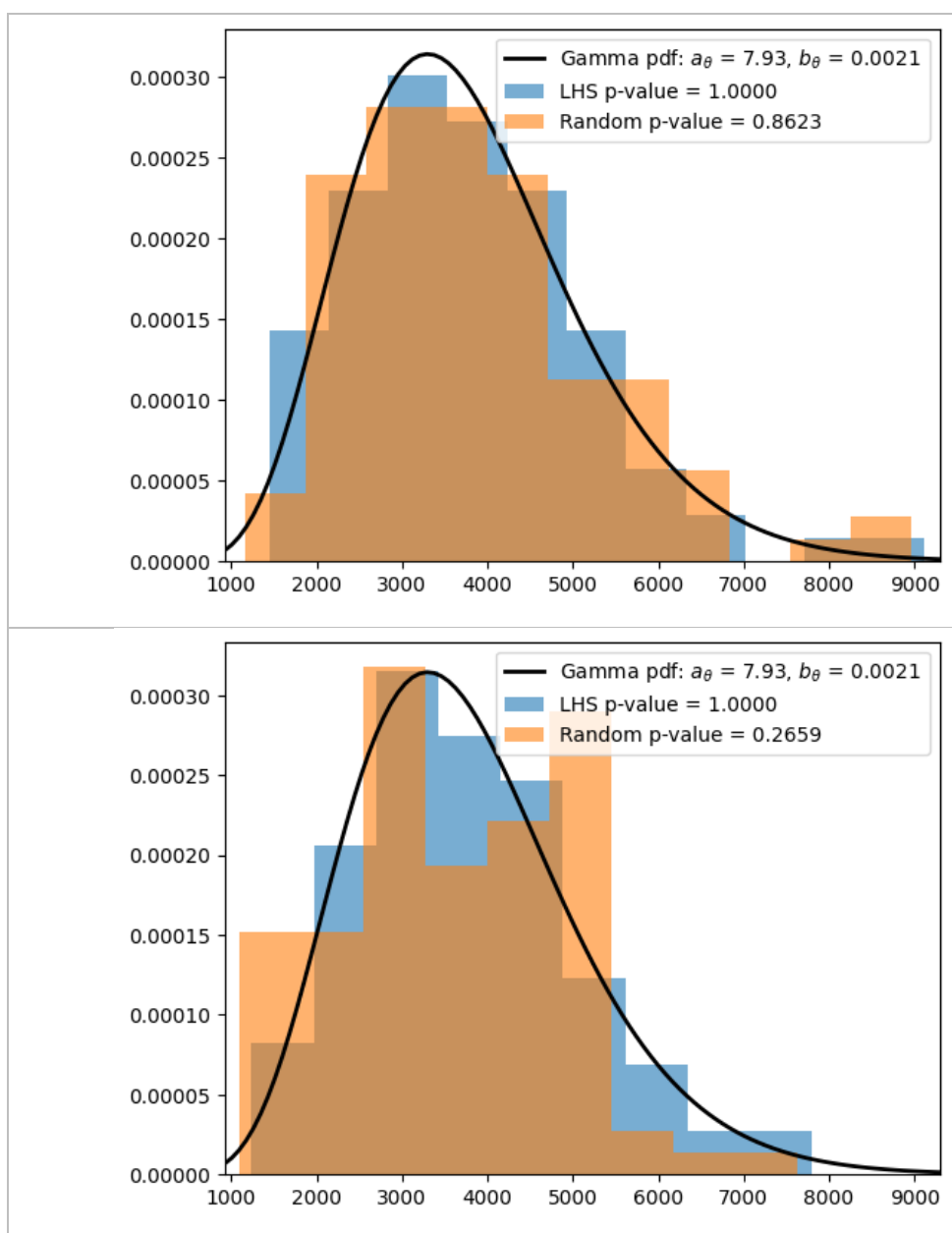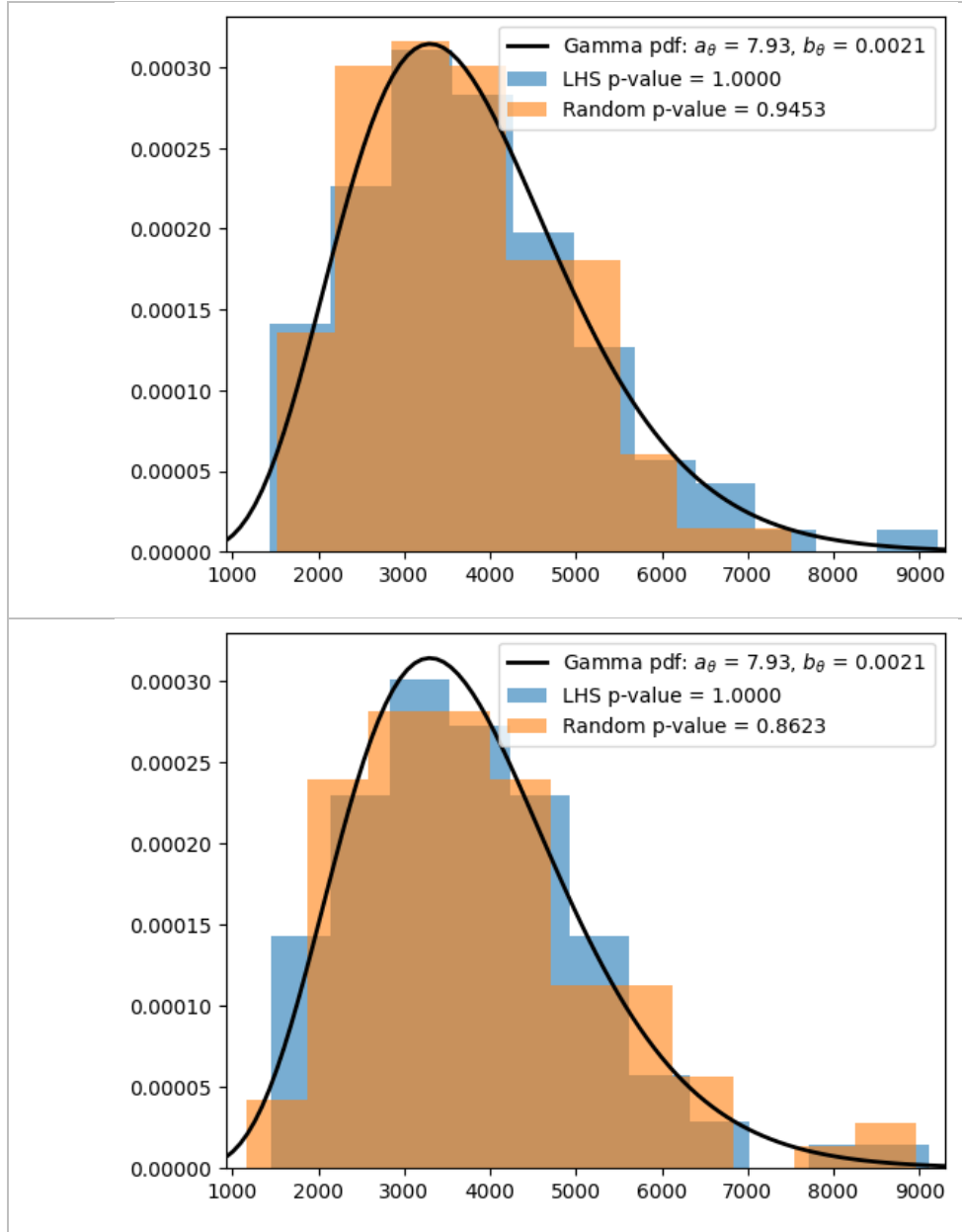
Source: The author (2025).

Looking closely at Equation (30), the term $P(E_i|\alpha_l, \beta_l)$ represents the Weibull counting process (McShane et al., 2008), as the evidence E_i is given as failure count data. Since the Poisson model assumes a constant failure intensity function, it is unsuitable for cases where failure intensity changes over time. Therefore, estimating this process requires a more appropriate method. For instance, if the systems undergo perfect repairs after each failure, the situation can be modeled as a renewal process. An effective way to approximate this counting process is through Monte Carlo simulation, with the corresponding algorithm outlined in Figure 9.

Figure 9 – Weibull counting process via Monte Carlo simulation

```
Algorithm 1 Weibull counting process via Monte Carlo simulation
    Input: α_l, β_l
    Initialize N ← 0
    for m = 1 to M do
        Initialize cumulative failure time X ← 0
        Initialize failure count n ← 0
        while X ≤ t_ij do
            Generate Failure Time: x ~ Weibull(α_l, β_l)
            Update: X ← X + x
            if X ≤ t_ij then
                n ← n + 1
            end if
        end while
        if n = k_ij then
            Update: N ← N + 1
        end if
    end for
    Return P(E_i|α_l, β_l) = N/M
```

Source: The author (2025).

The problem is that this method is extremely time-consuming, especially when attempting to reduce the variability inherent in the simulation process by increasing the number of iterations (Mundform et al., 2011). An alternative approach proposed by (McShane et al., 2008) involves a numerical procedure that approximates the accumulated conditional Weibull distribution using a power series. While this method addresses both the computational time and variability issues, tests revealed some inconsistencies in the convergence of the series.

Although the assumption of perfect repair is widely applied, it often yields improper results (BahooToroody et al., 2020). For systems that are degrading or improving, a time-dependent model is more appropriate. If we assume that the systems undergo minimal repair, the process can be modeled as a NHPP (das Chagas Moura et al., 2014; Kelly & Smith, 2009). By considering that the times to failure follow a Weibull distribution, an analytical solution for this model exists (Basu & Rigdon, 2001), which reduces both the computational cost and the variability of the results compared to Monte Carlo simulation. Then, the estimation of $P(E_i|\alpha_l, \beta_l)$ via NHPP for the Weibull distribution is provided in Equation (31).

$$P(E_i|\alpha_l, \beta_l) = \prod_{j=1}^{NS_i} \frac{\left[\left(\frac{t_{ij}}{\alpha_l}\right)^{\beta_l}\right]^{k_{ij}} e^{-\left(\frac{t_{ij}}{\alpha_l}\right)^{\beta_l}}}{k_{ij}!} \tag{31}$$

Now that we have established all the methods involved in the estimation of Equation (29), which determines the likelihood of the evidence of each subpopulation E_i, we can move on to analyzing the entire population. So, Equation (32) shows the likelihood of the evidence from the entire population E obtained by multiplying Equation (29) across the NP subpopulations.

$$L(E|a_\alpha, b_\alpha, a_\beta, b_\beta) = \prod_{i=1}^{NP} L(E_i|a_\alpha, b_\alpha, a_\beta, b_\beta) \qquad (32)$$

Taking the log-likelihood is a common practice in MLE because it helps with numerical stability by converting the products in Equation (32) into sums (Casella & Berger, 2002). This transformation prevents numerical underflow, which can occur when multiplying very small numbers, as is often the case with probabilities. Therefore, the MLE procedure seeks to find the set of hyperparameters that maximizes Equation (33).

$$log\, L(E|a_\alpha, b_\alpha, a_\beta, b_\beta) = \sum_{i=1}^{NP} \log L(E_i|a_\alpha, b_\alpha, a_\beta, b_\beta) \qquad (33)$$

Given the complexity of the log-likelihood in Equation (33) the Particle Swarm Optimization (PSO) meta-heuristic is here adopted because of its effectiveness in searching optimal solutions of non-linear equations in a real-valued search space (Bratton & Kennedy, 2007). Using Equation (33) as the objective function, PSO aims to find the set of hyperparameters $\{a_\alpha, b_\alpha, a_\beta, b_\beta\}$ that maximizes this function. To achieve this, it is necessary to define the search space for $\{a_\alpha, b_\alpha, a_\beta, b_\beta\}$, i.e. lower and upper boundaries for the hyperparameters. However, defining the search space can be challenging, as the hyperparameters do not provide direct insight into the distribution's scale or range.

The Method of Moments can help overcome this challenge by changing the decision variable for this problem. According to the Method of Moments, one can find $\{a_\theta, b_\theta\}$ through the expected value ($E[.]$) and the standard deviation ($S[.]$) of the Gamma distribution, as shown in Equation (34) and Equation (35) where $\theta = \alpha$ or $\theta = \beta$.

$$a_\theta = \frac{E[\theta]^2}{S[\theta]^2} \tag{34}$$

$$b_\theta = \frac{E[\theta]}{S[\theta]^2} \tag{35}$$

Hence, it is more intuitive to define the search space based on the expected value and the standard deviation of α and β, rather than based directly on the hyperparameters. Therefore, the decision variables for PSO will be the set $\{E[\alpha], S[\alpha], E[\beta], S[\beta]\}$, and the necessary transformations in Equation (34) and Equation (35) will be applied to obtain the corresponding set of hyperparameters $\{\hat{a}_\alpha, \hat{b}_\alpha, \hat{a}_\beta, \hat{b}_\beta\}$.

Figure 10 presents the flowchart of the implemented PSO algorithm, with gBest representing the global best solution obtained so far by any particle (Hamdan, 2008). The connections between the nomenclature of PSO and the MLE model are described below for clarity.

- Particles' positions: the values assigned to $\{E[\alpha], S[\alpha], E[\beta], S[\beta]\}$, which are the decision variables.
- Search space: lower (lb) and upper (ub) boundaries for each decision variable, here addressed as $[lb_{E(\alpha)}, ub_{E(\alpha)}, lb_{S(\alpha)}, ub_{S(\alpha)}, lb_{E(\beta)}, ub_{E(\beta)}, lb_{S(\beta)}, ub_{S(\beta)}]$.
- Fitness: objective function defined by the log-likelihood in Equation (33).
- Stop criteria: the maximum number of iterations.

Figure 10 – PSO algorithm
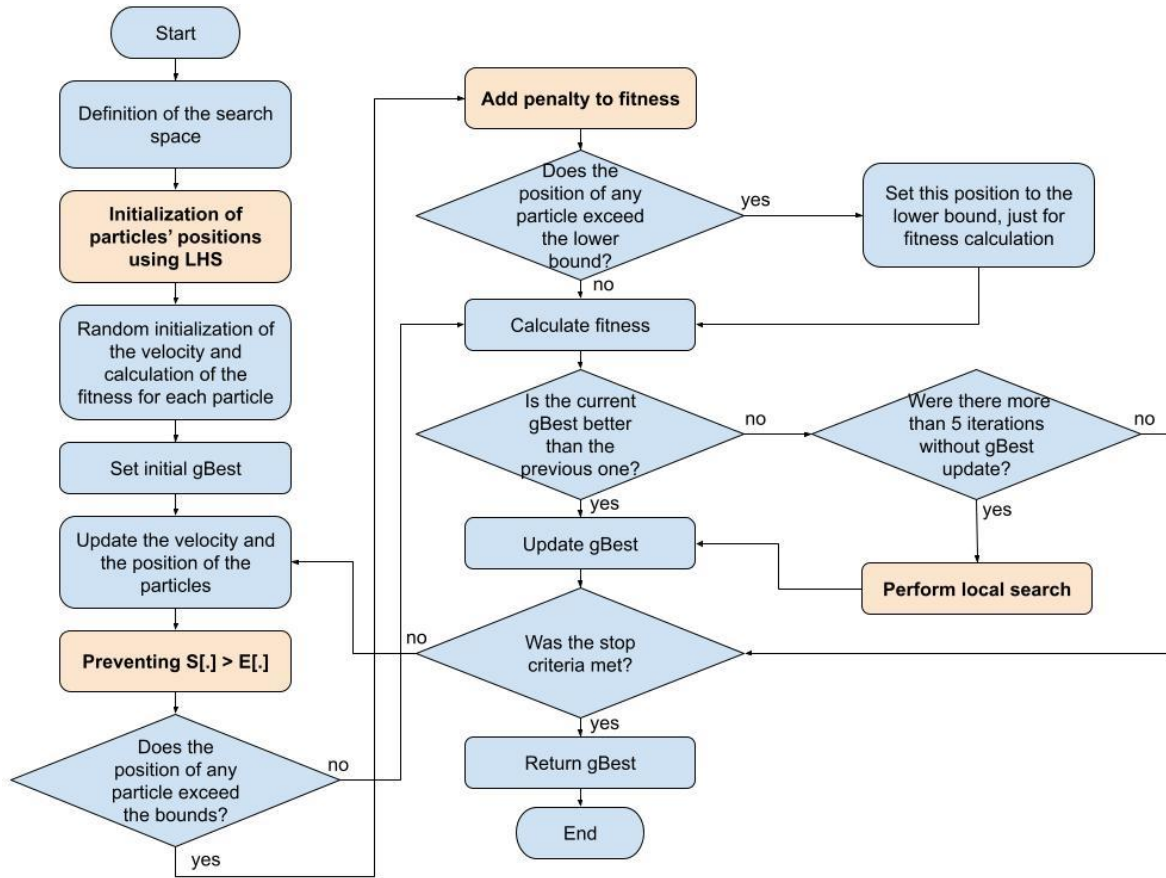
Source: adapted from (Sales da Cunha et al., 2024).

The PSO algorithm was implemented using the PySwarm Python library (https://pythonhosted.org/pyswarm/), with custom modifications made to the source code. The key differences from the traditional PSO algorithm are highlighted in orange in Figure 10 and described below.

- Initialization of particles' positions using LHS: The search space for each decision variable is divided into np intervals, with one value generated from each interval using LHS. This method ensures comprehensive coverage of the search space, as addressed before.

- Add penalty to fitness: A dynamic penalty is added to the fitness of particles that end up outside the search space, discouraging them from venturing into these regions. It is dynamic because the current iteration (it) is factored in, as shown in Equation (36) where x represents the vector of decision variables $\{E[\alpha], S[\alpha], E[\beta], S[\beta]\}$.

$$penalty = (it + 1) \left\{ H_{ub}(x) \cdot \frac{|x - ub_x|}{ub_x} + H_{lb}(x) \cdot \frac{|x - lb_x|}{lb_x} \right\} \tag{36}$$

$$H_{lb}(x) = \begin{cases} 1, & x < lb_x \\ 0, & x \geq lb_x \end{cases} \tag{37}$$

$$H_{ub}(x) = \begin{cases} 1, & x > ub_x \\ 0, & x \leq ub_x \end{cases} \tag{38}$$

- Perform local search: If the algorithm remains stuck at the same global best (gBest) position for 5 iterations, a local search is performed using the Nelder-Mead optimization method, as suggested by (Noel, 2012). This helps fine-tuning the solution when the global search is stuck.

- Preventing $S[.] > E[.]$: When $a_\theta < 1$, the Gamma distribution is asymptotic to both the vertical and horizontal axes, as shown in Figure 11 with $a_\theta = 0.16$ and $b_\theta = 0.0001$ as an example. Generating Gamma-distributed random values with a shape parameter $a_\theta < 1$ is a well-known challenge in the literature, with only a few methods available to address it (Tanizaki, 2008; Xi et al., 2013). These methods typically rely on majorization functions and the acceptance–rejection principle (Kundu & Gupta, 2007). In practice, when $a_\theta < 1$, the generated random values for θ can become very small. If this occurs with α, it implies that simulating the time to failure would result in an extremely short lifespan, causing the system to spend more time in a failure state than in operation. Consequently, the simulation would take an excessively long time to reach the time horizon and conclude, which is not a desirable or an expected outcome. Note that $a_\theta < 1$ corresponds to situations where $S[.] > E[.]$. Therefore, if any particle shows $S[.] > E[.]$, its fitness is set to infinity, effectively discouraging exploration of these regions.

Figure 11 – Example of Gamma distribution with $a_\theta < 1$.
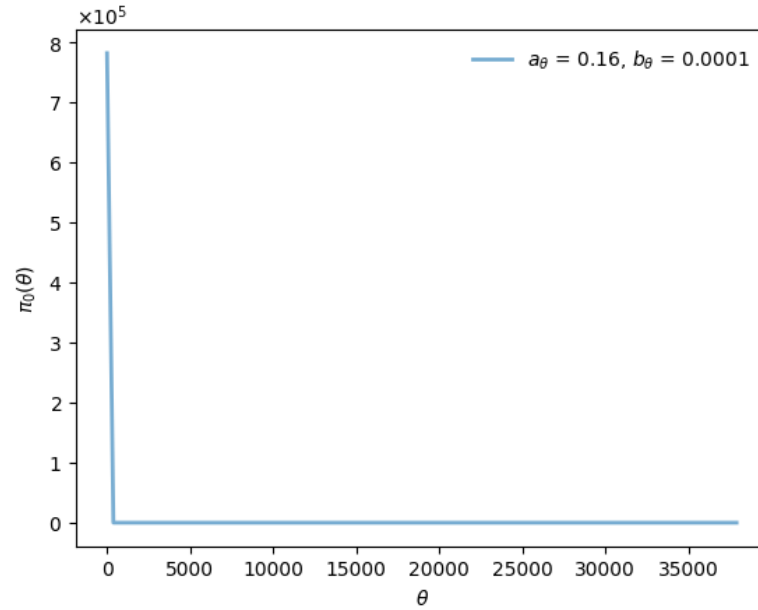
Source: The author (2025).

The main contribution of this work lies in the prior distribution estimation. The use of the Weibull distribution, known for its flexibility in modeling different failure behaviors, in combination with the NHPP to reflect minimal repair scenarios and all the methods involved in estimating and optimizing the likelihood function, represents a novel integration of techniques. This approach fills a gap in the literature and provides a more accurate and practical framework for reliability estimation in the O&G industry.

Since the model is subject to various sources of variability—such as the calculation of the likelihood of each subpopulation via Importance Sampling (Equation (30)) and the inherent variability of the PSO algorithm—conducting multiple independent PSO runs and comparing the results is a good strategy. To assess the quality of the PSO results, it is essential to use an appropriate metric. While the likelihood is an obvious metric, using an additional metric for further validation is helpful. Here, the NRMSE (das Chagas Moura et al., 2014) will be used as the metric to evaluate the accuracy of the prior distribution estimation across these runs (Equation (39)).

$$NRMSE = \sqrt{\frac{\sum_{i=1}^{NP} \sum_{j=1}^{NS_i} \{k_{ij} - E[N(t_{ij})]\}^2}{\sum_{i=1}^{NP} \sum_{j=1}^{NS_i} (k_{ij})^2}}$$

(39)

The NRMSE is defined in Equation (39), where k_{ij} is the actual number of failures for each system j of subpopulation i and $E[N(t_{ij})]$ is the expected number of failures over an observation time t_{ij} estimated via Monte Carlo simulation, following the steps outlined in Figure 12.

Figure 12 – Expected number of failures via Monte Carlo simulation.

Algorithm 2 Expected number of failures via Monte Carlo simulation

Input: $a_\alpha, b_\alpha, a_\beta, b_\beta$
Initialize $N \leftarrow 0$
for $m = 1$ to M **do**
    Sample: $\alpha \sim Gamma(a_\alpha, b_\alpha)$ and $\beta \sim Gamma(a_\beta, b_\beta)$
    for $z = 1$ to Z **do**
        Initialize cumulative failure time $X \leftarrow 0$
        Initialize failure count $n \leftarrow 0$
        while $X \leq t_{ij}$ **do**
            Generate Failure Time: $x \sim Weibull(\alpha, \beta)$
            Update: $X \leftarrow X + x$
            if $X \leq t_{ij}$ **then**
                $n \leftarrow n + 1$
            end if
        end while
        Update: $N \leftarrow N + n$
    end for
end for
Return $E[N(t_{ij})] = \frac{N}{Z \times M}$

Source: The author (2025).

The drawback of this metric is its high computational cost, as it requires two nested Monte Carlo loops: one for sampling the parameters α and β, and another for sampling the failure times. To enhance the quality of the estimates, LHS was also used here for generating the α and β samples. Each PSO run yields a potential solution for the Gamma prior distribution of α and β. In addition to evaluating the log-likelihood of each potential solution, the minimum NRMSE will be used as the criterion to identify the best solution, which will then be selected as the prior distribution for the next step. Figure 13 illustrates how the methods discussed so far are integrated to estimate the prior distribution.

Figure 13 – Prior distribution estimation process.
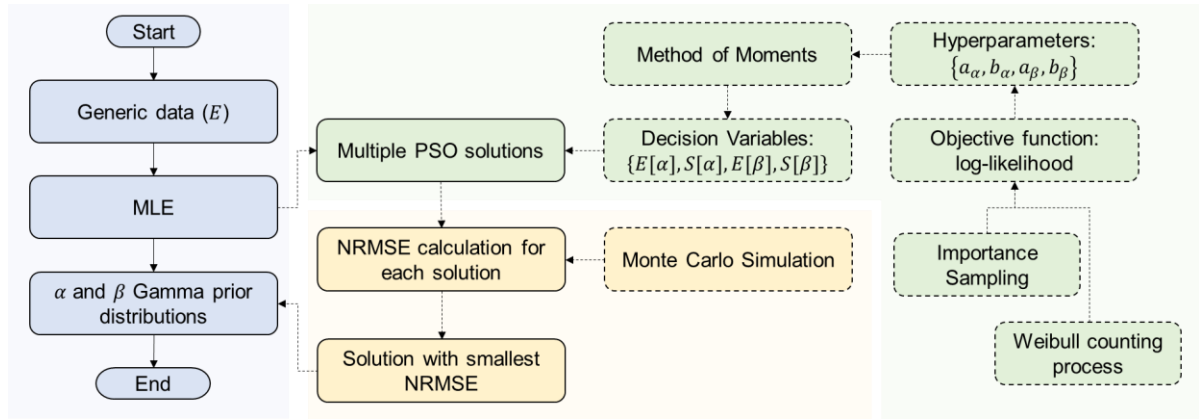
3.2 Posterior Distribution Estimation

Referring back to the Bayesian model for the non-constant failure intensity case described in Section 2.3, Equation (21) determines the posterior distribution $\pi_1(\alpha, \beta|S)$, representing the updated knowledge about α and β, given the specific data S. Here, the specific data is in the form of failure times, as illustrated in Figure 14.

Figure 14 – A timeline illustrating a hypothetical specific data in the form of failure times, where t^m represents the mth time of failure.
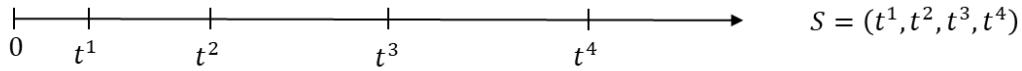
This update occurs through the application of the likelihood function of the Weibull model, denoted as $L(S|\alpha, \beta)$ (see Equation (21)), to the prior distributions of α and β, estimated following the methodology described in Section 3.1. Since α and β follow independent continuous distributions, there is no family of conjugate distributions with the Weibull likelihood (Erto & Giorgio, 2013). Therefore, Equation (21) does not have an analytical solution and deriving the posterior distribution requires the use of numerical methods. MCMC algorithms can be used to derive a non-parametric posterior distribution (Bolstad, 2009).

For the implementation of MCMC, the Stan platform (Carpenter et al., 2017) was leveraged through its Python interface, PyStan (Riddell et al., 2021). Stan employs advanced MCMC techniques, such as Hamiltonian Monte Carlo (HMC) (Betancourt, 2017; Neal, 2011) and no-U-turn sampling (NUTS) (Hoffman & Gelman, 2014), which provide efficient approaches to Bayesian analysis. One of the key advantages of NUTS is its ability to automatically adjust parameters, leading to faster convergence and

reduced computational costs. Due to these benefits, NUTS was adopted here for analysis.

Stan reports the $\hat{R}$ statistic, which is the maximum value between the rank-normalized split-$\hat{R}$ and the rank-normalized folded-split-$\hat{R}$, offering insights into convergence (see Section 2.3.1 for details). Although NUTS handles many parameters automatically, some parameters were manually set based on guidelines and practices from the literature (Gelman et al., 2013). The number of iterations was set to 50,000 iterations, considering the complexity of the problem. To monitor convergence and assess the behavior of the chains, four separate chains were considered. The initial portion of each chain, known as the warmup or "burn-in" phase, was set to 25% of the total iterations, during which samples are discarded to help the chains reach their stationary distribution.

MCMC simulations generally produce a non-parametric posterior distribution, summarized using key statistics such as the mean, variance, and credible intervals. In some instances, parametric estimators for marginal posterior densities can be applied to MCMC samples, offering computational benefits or improved interpretation, especially in complex models (Chen et al., 2000). The parametric distribution considered in this study were: Gamma, Weibull, and Lognormal. The goodness of fit was evaluated using the Kolmogorov-Smirnov (KS) test (Corder & Foreman, 2014), which computes the D statistic, indicating the maximum distance between the empirical and theoretical cumulative distribution functions. The p-value from the KS test, calculated based on the D statistic, will be used to determine the parametric distribution that better describes the posterior distributions of α and β.

An example of such approach is illustrated in Figure 15, where MCMC was applied to a hypothetical case. The prior distribution of α and β was set with hyperparameters $a_\alpha = 7.4$, $b_\alpha = 0.002$, $a_\beta = 10.42$, $b_\beta = 6.446$, and the following failure times were used as specific data $S = (2562.3268, 3564.5464, 7919.6538, 8981.7632, 18670.2207)$ for the likelihood function in Equation (20). In this example, the Lognormal distribution provided the best fit for α and the Gamma distribution for β, according to the p-values of the KS test. The p-values above 0.05 indicate that the hypothesis that the non-parametric distribution

resulting from MCMC (green histograms in Figure 15) and the fitted parametric distributions originate from the same distribution is not rejected.

Figure 15 – Non-parametric posterior distributions of (a) α and (b) β, fitted to Gamma, Weibull, and Lognormal distributions.

(a)

(b)
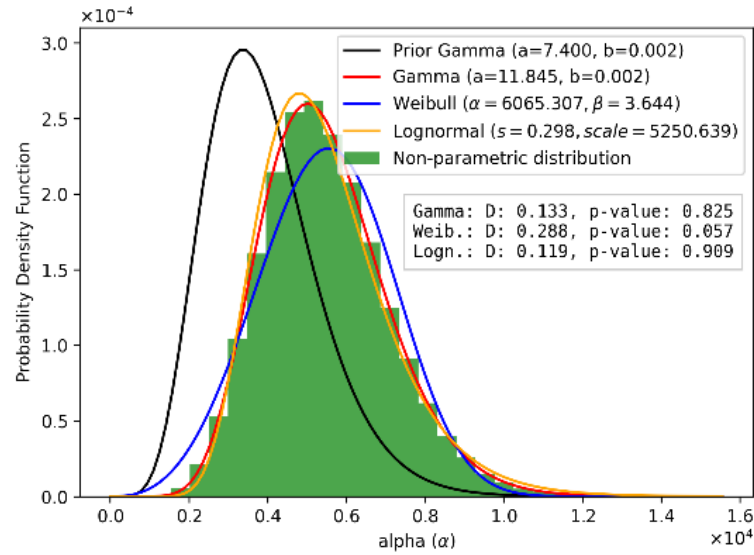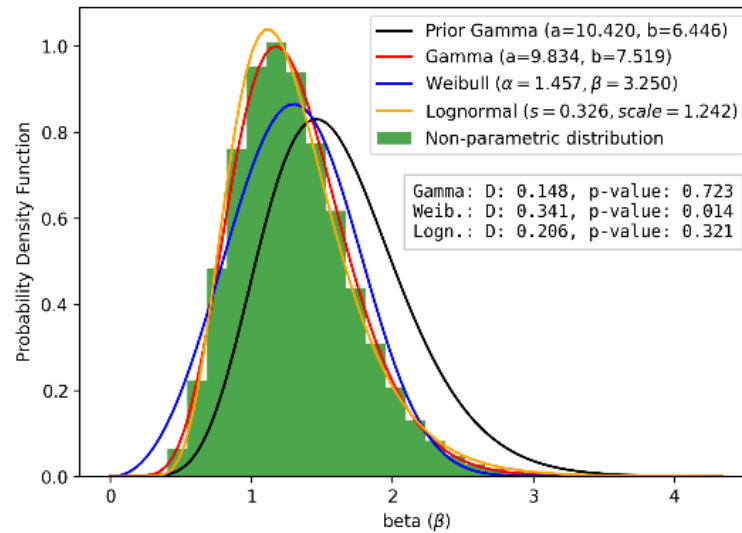
Source: The author (2025).

In summary, the application of Bayesian inference through MCMC techniques allowed for the estimation of non-parametric posterior distributions for the Weibull model parameters α and β, overcoming the lack of conjugate priors. This is a well known problem in the literature and by leveraging the computational capabilities of the

Stan platform and its advanced sampling algorithms like NUTS, the study achieved efficient and accurate parameter estimation. Fitting a parametric distribution to the MCMC samples, validated by the Kolmogorov-Smirnov test, is an addition to traditional approaches as it enhances interpretability and provides the means for integration with further reliability analysis.

4 EXPERIMENTAL EVALUATION

4.1 Prior Distribution Estimation

This section presents the tests performed for the prior distribution estimation model described in Section 3.1. To evaluate the quality of the estimation, simulated data will be used, providing a basis for comparison. The simulation process will be detailed, and the parameters used will be presented. The main results will be shown along with a sensitivity analysis, varying the amount of available data to assess the model's robustness.

4.1.1 Data Simulation Process

To assess the model's ability to estimate prior distributions, synthetic data sets were generated in the form of paired entries (k_{ij}, t_{ij}). The synthetic data will be used as generic data in the prior distribution estimation model. This approach allows for direct comparison between the model's results and the theoretical distribution from which the data was generated. The simulation process begins with the specification of the following parameters: $\{a_\alpha^T, b_\alpha^T, a_\beta^T, b_\beta^T\}$ (hyperparameters of the theoretical prior distribution), NP (number of subpopulations), NS_i (number of systems per subpopulation i), and G (number of groups per subpopulation).

Assuming that $\alpha \sim Gamma(a_\alpha^T, b_\alpha^T)$ and $\beta \sim Gamma(a_\beta^T, b_\beta^T)$, NP pairs (α_i, β_i) are generated, each representing a subpopulation i (where $i = 1 \dots NP$). This means that all equipment within a subpopulation share the same failure time distribution, while different subpopulations exhibit non-homogeneous characteristics, as defined in Section 2.3. Each subpopulation is then divided into G groups, with each group g containing $\frac{NS_i}{G}$ systems, observed over a time t_g (for $g = 1 \dots G$). For each system j (where $j = 1 \dots NS_i$) in subpopulation i, failure times are randomly generated up to t_g following a Weibull distribution with parameters (α_i, β_i). This subgrouping within subpopulations ensures the capture of stochastic failure behavior across different observation periods, offering a comprehensive view of the parameters' distributions.

Next, the number of failures is counted to form the pairs (k_{ijg}, t_g), which are equivalent to (k_{ij}, t_{ij}), with the group index g used here primarily for illustration in the simulation process. Figure 16 illustrates this simulation process for $NP = 3$, $NS_i = 15 \, \forall i$ and $G = 3$. With this simulation scheme, it is possible to compare the result of

the prior distribution estimation model with the theoretical prior distribution that generated the data, i.e., $\{a_\alpha^T, b_\alpha^T, a_\beta^T, b_\beta^T\}$.

Figure 16 – Data simulation example, with $NP = 3$, $NS_i = 15$ $\forall i$ and $G = 3$.

$\alpha \sim Gamma(a_\alpha^T, b_\alpha^T)$ $\beta \sim Gamma(a_\beta^T, b_\beta^T)$

(α_1, β_1) (α_2, β_2) (α_3, β_3)

	Subpopulation 1	Subpopulation 2	Subpopulation 3
system 1	$(k_{1,1,1}; t_1)$	$(k_{2,1,1}; t_1)$	$(k_{3,1,1}; t_1)$
system 2	$(k_{1,2,1}; t_1)$	$(k_{2,2,1}; t_1)$	$(k_{3,2,1}; t_1)$
system 3	$(k_{1,3,1}; t_1)$	$(k_{2,3,1}; t_1)$	$(k_{3,3,1}; t_1)$
system 4	$(k_{1,4,1}; t_1)$	$(k_{2,4,1}; t_1)$	$(k_{3,4,1}; t_1)$
system 5	$(k_{1,5,1}; t_1)$	$(k_{2,5,1}; t_1)$	$(k_{3,5,1}; t_1)$
system 6	$(k_{1,6,2}; t_2)$	$(k_{2,6,2}; t_2)$	$(k_{3,6,2}; t_2)$
system 7	$(k_{1,7,1}; t_2)$	$(k_{2,7,1}; t_2)$	$(k_{3,7,1}; t_2)$
system 8	$(k_{1,8,2}; t_2)$	$(k_{2,8,2}; t_2)$	$(k_{3,8,2}; t_2)$
system 9	$(k_{1,9,2}; t_2)$	$(k_{2,9,2}; t_2)$	$(k_{3,9,2}; t_2)$
system 10	$(k_{1,10,2}; t_2)$	$(k_{2,10,2}; t_2)$	$(k_{3,10,2}; t_2)$
system 11	$(k_{1,11,3}; t_3)$	$(k_{2,11,3}; t_3)$	$(k_{3,11,3}; t_3)$
system 12	$(k_{1,12,3}; t_3)$	$(k_{2,12,3}; t_3)$	$(k_{3,12,3}; t_3)$
system 13	$(k_{1,13,3}; t_3)$	$(k_{2,13,3}; t_3)$	$(k_{3,13,3}; t_3)$
system 14	$(k_{1,14,3}; t_3)$	$(k_{2,14,3}; t_3)$	$(k_{3,14,3}; t_3)$
system 15	$(k_{1,15,3}; t_3)$	$(k_{2,15,3}; t_3)$	$(k_{3,15,3}; t_3)$

Source: The author (2025).

It is crucial to evaluate the impact of dataset size on the results, as larger datasets tend to produce estimates that closely align with the theoretical prior distribution. However, given that data are often scarce, it is essential to assess the model's robustness as the sample size decreases. So, the proposed methodology was assessed with tests generated following the procedure described above, by varying the synthetic data sets size. Each test will be addressed by its size as (NP, NS_i). As an illustration, the example in Figure 16 would be addressed as (03,15).

The estimated prior distributions will be compared with the theoretical one through: visual inspection of the distribution graph, the log-likelihood value and the

NRMSE metric discussed in Section 3.1. Beyond that, an additional comparison is suggested. This involves contrasting the $\left(\frac{\sum_j k_{ij}}{NS_i}, \frac{\sum_j t_{ij}}{NS_i}\right)$ pairs from the generic data with the $\left(\frac{\sum_j \hat{k}_{ij}}{U}, \frac{\sum_j \hat{t}_{ij}}{U}\right)$ pairs generated using the resulting prior distribution estimation, i.e. estimated parameters $\{\hat{a}_\alpha, \hat{b}_\alpha, \hat{a}_\beta, \hat{b}_\beta\}$. The same simulation procedure as described above will be used to generate the data with $\{\hat{a}_\alpha, \hat{b}_\alpha, \hat{a}_\beta, \hat{b}_\beta\}$, and this data will be addressed as "estimated data". The estimated data will consist of 100 subpopulations ($U = 100$). The Kolmogorov-Smirnov test (Corder & Foreman, 2014) will be used to determine whether the distribution of the mean number of failures across subpopulations in the generic data matches that of the estimated data. This serves as an additional metric to measure the quality of the prior estimation procedure.

4.1.2 Parameters definition

The parameters of the prior distribution model are presented in Table 1 (for detailed descriptions, please refer to the List of Symbols). Arbitrary values for the hyperparameters of the theoretical prior distribution $\{a_\alpha^T, b_\alpha^T, a_\beta^T, b_\beta^T\}$ were selected to produce typical values for α e β, ensuring realistic modeling. With the values for the theoretical prior distribution defined in Table 1, α is defined in a range approximately from 1320 to 7272 and β approximately from 0.71 to 2.68. The values for t_g were based on expert knowledge suggesting these intervals as critical points for failure analysis and practical validity. Moreover, selecting evenly spaced time intervals allows for a clear analysis of failure trends, facilitating statistical modeling and assessments.

Table 1 – Parameters.

Category	Parameter	Value
Data Simulation	$\{a_\alpha^T, b_\alpha^T, a_\beta^T, b_\beta^T\}$	$[8, 0.0022, 13, 8.5]$
Data Simulation	t_g	$[2000, 4000, 6000, 8000, 10000]$
PSO	np	50
PSO	c_1	2.05
PSO	c_2	2.05
PSO	w	0.729843
PSO	it	100
PSO	$[lb_{E(\alpha)}, ub_{E(\alpha)}]$	$[2 \cdot z, 4 \cdot z]$
PSO	$[lb_{S(\alpha)}, ub_{S(\alpha)}]$	$[cs_{min} \cdot \sqrt{z}, cs_{max} \cdot \sqrt{z}]$
PSO	$[lb_{E(\beta)}, ub_{E(\beta)}]$	$[0.5, 5]$
PSO	$[lb_{S(\beta)}, ub_{S(\beta)}]$	$[0.5, 5]$

Source: The author (2025).

The PSO parameters—such as np, c_1, c_2, and w —were selected based on established practices in the field, as these values are commonly recognized for their stability. For more information on these parameters, please refer to Bratton & Kennedy (2007). The number of iterations was determined through preliminary experimentation to ensure effective convergence of the optimization process in a timely manner, and this analysis will be presented later on.

Additionally, the PSO heuristic approach requires defining the search space that satisfies the condition in Equation (40). For the scale parameter α, a case-specific search space was defined, since α is related to the characteristic life of equipment. The search space is defined such that $S[\alpha] \in [cs_{min} \cdot \sqrt{z}, cs_{max} \cdot \sqrt{z}]$, where $z = \frac{\sum_i \sum_j t_{ij}}{\sum_i \sum_j k_{ij}}$. The idea is that z serves as a rough estimate of the mean time between failures. This approach allows the PSO model to dynamically adjust the search intervals for each new data set while ensuring that the estimated distributions have suitable variance for use as prior distributions.

$$\{lb_{E(\alpha)}, lb_{S(\alpha)}, lb_{E(\beta)}, lb_{S(\beta)}\} \le \{E[\alpha], S[\alpha], E[\beta], S[\beta]\}$$
$$\le \{ub_{E(\alpha)}, ub_{S(\alpha)}, ub_{E(\beta)}, ub_{S(\beta)}\}$$

(40)

A sensitivity analysis of the search space parameters for α will be detailed in Section 4.1.3. On the other hand, β controls the failure intensity behavior, which is influenced by underlying physical and operational constraints. Thus, a well-defined range can capture the likely behavior of the system without needing a broad or dynamically varied search space. The value of β usually falls within a known range, characterizing the following failure behaviors (Nelson, 1982):

- $\beta < 1$: decreasing failure intensity (infant mortality or early-life phase).
- $\beta = 1$: constant failure intensity, representing the Exponential distribution.
- $\beta > 1$: increasing failure intensity (wear-out phase).

In practical applications, it is unlikely to observe β values above five. Limiting the search space for β improves computational efficiency without sacrificing accuracy, as the likely β values can be reasonably anticipated based on prior experience or domain knowledge.

4.1.3 PSO search space assessment

The initial analysis involves evaluating the PSO's performance by varying the search space parameters for α. Recall that the search space for $S[\alpha]$ is defined by the constants $\{cs_{min}, cs_{max}\}$. Different values for these constants are here considered: $\{10,55\}$, $\{20,65\}$, $\{30,75\}$, $\{40,95\}$. The focus is on varying the constants that determine the bounds for the standard deviation of α, as these are more challenging to fine-tune.

This analysis was executed with a round of test data sets, by varying the data set size. Each test will be addressed by its size as (NP, NS_i), with NP and NS_i varying between 5, 25 and 50 $\forall i = 1, \dots, NP$. Thus, we set $\{a_\alpha^T, b_\alpha^T, a_\beta^T, b_\beta^T\}$ as shown in Table 1 and use these hyperparameters to generate the biggest test set following the procedure described in Section 4.1.1, in this case (50,50), and then sample data from this test to the others, so smaller tests are contained on the bigger ones. Considering that $G = 5$, the other tests follow the same logic presented above in Figure 16, that is, for the test (_,25) there are 5 devices in each of the 5 groups, and for the test (_,05) there is 1 device in each of the 5 groups, regardless of the number of subpopulations considered. Testing was structured this way to assess the model's performance in scenarios with limited data and to evaluate its behavior as data availability increases.

The prior distribution estimation model was executed five times for each test size. The results are detailed in Appendix A. Here, results are primarily compared using the NRMSE to determine if the quality of the solution is influenced by the search space, which in turn affects the selection of the best prior distribution for subsequent steps in the methodology. Table 2 summarizes the results using a two-factor ANOVA test (Montgomery & Runger, 2013). The first factor is the bounds of the search space, and the second is the dataset size. The test examines three hypotheses: whether the means of the columns differ, whether the means of the rows differ, and whether there is an interaction between the factors. The p-value in Table 2 represents the outcome of the two-factor ANOVA test.

Table 2 – Summary of two factor ANOVA test for varying search space bounds.

Source of variation	F_{crit}	F	p-value
Bounds of the search space	2.00325	1.16659	0.32334
Dataset size	2.66744	0.94379	0.42127
Interaction among factors	1.59323	0.94501	0.54179

Source: The author (2025).

There is no evidence to reject either null hypothesis, as the F-values do not exceed the critical value F_{crit}. Therefore, the NRMSE values remain consistent despite changes in the bounds of the search space, demonstrating the model's flexibility and robustness in identifying suitable prior distributions. Consequently, the values $\{20,65\}$ were selected, as they produced superior visual results.

4.1.4 Results

The proposed methodology underwent a series of tests generated as described in Section 4.1.1. Each test set is identified by its size, denoted as (NP, NS_i). Initially, NP was set to 5 and NS_i assumed the values 5, 25, or 50 $\forall i = 1, ..., NP$. These selections were made to focus on scenarios with limited data while also assessing the model's performance under conditions of increased data availability.

For each test set, the prior distribution estimation model was executed 20 independent times (runs) to evaluate the convergence and robustness of the model, given that the results are subject to variability as mentioned in Section 3.1. This number of runs was chosen as it provides a sufficient sample size to investigate variability in the results. Figure 17 displays the results for each test, comparing the distribution of α and β that generated the data ("theoretical distribution"), indicated by the black dotted curve, with the results of all runs in terms of the log-likelihood in Equation (33) (log-L),

p-value, and NRMSE in Equation (39). The p-value is obtained through the Kolmogorov-Smirnov test, used to ascertain whether the mean number of failures across subpopulations in the generic data and the estimated data adhere to the same distribution, as described in Section 4.1.1.

Figure 17 – Prior estimation results for (a) (05,05), (05,25), and (c) (05,50).

(a)

(b)

(c)
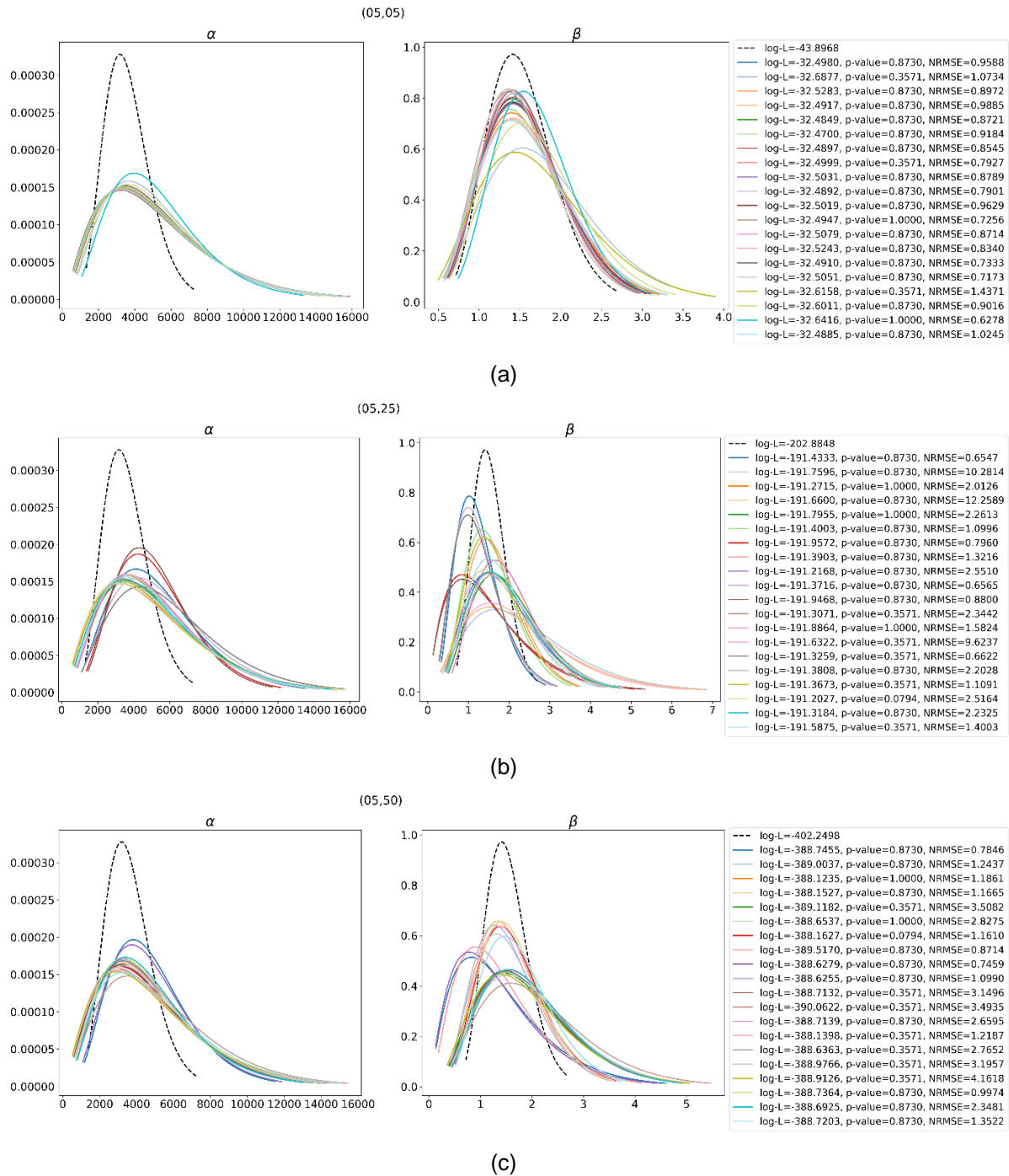
Source: The author (2025).

There is some variability among the resulting distributions, even though they are very close in terms of log-likelihood. This variability may stem from the similarity in log-

likelihood values across different sets of hyperparameters, which causes the search algorithm to explore various regions of the hyperparameter space, leading to nearly identical log-likelihood outcomes. Additionally, the nature of the data—failure count data, which is inherently less informative—can contribute to this variability.

Despite this variability, the results consistently outperform the theoretical distribution in terms of log-likelihood, as the goal is to maximize this metric. Given the small sample size, it's possible that other distributions might model the data more effectively compared with the theoretical distribution. In particular, the available data is limited to failure counts over intervals, which inherently involves a loss of information. The proposed approach accounts for this limitation by incorporating an appropriate likelihood function and tailored sampling methods, which better capture the characteristics of the observed data. Even though the type of data and the small sample size introduces some variability, the mode of the posterior distributions for both α and β remains stable across different runs, demonstrating consistency when compared to the theoretical model.

For all tests, the hypothesis that the generic data sample and the estimated sample (across each PSO run) belong to the same distribution is not rejected, as evidenced by the p-values in Figure 17. In fact, most p-values exceed 0.3, with some approaching 1. This reinforces the effectiveness of the proposed methodology, which assumes Weibull-distributed equipment failure times subject to minimal repair with failure count data.

Table 3 presents the mean and standard deviation (std) of key values and metrics obtained from the experiments. The std of the log-likelihood was consistently low across all datasets in the 20 runs. This result aligns with the nature of the algorithm searching for solutions in a continuous space, where different combinations of hyperparameter values often yield very similar log-likelihood results. Consequently, variations in the resulting distribution across different runs do not cause a considerable variation on the log-likelihood result. Additionally, it is noteworthy that for the smallest dataset (05,05), each run took an average of 706 seconds, with larger datasets requiring more time, averaging around 1124 seconds per run.

Table 3 – Result's mean and standard deviation (std).

Test	metric	$\hat{a}_\alpha$	$\hat{b}_\alpha$	$\hat{a}_\beta$	$\hat{b}_\beta$	log-L	p-value	NRMSE	time
(05,05)	mean	2.7677	0.0005	8.6163	5.3681	-32.5257	0.8083	0.8930	706.8100
	std	0.3411	0.0001	1.2807	0.9019	0.0602	0.1983	0.1696	89.2320
(05,25)	mean	3.2846	0.0006	4.3758	2.4971	-191.5106	0.7234	2.9224	981.7141
	std	0.8362	0.0002	1.2027	0.9138	0.2468	0.2867	3.4526	278.8055
(05,50)	mean	2.9733	0.0006	4.5338	2.5928	-388.7517	0.6655	1.9968	1124.1563
	std	0.5878	0.0001	1.2995	0.8472	0.4653	0.2959	1.1136	158.6732

Source: The author (2025).

Figure 18 shows the evolution of the optimal log-likelihood found by the PSO algorithm over the course of 100 iterations. Note that although this is a maximization problem, the algorithm was coded for minimization; hence, the objective function is represented as the negative log-likelihood and the desired behavior is to see it decreasing over the iterations.

Figure 18 – Negative log-likelihood evolution throughout PSO iterations for (a) (05,05), (05,25), and (c) (05,50).

(a)

(b)

(c)
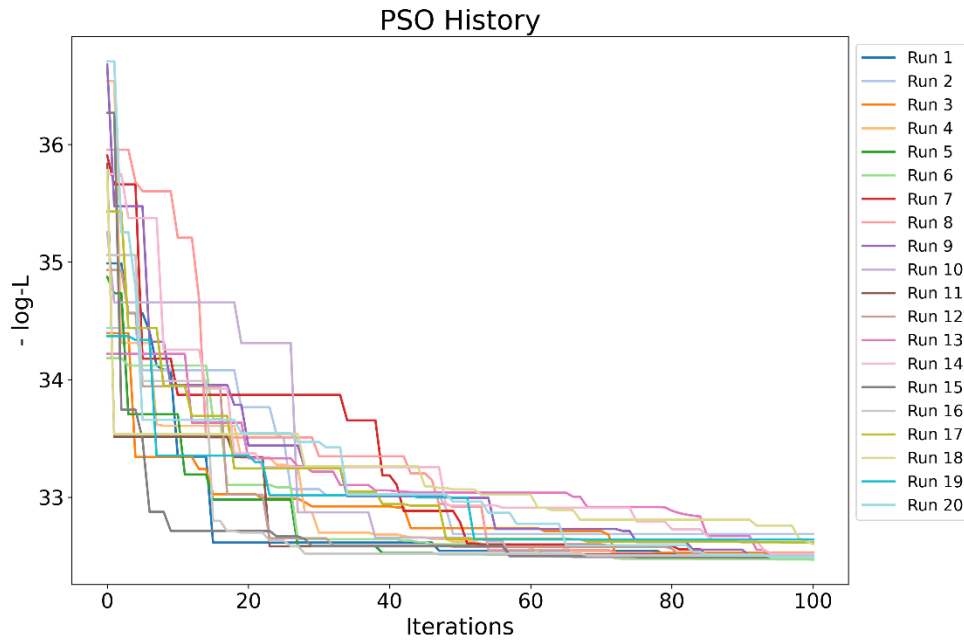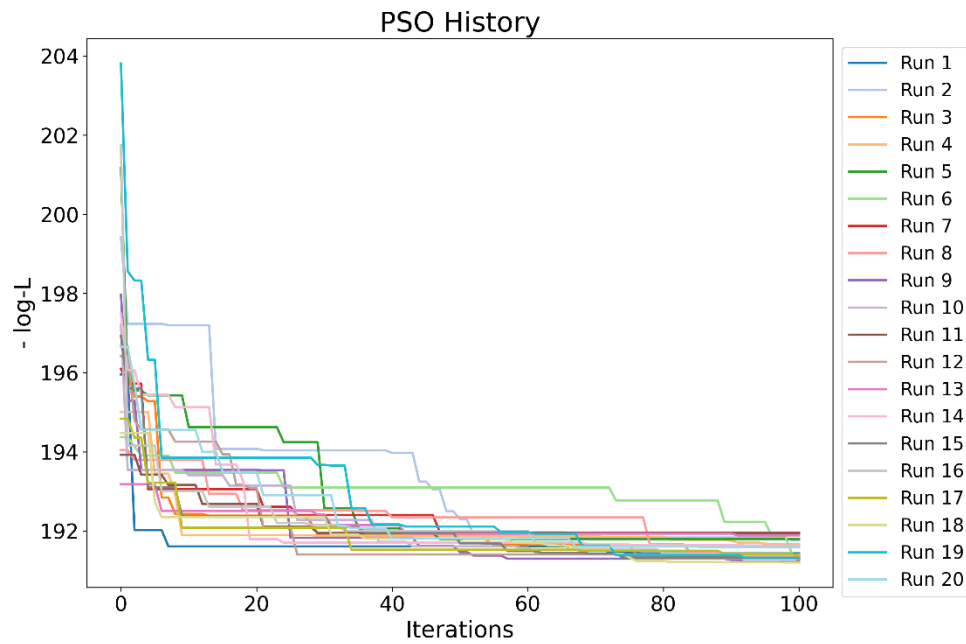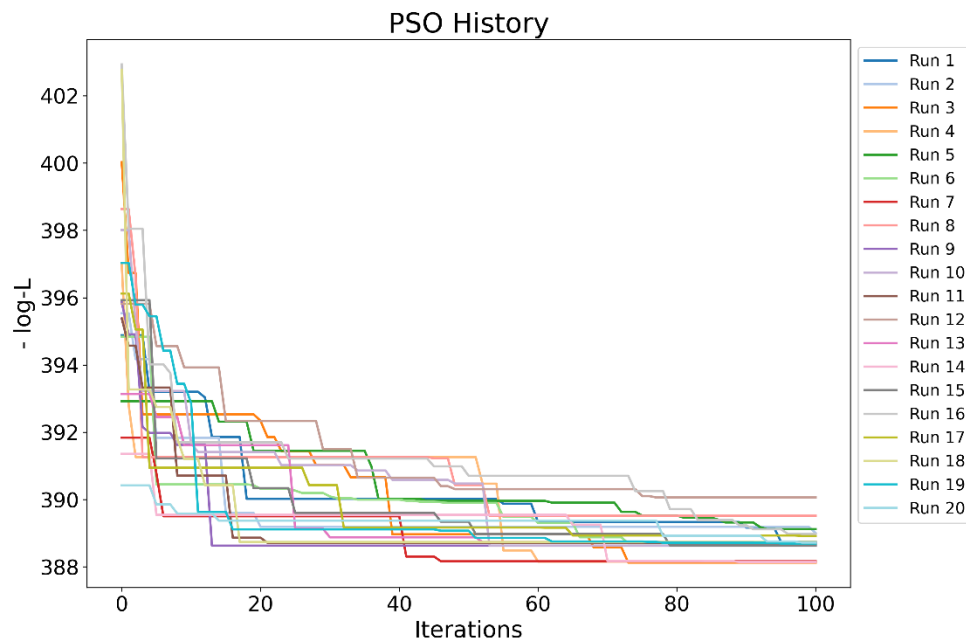
Source: The author (2025).

From Figure 18, defining a stopping criterion based solely on the number of iterations is challenging. Significant changes occur in the initial iterations, but only marginal reductions are seen sporadically over subsequent iterations. These changes are more abrupt, with more pronounced curve adjustments in the case of the (05,50) configuration. However, increasing the number of iterations further extends the already lengthy runtime, which averages nearly 20 minutes to obtain a result for the (05,50) configuration with just 100 iterations. However, an additional stopping criterion could

be introduced: halting the search after a certain number of consecutive iterations where results remain unchanged. This strategy is called patience and it offers an alternative to simply increasing the iteration count.

While the prior distribution estimation model was executed in 20 independent runs for each test, only one solution should be selected. Deciding based on log-likelihood and p-value is challenging as they may lead to opposite conclusions. Additionally, in some instances (see Figure 17), more than one solution could yield the same p-value, complicating the use of such criterion. Thus, the NRMSE serves as the primary decision criterion due to its dimensionless nature, quantifying the proximity between empirical data (generic data) and expected values (see Section 3.1 for more details). Therefore, the solution with the smallest NRMSE is chosen to proceed for the posterior distribution estimation.

4.1.5 Sensitivity Analysis

In addition to the results presented above, larger data sets were also simulated, with 25 and 50 subpopulations and varying number of systems per subpopulation as before. Although this amount of data is not expected in real scenarios, the goal here is to evaluate the model robustness. For each test, 5 independent runs were executed to estimate the prior distributions and to evaluate the convergence and robustness of the PSO. Figure 19 displays the results for the test sets comprise of more than 5 subpopulations, since the results with 5 subpopulations have previously been addressed. Here again the theoretical distribution is indicated by the black dotted curve, and the results are expressed in terms of the log-likelihood (log-L), p-value, and NRMSE.

Figure 19 – Prior estimation results for (a) (25,05), (b) (25,25), (c) (25,50), (d) (50,05), (e) (50,25), (f) (50,50).

(a)

(25,25)

(b)

(25,50)

(c)

(50,05)

(d)

(e)

(f)
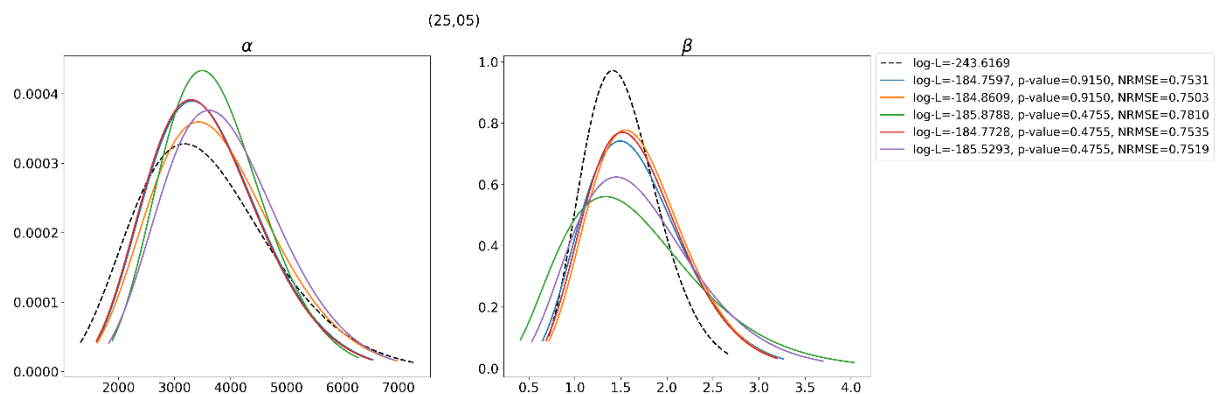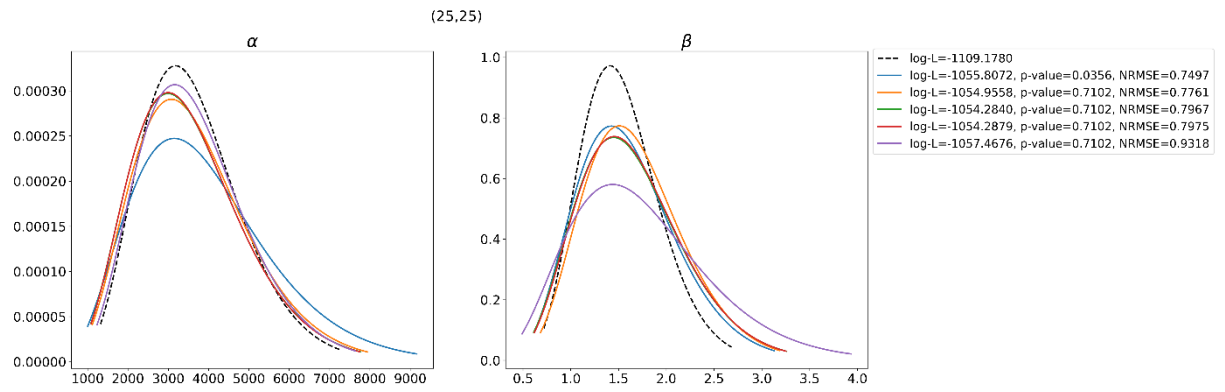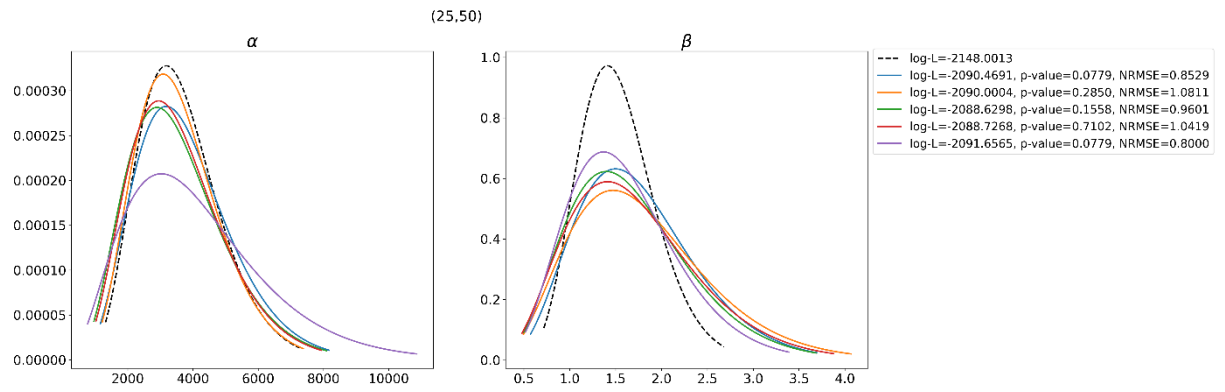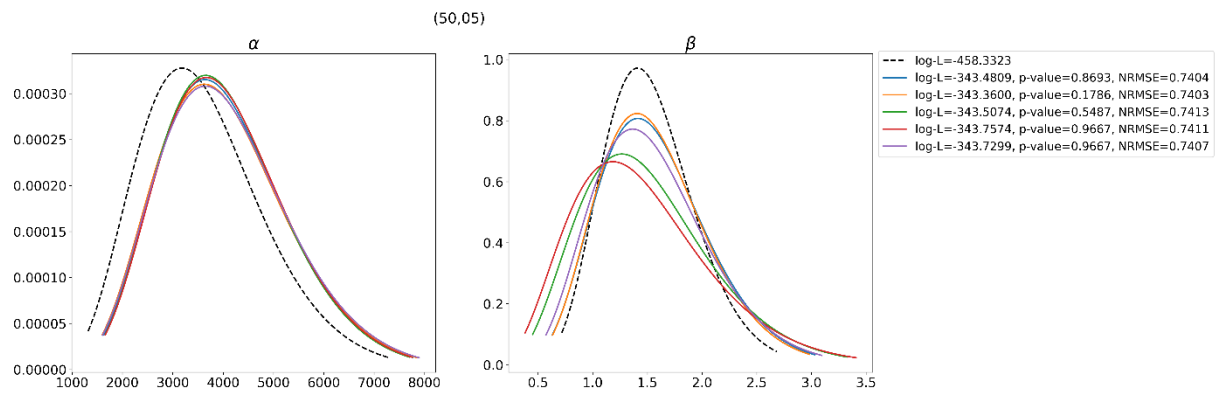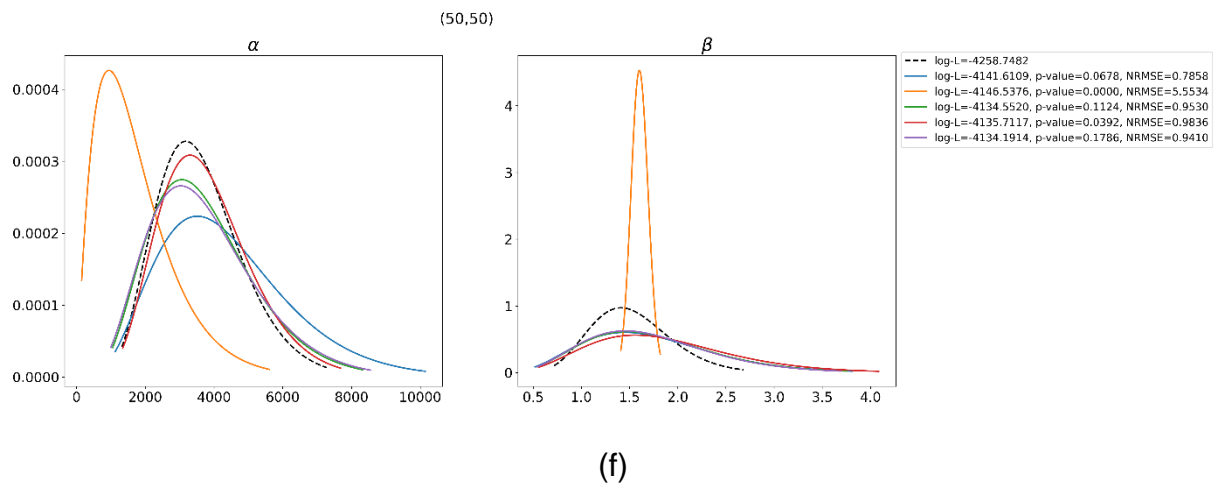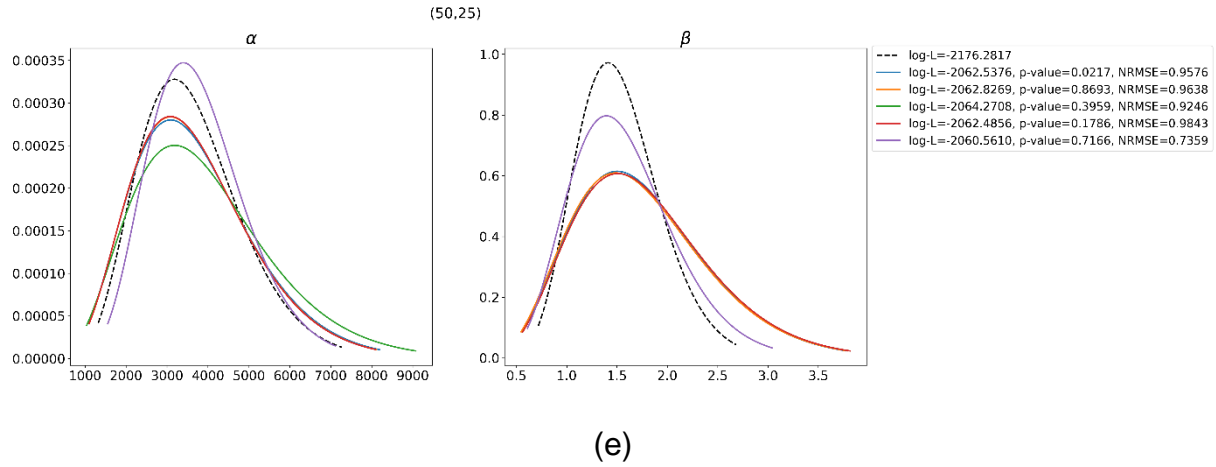
Source: The author (2025).

In general, there is variability in the results, both concerning the theoretical distribution and the results among themselves, although the latter is generally smaller. Despite this variability, all results exhibit close log-likelihood values. The complexity of this optimization problem, with multiple local minima, coupled with the fact that we determine the solution through the mean and standard deviation of the hyperparameters, means that even small variations in these values can lead to significant changes in the hyperparameters result.

Another notable observation pertains to the test size. It is evident that the number of subpopulations has a more pronounced impact on both the reduction of variability and proximity to the original distribution compared to the number of units per subpopulation. However, with more data available, results can be biased to a specific region of the parameter space and with less amplitude, as demonstrated by the orange curve in Figure 19f for the larger test set, i.e. (50,50). Remarkably, even with just 5 subpopulations, as exemplified by the (05,05) case, the results exhibited robust

performance concerning both p-values and NRMSE, as illustrated in Figure 17. These results validate the proposed methodology for the prior distribution estimation.

4.2   Posterior Distribution Estimation

This section presents the tests performed for the posterior distribution estimation model described in Section 3.2. The solution with the lowest NRMSE among all the runs obtained for each test presented before is selected to proceed with posterior distribution estimation. In this step, synthetic data is generated by simulating failure times based on a Weibull distribution with parameters $\alpha = 3566.61$ and $\beta = 1.34$ ("original parameters"). Using these parameters, five failure times are randomly generated for a hypothetical system $T = \{2562.3268, 1002.2196, 4355.1074, 1062.1094, 9688.4575\}$.

Here, the non-parametric posterior distribution estimations for α and β will be presented for both the largest and smallest prior tests, specifically (05,05) and (50,50), in Table 4 and Table 5, respectively. From the test (05,05) the best result was from Run 19, with NRMSE of 0.6278, and hyperparameters {3.9657,0.0007,11.3901,6.7363}. From the test (50,50) the best result was from Run 1, with NRMSE of 0.7858, and hyperparameters {5.0402,0.0011,6.0159,3.4480}. All posterior estimations conducted using MCMC algorithms showed convergence, as indicated by a split-$\hat{R}$ value of 1.

Table 4 – Non-parametric posterior distribution estimation from prior test (05,05).

	Original	Mean	SD	5%	25%	50%	75%	95%	$\hat{R}$
α	3566.61	8373	2829	4386	6362	8018	10020	13510	1
β	1.34	1.577	0.4866	0.8742	1.2255	1.5255	1.8740	2.4546	1

Source: The author (2025).

Table 5 – Non-parametric posterior distribution estimation from prior test (50,50).

	Original	Mean	SD	5%	25%	50%	75%	95%	$\hat{R}$
α	3566.61	6591	2234	3299	4997	6390	7955	10570	1
β	1.34	1.3796	0.6138	0.5712	0.9343	1.2784	1.7457	2.5282	1

Source: The author (2025).

These tables demonstrate that when starting from a more informative prior, estimated with a larger data set and thus characterized by lower uncertainty (narrower amplitude), the posterior distribution aligns more closely with the original parameters. This is evident when comparing the mean values of α and β in Table 5 to the values used to generate the data. Additionally, Figure 20 and Figure 21 – Posterior estimation results for (50,50) (a) α and (b) β. visually highlights that the posterior distributions shift away from the prior, becoming more skewed toward the observed data, especially for α. Besides the non-parametric distribution results of MCMC, goodness-of-fit was evaluated for three distributions: Gamma, Weibull, and Lognormal.

Figure 20 – Posterior estimation results for (05,05) (a) α and (b) β.

(a)

(b)
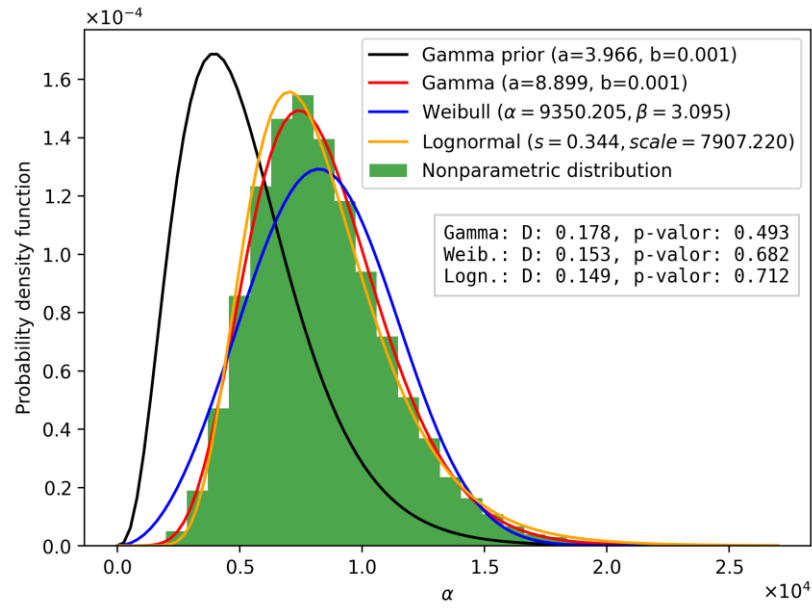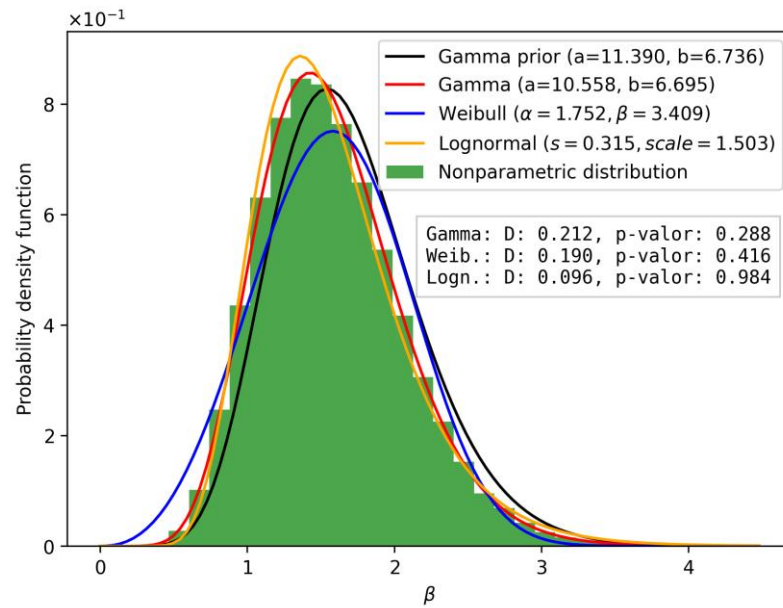
Source: The author (2025).

Figure 21 – Posterior estimation results for (50,50) (a) α and (b) β.

(a)

(b)
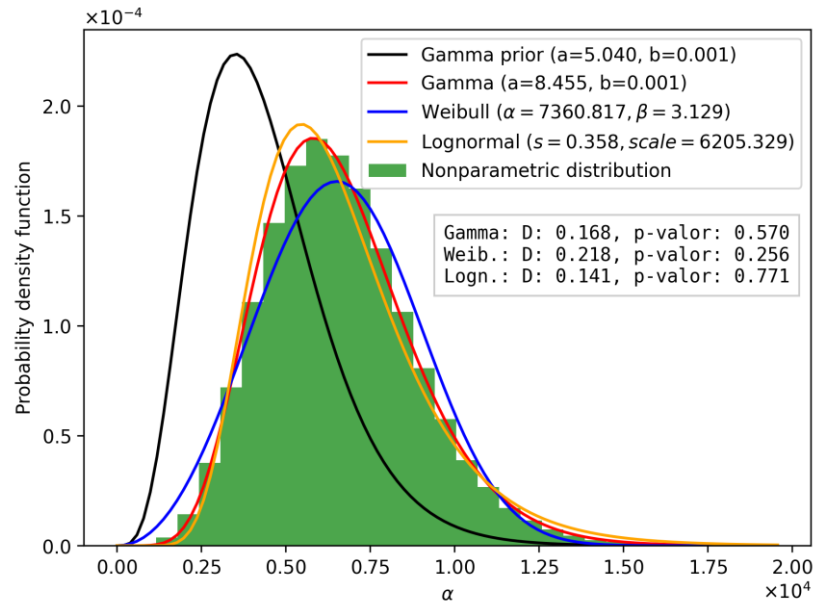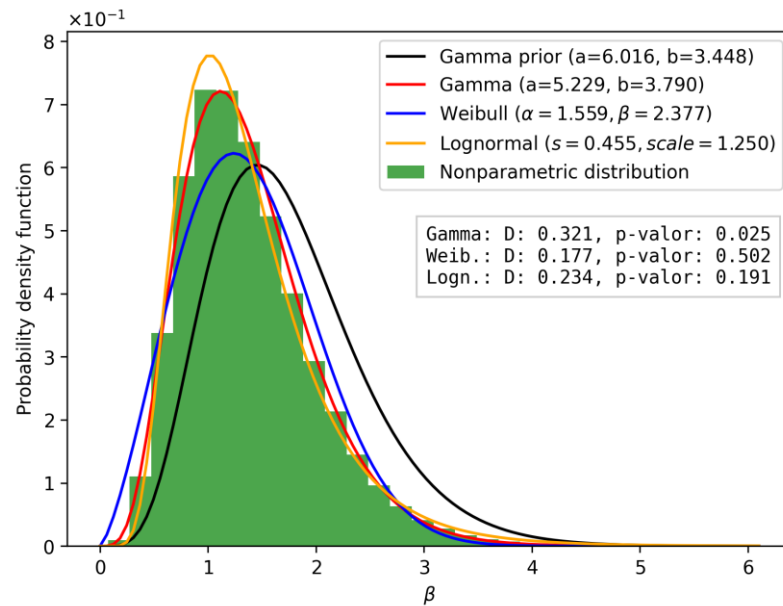
Source: The author (2025).

Although the Bayesian approach is primarily designed for scenarios with limited data, we also explored how the posterior distribution estimation behaves when a larger data set for the same equipment is available. Using the same parameters, a larger sample of 50 failure times was randomly generated, and the posterior estimation was performed. Here, it was considered the following specific data $T = [452.6755,$ 3212.2231, 10339.5549, 2732.3630, 1559.4472, 978.2900, 650.6369, 1288.1594, 3604.3077, 350.3445, 1565.8423, 457.3386, 1512.4945, 1221.3923, 450.5418,

5282.0135, 69.9732, 366.4941, 1188.8401, 1760.4854, 507.9755, 899.0795, 163.8543, 139.6921, 1017.6044, 873.1068, 2001.2863, 328.9064, 673.8020, 300.3467, 277.0293, 762.6984, 916.0977, 730.1398, 663.9674, 369.1867, 2250.9381, 489.8128, 645.2946, 950.0629, 1181.8031, 147.0101, 2641.1091, 489.4880, 559.1566, 53.6336, 1420.9667, 623.0107, 1119.6867, 1595.1587].

Table 6 – Non-parametric posterior distribution estimation from prior (50,50) with 50 failure times.

	Original	Mean	SD	5%	25%	50%	75%	95%	$\hat{R}$
α	3566.61	4711	1923	2130	3297	4424	5800	8285	1
β	1.34	1.8217	0.7213	0.8132	1.2988	1.7279	2.2446	3.1393	1

Source: The author (2025).

As anticipated, all results converged, reaffirming the impact of data quantity on estimation accuracy. Table 6 highlights this trend, showing that with more specific data, the mean estimates for α were significantly closer to the original parameters when using 50 failure times, compared to the results from only 5 failure times in Table 5.

4.3   Industrial Application Example

The previously presented results were obtained using simulated data, with a reference distribution available for comparison. Now, the methodology will be applied to a real case study pertaining to booster pumps in an O&G company. A booster pump is a type of pump used to increase the pressure of fluids within a pipeline. These pumps are typically deployed in situations where the natural pressure in the pipeline is insufficient to maintain an optimal flow rate.

The data in Table 7 was sourced from OREDA, and pertains to booster pumps across five different installations, encompassing a total of 25 systems. Each row corresponds to one system, with the first column indicating which subpopulation the system belongs to, and K indicates the number of failures during time horizon T. This data will be used to estimate the prior distribution. Each installation will be treated as a subpopulation, as the systems within the same installation are subject to similar operating conditions, which may differ across installations. The data includes both critical failures that result in shutdowns and degraded failures, where the equipment operates below nominal levels but does not cause a complete operational stoppage.

Table 7 – Generic data from booster pumps, sourced from OREDA.

Subpopulation	K	T
1	1	22201
1	4	8345
1	4	21103
2	6	24437
2	6	10226
2	8	8387
2	5	4906
2	5	11208
2	7	12734
3	1	19870
3	1	15131
3	1	18377
3	0	17082
3	1	20542
3	3	16512
3	6	20187
3	2	12836
4	4	19585
4	1	22292
4	1	15674
4	4	12855
5	4	10485
5	3	13838
5	3	19301
5	1	22179

Source: The author (2025).

Figure 22 presents the results of three prior distribution estimations for this industrial application example. While some variability is observed among the results, the modes are very close, and all NRMSE values are similarly low. Based on the NRMSE, the first run is selected for use in the next step.

Figure 22 – Prior distribution estimation for the industrial application example.
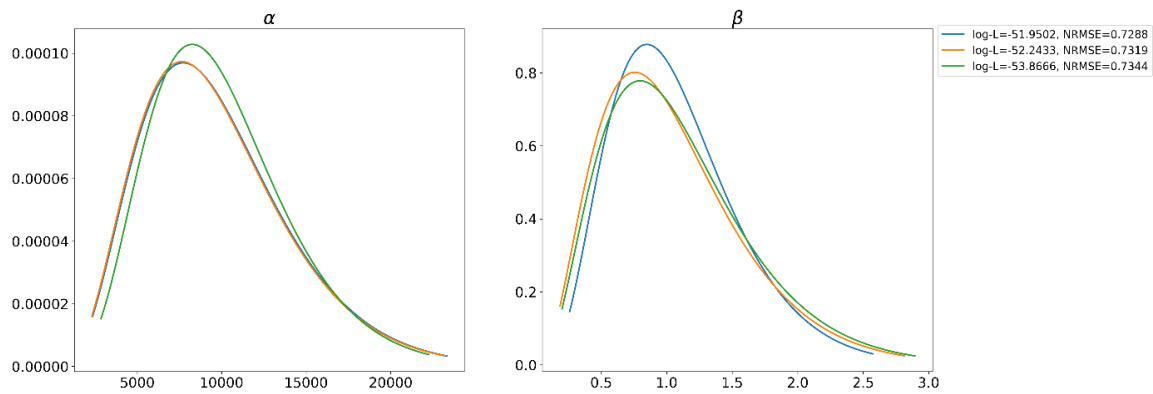
Source: The author (2025).

With the result presented above, the next step is the posterior distribution estimation. The specific failure data used in this stage comes from a booster pump operating in an Oil & Gas company, where reliability estimates are needed. The dataset consists of the following failure times: T =[1728, 1781, 3533, 4901, 5237, 5789, 6341, 7109, 7157, 7205, 8501, 9029, 9821, 12557, 16013, 17453, 23357, 26021, 30365, 31781, 35981], which includes both critical and degraded failures.

As illustrated in Figure 23, the time intervals between failures tend to increase, suggesting an improvement in the system's condition over time. This behavior indicates that the Weibull shape parameter β is likely to be less than or close to 1. Although such values are already considered within the prior distribution, we now expect the posterior to be more concentrated in this region.

Figure 23 – Specific data from a booster pump in an O&G company.
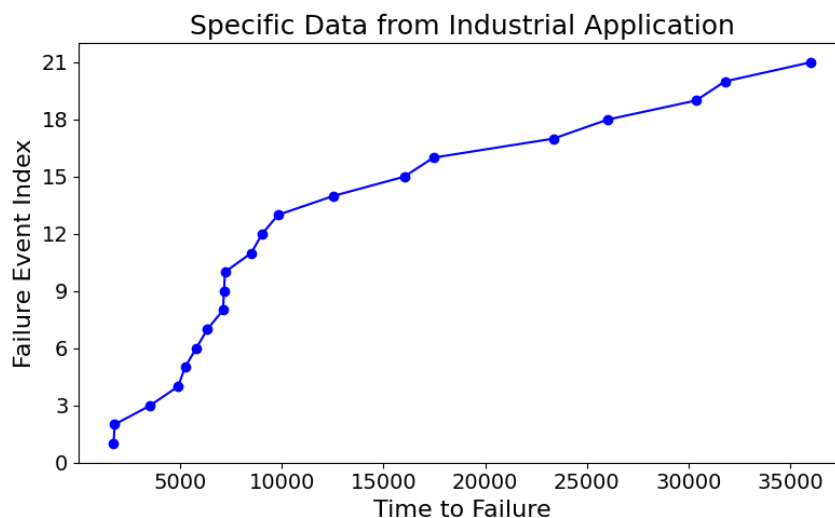
Source: The author (2025).

Figure 24 shows the result of the posterior distribution estimation, showing a shift in the mode compared to the prior distribution, for both α and β. The posterior distribution of β is now more skewed toward values below 1 and has reduced variance. According to the p-value, the posterior of α can be modeled by a Lognormal distribution and β by a Gamma distribution.

Figure 24 – Posterior estimation results for the industrial application example for (a) α and (b) β.

(a)

(b)
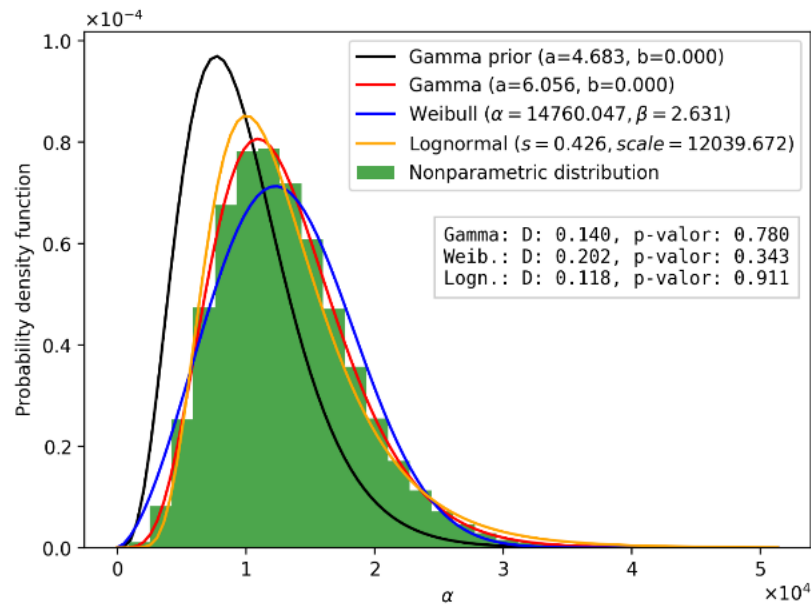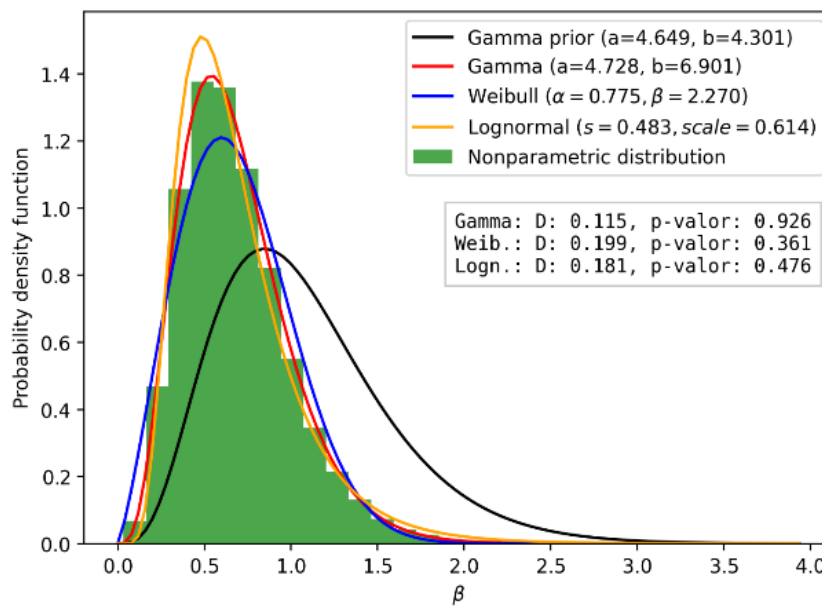
Source: The author (2025).

These results reflect the use of the Weibull model. If the same dataset were analyzed under the assumption of a constant failure rate (i.e., using the Exponential model described in Section 2.2), the estimated mean of λ would be 0.00051325. As shown in Figure 25, the failure intensity function of the Weibull model is consistently lower than that of the Exponential model. This implies that the Weibull-based estimates are more optimistic and may lead to fewer maintenance actions. In contrast, assuming a constant failure rate could result in unnecessary maintenance activities and higher operational costs.

Figure 25 – Comparison between the Failure Intensity Function for the Weibull and the Exponential models.
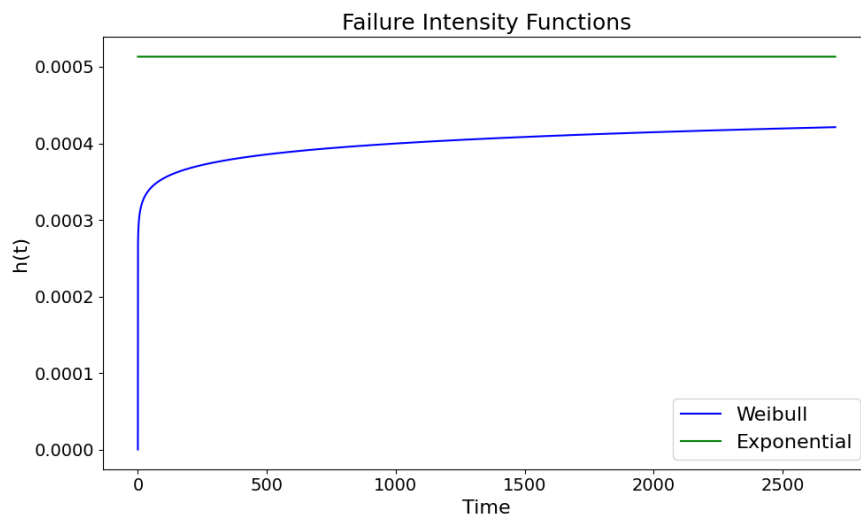
Source: The author (2025).

Mahmoudi (2021) underscored the critical need for collaboration between academia and industry to drive innovation and practical applications in the O&G sector. Building on this foundation, the current study expands traditional methodologies by incorporating a non-constant failure intensity into the analysis, thereby enhancing the applicability of reliability models to real-world scenarios. This approach not only addresses the limitations of assuming a constant failure intensity function but also offers more accurate insights when applied to complex systems in industry, bridging the gap between theoretical research and practical implementation.

5   SOFTWARE DEVELOPMENT

Petrobayes is a user-friendly, web-based software designed to support decision-makers in performing Bayesian analysis and reliability estimation with ease. It is registered in the *Instituto Nacional da Propriedade Industrial* (INPI) under the number BR512024000005-4. The platform is organized into different modules tailored to specific user needs and problem domains. Currently, the only language supported is Portuguese. Petrobayes has been presented in the article (Santana et al., 2022).

The software was built using Streamlit (https://streamlit.io/), an open-source Python framework for building web applications. Streamlit was specifically designed to address the growing demand among data scientists for creating user-friendly interfaces quickly and efficiently. Petrobayes leverages this intuitive framework to provide an easy-to-use interface for the Bayesian model presented in Section 3 and extend its analysis. An overview of the software's modules and a step-by-step usage guide are provided next.

5.1   Features and Functionalities

While Petrobayes primarily focuses on Bayesian inference, it also includes other supporting modules. In total, Petrobayes comprises three modules: Statistical, Bayesian, and Availability. Each module functions independently but becomes particularly powerful when used in combination.

The Statistical module conducts fit and homogeneity tests on failure data. Fit tests help identify the parametric probability distribution that best represents the data, while homogeneity tests determine whether different data sets originate from the same population (i.e., are generated by the same statistical phenomenon). These capabilities are especially valuable for the Bayesian Modules, as they allow for testing hypotheses on the homogeneity of diverse data sets and grouping them into different subpopulations.

The Bayesian module is dedicated to Bayesian analysis and is divided into two sub-modules: Bayesian Exponential and Bayesian Weibull. These modules differ in their assumptions regarding the distribution of failure times. The Bayesian Exponential module assumes that failure times follow an Exponential distribution with a constant failure intensity function (represented by λ), implementing the model described in Section 2.2, while the Bayesian Weibull module is based on a Weibull distribution,

characterized by scale (α) and shape (β) parameters, implementing the model described in Section 3. Each module's analysis unfolds in two main stages:

- Stage 1: Estimates the prior distribution, reflecting the PVD based on generic data, like OREDA.
- Stage 2: Estimates the posterior distribution by updating the prior information with specific data from the system of interest, known as the likelihood distribution.

The Availability module allows for analyzing the behavior of a system over time. This modeling is tailored to the system's characteristics and the level of detail provided by the user, offering the flexibility to choose from three distinct modeling approaches: 2 states, 3 states and 4 states. The model employs discrete states of a Semi-Markov Process to describe the operational state of the system, ranging from perfect functioning to critical failure. This module has been presented in the article (Siqueira et al., 2023). The result from the Bayesian module can be used here to estimate the availability of the system under analysis.

Figure 26 illustrates how these modules can be integrated, demonstrating how the output of one module can serve as input for another. While this overview highlights the software's overall potential, the focus here will be on the Bayesian Weibull module, which implements the methodology outlined in Section 3.

Figure 26 – Petrobayes' modules connection.
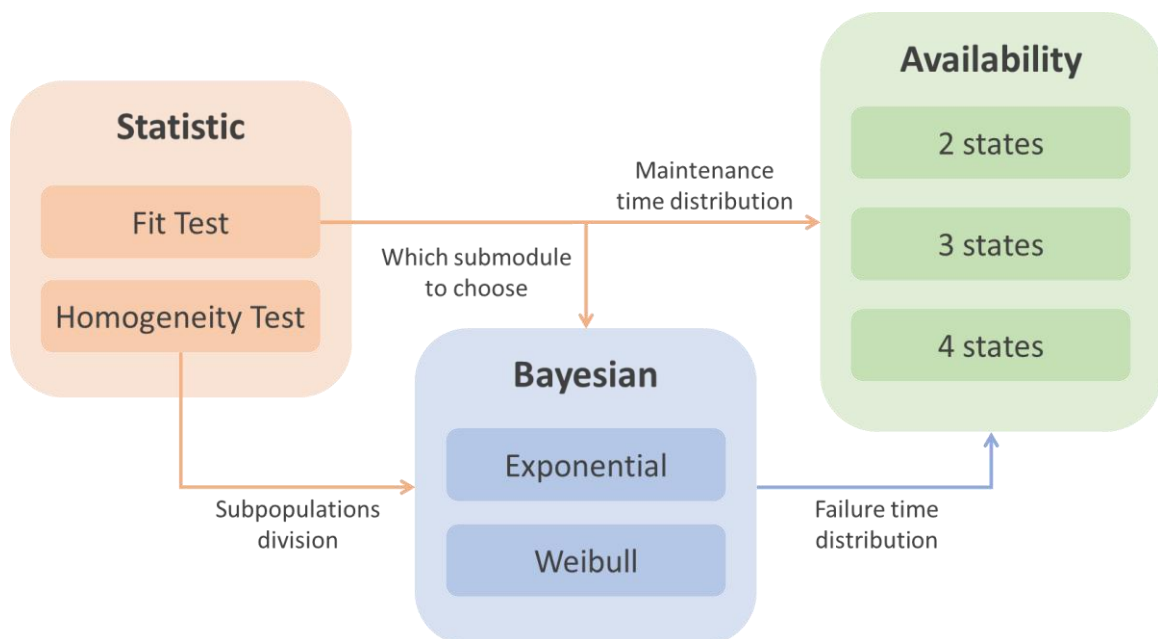
Source: The author (2025).

5.2   Application and Usability

This section demonstrates the step-by-step use of the Bayesian Weibull module, which implements the methodology outlined in Section 3. It is important to note that η in the software is equivalent to α in the developed model. As previously mentioned, the software only supports Portuguese. However, in here, visual cues and explanations will be provided to guide the reader through the content. Figure 27 shows the home screen of this module, where you can find a brief explanation in the "Guia rápido" tab and see the required file format in the "Exemplo de formatação dos dados" tab.

Figure 27 – Home screen of the Bayesian Weibull Module.

Source: The author (2025).

To begin the analysis, data for Stage 1 must be uploaded via a '.xlsx' file, where each row corresponds to one system, with the first column indicating which subpopulation the system belongs to, and K indicates the number of failures during time horizon T, as shown in Figure 28. The interface accepts data with or without headers, as pointed by the red arrow in Figure 28. The header corresponds to the first row of the '.xlsx' file. Selecting "Sim" indicates that the first row should be used as the column title, while selecting "Não" means that the first row is treated as data.

Figure 28 – Data Input for Stage 1 of the Bayesian Weibull Module.

O arquivo contém cabeçalho?
- ● Sim
- ○ Não

Selecione um arquivo XLSX

Drag and drop file here
Limit 200MB per file • XLSX

Browse files

dados_weibull.xlsx 10.2KB ✕

	Subpopulação	K	T
0	1	0	2000
1	1	0	4000
2	1	0	6000
3	1	0	8000
4	1	0	10000
5	2	0	2000
6	2	0	4000
7	2	1	6000
8	2	1	8000
9	2	1	10000

O arquivo importado contém dados de 25 equipamentos contidos em 5 subpopulações.

Configurações avançadas ⌄

Executar Estágio 1

Source: The author (2025).

Table 8 shows the data that will be used as an application example. It's important to note that the model is compatible with various data configurations, all of which must consist of positive integers, except for the "T" parameter, which should be a positive real number. Additionally, it's possible to include information on equipment that did not fail. For example, the first piece of equipment in Table 8 did not experience any failures during the 2000-unit observation period, so its "K" value is 0.

Table 8 – Application example for the Stage 1 of the Bayesian Weibull Module.

Subpopulation	K	T
1	0	2000
1	0	4000
1	0	6000
1	0	8000
1	0	10000
2	0	2000
2	0	4000
2	1	6000
2	1	8000
2	1	10000
3	0	2000
3	0	4000
3	0	6000
3	0	8000
3	1	10000
4	0	2000
4	0	4000
4	0	6000
4	0	8000
4	0	10000
5	0	2000
5	0	4000
5	0	6000
5	0	8000
5	1	10000

Source: The author (2025).

The interface now includes additional input options under the "Configurações avançadas" tab (Figure 29), related to the PSO, such as the number of rounds to execute, the maximum number of iterations per round, and patience (the maximum number of iterations where consecutive results do not change after which the search should stop). The default values are suggestions based on the extensive testing provided in Section 4, but they can be adjusted according to the user needs. Once these settings are configured, clicking "Executar Estágio 1" in Figure 28 will initiate the prior distribution estimation according to Section 3.1.

Figure 29 – Advanced Settings for the Stage 1 of the Bayesian Weibull Module.

Source: The author (2025).

It's important to note that even when using the same input data and following the same steps, the randomness inherent to the method can lead to different solutions in different tests. However, in theory, the solutions tend to converge within a specific region of the search space. Since multiple rounds of PSO are conducted, different solutions are generated, and their graphs are displayed.

Additionally, the solutions are evaluated based on the log-likelihood and the NRMSE (see Section 3.1). The values of these measures are shown alongside each solution. The results are presented in terms of the prior distributions of η e β (Figure 30), assuming each follows a Gamma distribution with its respective hyperparameters. The graphs initially display only one solution: the best among all rounds (i.e., the one with the lowest NRMSE). However, these graphs are dynamic, allowing the user to view all solution curves by clicking on their names on the right side of the screen, as indicated by the arrow.

Figure 30 – Stage 1 Result: Prior Distribution of η e β.

Source: The author (2025).

All graphs are interactive, allowing users to select and hide curves as desired. The graph's axis scale can also be adjusted according to the following options:

- Linear: Uses exact values on the axes.
- Log-x: Converts the X-axis to a logarithmic scale.
- Log-y: Converts the Y-axis to a logarithmic scale.
- Log-xy: Converts both the X and Y axes to logarithmic scales.

It's important to emphasize that the solution with the lowest NRMSE will be passed on to Stage 2 if the option to use Stage 1 results is selected, as will be shown later. In addition to the information related to the prior distribution, the mean values of of η e β along with other results are presented, such as: the probability density function f(t) (Figure 31), the failure intensity over time h(t) (Figure 32), and the reliability R(t) with the confidence interval (Figure 33). All graphs show the Mean Time To Failure

(MTTF) with a vertical dashed line, presented in the same time unit as the data provided.

Figure 31 – Stage 1 Result: Probability Density Function f(t).

Source: The author (2025).

Figure 32 – Stage 1 Result: Failure intensity h(t).
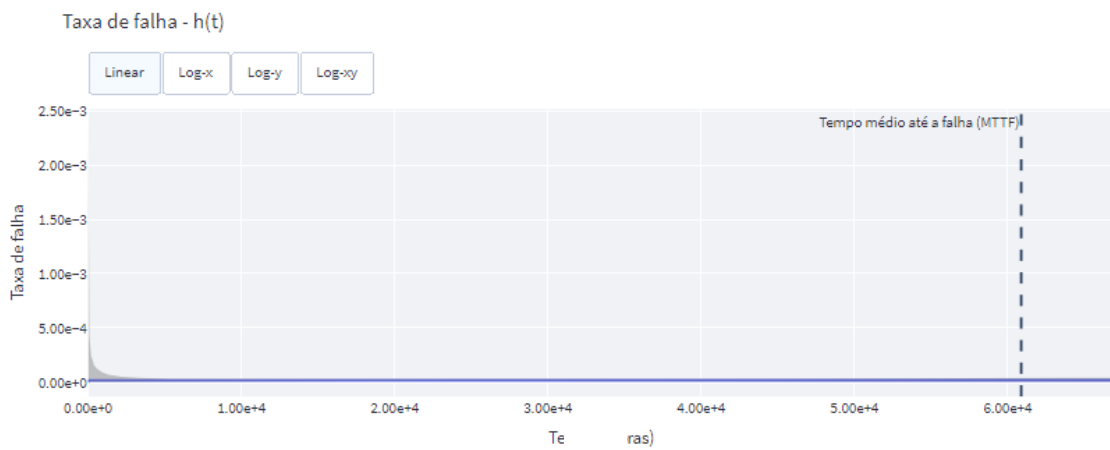
Source: The author (2025).

Figure 33 – Stage 1 Result: Reliability R(t).
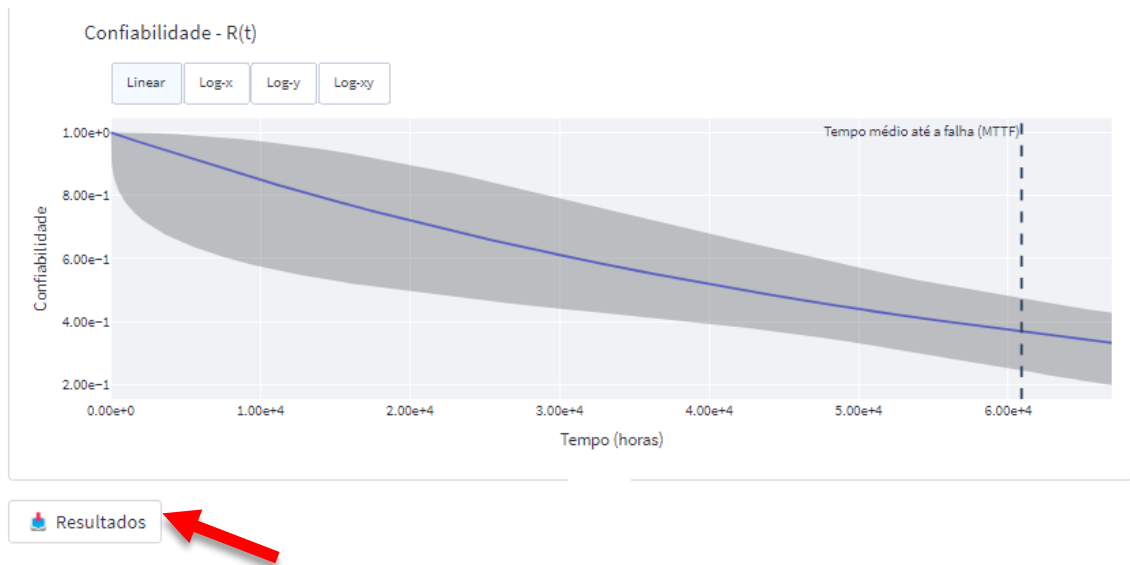
Source: The author (2025).

At the end of the execution, the prior distribution results can be saved, as indicated by the red arrow in Figure 33. This allows the analysis to be resumed later without needing to re-run Stage 1. Next, if not using the Stage 1 result, it is necessary to manually enter the hyperparameters of the Gamma distribution of η and β (Figure 34). If the Stage 1 result is used, as shown in Figure 35, there is no need to provide hyperparameter values.

Figure 34 – Prior distribution Input for Stage 2, when not using the result of Stage 1.

Estágio 2: Dados específicos para a verossimilhança

Utilizar o resultado do Estágio 1?

○ Sim
● Não

Hiperparâmetros da distribuição *a priori* Gamma de η

a	⑦	b	⑦
1,0000	− +	1,0000	− +

Hiperparâmetros da distribuição *a priori* Gamma de β

a	⑦	b	⑦
1,0000	− +	1,0000	− +

Exemplo de formatação dos dados ⌄

Escolha o arquivo XLSX

Drag and drop file here
Limit 200MB per file • XLSX   Browse files

☑ Gerar relatório de resultados

Executar Estágio 2

Source: The author (2025).

Additionally, the specific data, given as failure times or censored times, must be entered via an '.xlsx' file. Figure 35 provides a preview of the specific data, as demonstrated in this example. A PDF report summarizing the main results can be generated at the end by selecting "Gerar relatório de resultados" as indicated by the red arrow in Figure 35.

Figure 35 – Data Input for Stage 2.

Source: The author (2025).

The user should then click "Executar Estágio 2" to initiate the posterior distribution estimation following the approach of Section 3.2. The example data used here is shown in Table 9, where the "Type" column indicates whether that time corresponds to a failure (F) or a censure (C).

Table 9 –Data for Stage 2.

Time	Type
1500	F
2000	F
5200	F
10000	F
15200	C
21300	C

Source: The author (2025).

Executing Stage 2 deploys the MCMC model, which yields non-parametric distributions. To make these results applicable to other analyses, we fit specific parametric distributions, namely Gamma, Weibull, and Lognormal. Following the completion of Stage 2 of the Bayesian Weibull module, a table is presented containing information on the tested parametric distributions, as shown in Figure 36.

Figure 36 – Stage 2 Result: parametric distributions.

Distribuições *a posteriori* ajustadas

Ajuste dos parâmetros

Distribuição *a posteriori* de η

	Gamma	Weibull	Lognormal
Teste K-S (p-valor)	0.589	0.32734	0.979529
a	31.7761	61,880.2	0.1788
b	0.0006	5.8348	56,663
Média	57,566.4	57,318.4	57,575.7
Variância	1.04289e+08	1.2987e+08	1.07654e+08
Desvio-padrão	10,212.2	11,396	10,375.7
Mediana	56,963.7	58,112.7	56,663
LIC 5	41,859.9	37,194.2	42,227.4
LSC 95	75,329.2	74,682.2	76,033.5

Distribuição *a posteriori* de β

	Gamma	Weibull	Lognormal
Teste K-S (p-valor)	0.9774	0.623152	0.322194
a	7.8955	0.6159	0.3677
b	14.3565	2.9706	0.5155
Média	0.55	0.5497	0.5516
Variância	0.0383	0.0406	0.044
Desvio-padrão	0.1957	0.2016	0.2098
Mediana	0.5269	0.5444	0.5155
LIC 5	0.2722	0.2266	0.2816
LSC 95	0.9064	0.8911	0.9438

Source: The author (2025).

The posterior distribution for η and β are shown in Figure 37 and Figure 38, respectively. These graphs are interactive, so it is possible to see all the distributions tested, but the best fit (based on the p-value of the Kolmogorov-Smirnov test) is highlighted. In this example, the best fit for η was a Lognormal distribution and for β a Gamma distribution.

Figure 37 – Stage 2 Result: η posterior distribution.
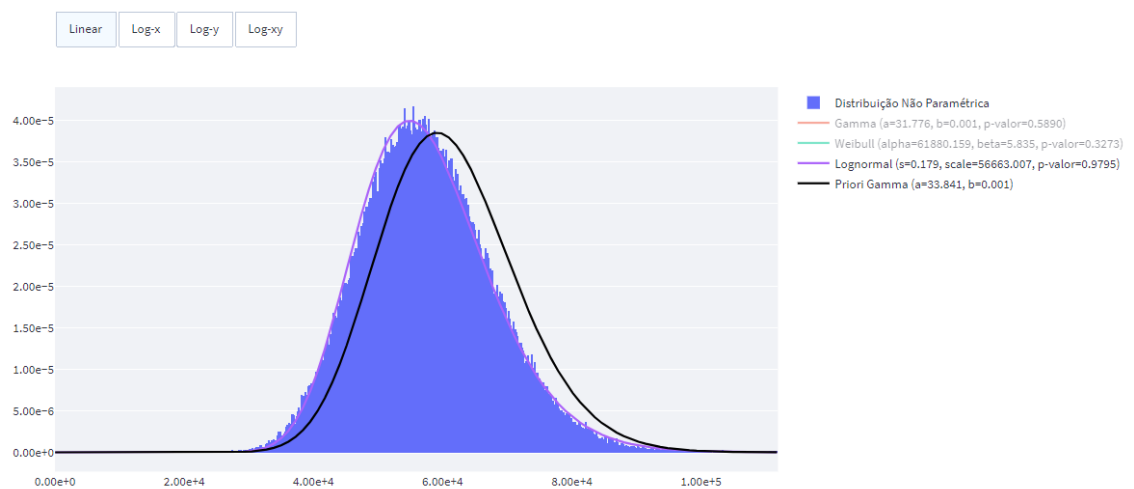
Source: The author (2025).

Figure 38 – Stage 2 Result: β posterior distribution.
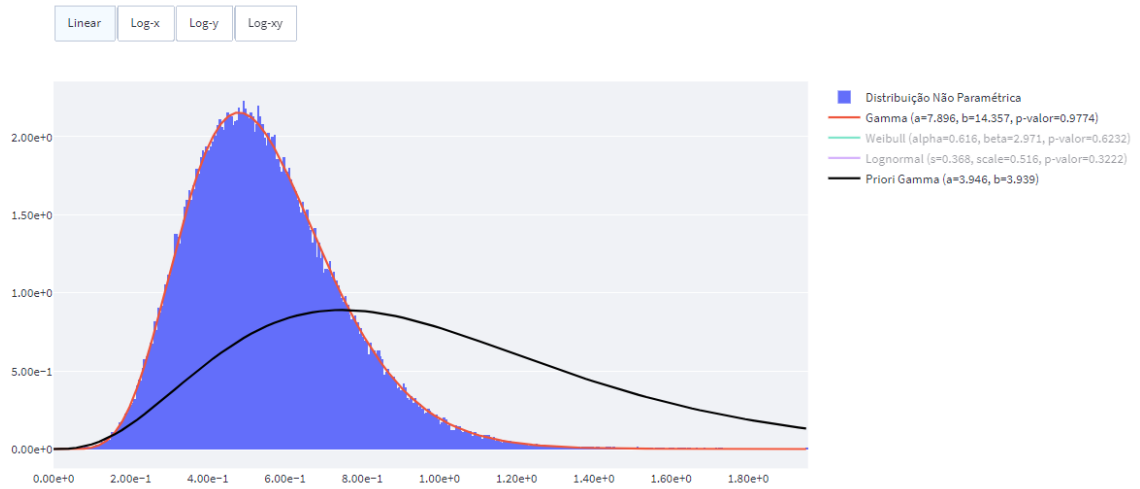
In addition to the information related to the posterior distribution, the mean values of η e β along with other results are presented, such as: the probability density function f(t) (Figure 39), the failure intensity over time h(t) (Figure 40), and the reliability R(t) with the confidence interval (Figure 41). All graphs indicate the MTTF with a vertical dashed line, presented in the same time unit as the data provided.

Figure 39 – Stage 2 Result: Probability Density Function f(t).

Figure 40 – Stage 2 Result: Failure intensity h(t).

Source: The author (2025).

Figure 41 – Stage 2 Result: Reliability R(t).
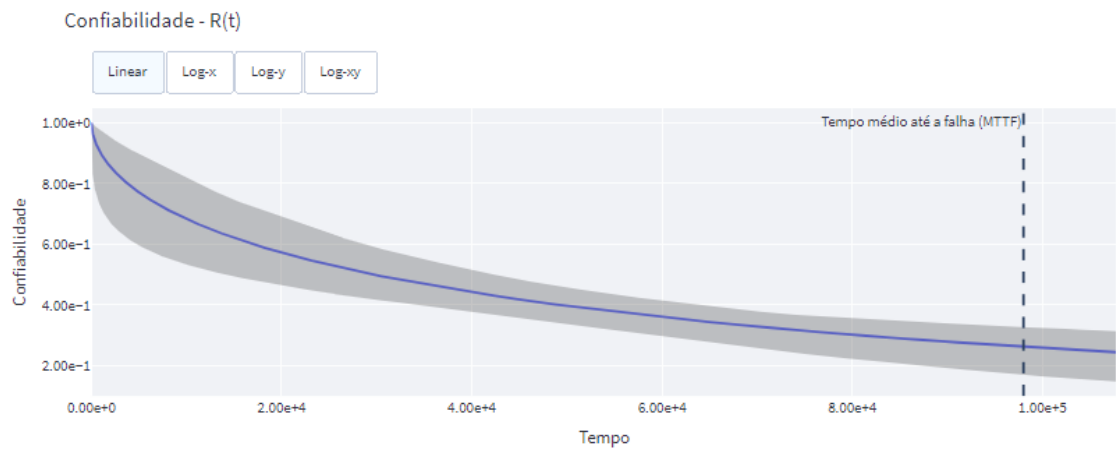
Source: The author (2025).

Figure 42 shows the joint posterior distribution graph of η e β generated based on the best parametric fit. At the conclusion of this stage, the user can click on the option to download a report to obtain a summary of the key metrics and relevant information from the analysis, as indicated by the red arrow in Figure 42.

Figure 42 – Stage 2 Result: Joint posterior distribution of η and β.
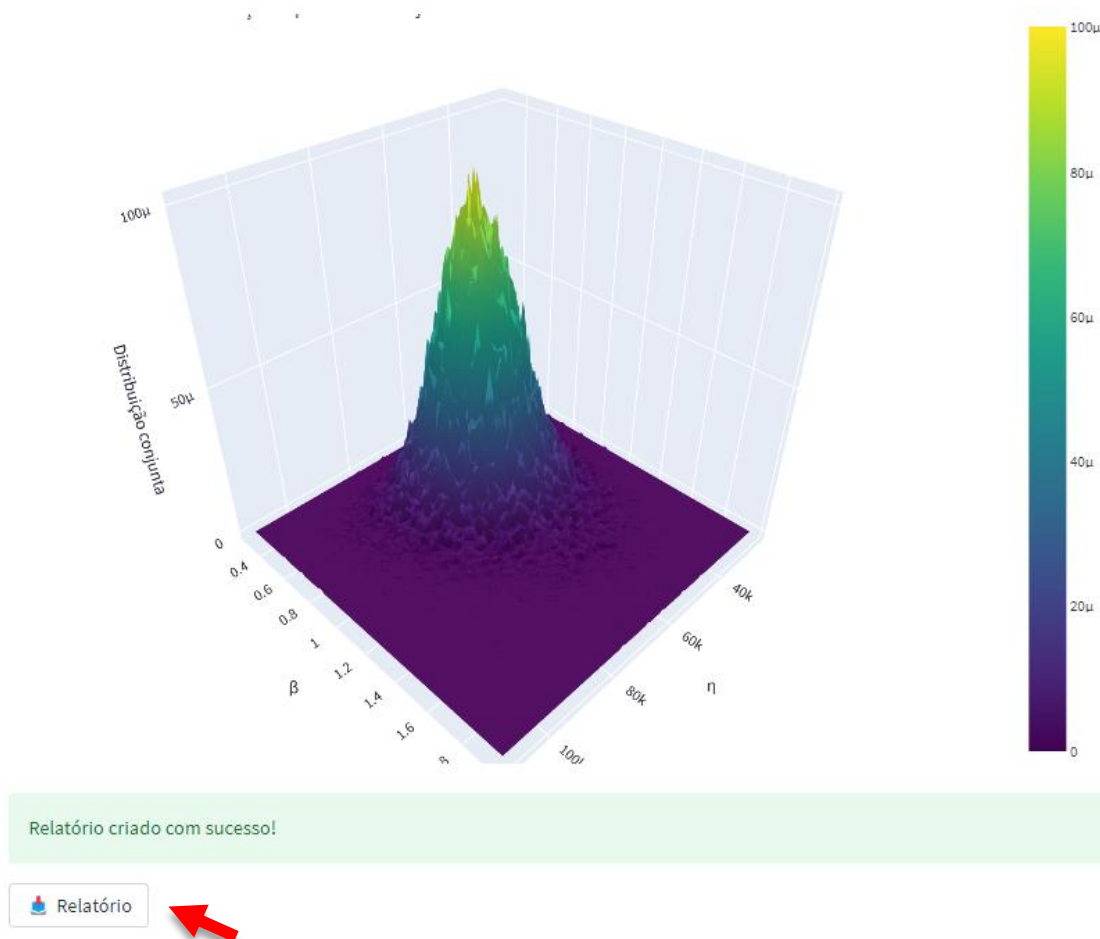
Source: The author (2025).

Figure 43 provides an example of the first page of the report generated at the end of the model's execution. This report allows the user to save the results obtained for future reference, instead of re-running the model when needed.

Figure 43 – Example from the first page of a report from the Bayesian Weibull module.

Petrobayes
Relatório - Weibull Bayesiano
Data: 25-10-2023 Hora: 16:07:13

Descrição da distribuição a priori de alpha

	a	b	Média	Desvio-padrão	Mediana	LIC 5	LSC 95
Rodada 1	36.3923	0.00	6.02e+04	9.97e+03	5.96e+04	4.48e+04	7.75e+04
Rodada 2	33.8411	0.00	6.09e+04	1.05e+04	6.03e+04	4.48e+04	7.91e+04

Visualização da distribuição a priori de alpha

Descrição da distribuição a posteriori de alpha

	Gamma	Weibull	Lognormal
Teste K-S (p-valor)	0.589	0.3273	0.9795
a	31.7761	61880.1588	0.1788
b	0.0006	5.8348	56663.007
Média	57566.4078	57318.3613	57575.7299
Variância	104288902.8163	129869650.5909	107654463.5763
Desvio-padrão	10212.1938	11396.0366	10375.6669

Source: The author (2025).

6   CONCLUSIONS

This work presents a two-stage Bayesian approach to address the challenges of reliability estimation for equipment in the O&G industry, under the hypothesis of non-constant failure intensity and minimal repair. The approach brings two key contributions: the incorporation of prior knowledge using failure count data within a Weibull model and the development of an accessible web application, Petrobayes, to make these models readily available to the industry. Indeed, these contributions are particularly significant as they address limitations with the use of OREDA database, especially concerning the challenging of non-constant failure intensity and minimal repairs, developing tailored methods to address the specific nuances of the dataset and the assumptions. Therefore, this work overcomes limitations in risk and reliability analysis in the O&G sector and supports effective risk management.

The study was structured around the two-stage Bayesian framework, divided into prior and posterior estimation phases. In the prior estimation phase, PSO was employed to solve the MLE problem and determine the optimal set of hyperparameters that model the PVD. The results were validated by three performance metrics: log-likelihood, p-value, and NRMSE. The use of p-value alone as a decision metric is challenging, as some solutions exhibited equivalent p-values, complicating the decision-making process. Therefore, the NRMSE is the primary metric utilized to identify the optimal solution.

The prior estimation is subjected to the following sources of variability: the calculation of the log-likelihood using Importance Sampling, the PSO heuristic and the fact that different combinations of hyperparameter values often yield very similar log-likelihood results. The last point was corroborated by a low standard deviation on the log-likelihood obtained in different runs, even though the hyperparameters themselves varied. This can be expected of an algorithm exploring solutions in a continuous space. However, the results of the study showed that despite some variability, the log-likelihood values consistently outperformed the theoretical distribution across different test sets. Even with small data sets, such as the (05,05) case, the model delivered strong performance in terms of p-value and NRMSE.

In this context, the Petrobayes software was designed as a user-friendly tool for Bayesian inference. It is organized into different modules: Statistical, Bayesian, and Availability. Each module functions independently but becomes particularly powerful

when used in combination. The focus here was on the Bayesian module that implements the model discussed here. Petrobayes extends the results presented here by using the posterior distribution to estimate the reliability and failure intensity.

So, this study built on prior research by introducing a non-constant failure intensity and minimal repairs, addressing limitations in traditional reliability models and offering more accurate insights when applied to real-world scenarios in the O&G industry. In scenarios where specific data is sparse or unavailable, the methodology demonstrated that it could effectively use more informative priors derived from generic data to generate posterior distributions closely aligned with real-world data. This finding is particularly important for industries such as O&G, where equipment is often expected to operate in extreme conditions for extended periods, and acquiring large amounts of specific data is time-consuming and expensive. The research thus offers a practical solution for improving reliability estimates under such conditions, allowing for better maintenance planning, cost estimation, and system availability predictions for the Oil & Gas sector.

Despite its contributions, this work has some limitations, particularly regarding computational efficiency when handling large datasets. Variability in prior estimation also persists, indicating the need to explore alternative heuristic methods, such as genetic algorithms, to improve the MLE process. Future research should focus on enhancing computational speed, reducing prior estimation variability, and applying the proposed framework to other complex systems beyond the oil and gas industry.

Future research could benefit from relaxing the assumption of minimal repair and adopting a Generalized Renewal Process (GRP) model (Wang & Yang, 2012) could further improve reliability analysis by better representing real-world maintenance practices. While the Gamma prior distribution was selected due to its widespread use in the literature, thanks to its support for positive real values and flexibility in modeling skewed behaviors, there is also potential to explore alternative parametric priors. These alternatives could be considered with appropriate adjustments.

Another important direction would be to develop an English version of the PetroBayes software, making it accessible to the global scientific community and fostering broader collaboration and application in diverse industrial contexts. This work

thus lays the foundation for more accurate and comprehensive reliability assessments, supporting better decision-making and operational efficiency in high-risk industries.

BIBLIOGRAPHY

Animah, I., & Shafiee, M. (2020). Application of risk analysis in the liquefied natural gas (LNG) sector: An overview. *Journal of Loss Prevention in the Process Industries*, *63*, 103980. https://doi.org/10.1016/j.jlp.2019.103980

BahooToroody, A., Abaei, M. M., Arzaghi, E., Song, G., De Carlo, F., Paltrinieri, N., & Abbassi, R. (2020). On reliability challenges of repairable systems using hierarchical bayesian inference and maximum likelihood estimation. *Process Safety and Environmental Protection*, *135*, 157–165. https://doi.org/10.1016/j.psep.2019.11.039

Barbu, A., & Zhu, S.-C. (2020). *Monte Carlo Methods*. Springer Singapore. https://doi.org/10.1007/978-981-13-2971-5

Basu, A. P., & Rigdon, S. E. (2001). Ch. 2. The weibull nonhomogeneous poisson process. In *Handbook of Statistics* (20th ed., pp. 43–68). Elsevier. https://doi.org/10.1016/S0169-7161(01)20004-2

Betancourt, M. (2017). *A Conceptual Introduction to Hamiltonian Monte Carlo*.

Bhattacharyya, B. (2021). Structural reliability analysis by a Bayesian sparse polynomial chaos expansion. *Structural Safety*, *90*. https://doi.org/10.1016/j.strusafe.2020.102074

Bolstad, W. M. (2007). Introduction to Bayesian Statistics. In *Applied Statistics* (2nd ed., Issue 1). JohnWiley & Sons.

Bolstad, W. M. (2009). *Understanding Computational Bayesian Statistics* (Vol. 4, Issue 1). John Wiley & Sons. https://doi.org/10.1002/9780470567371

Bratton, D., & Kennedy, J. (2007). Defining a Standard for Particle Swarm Optimization. *2007 IEEE Swarm Intelligence Symposium*, *Sis*, 120–127. https://doi.org/10.1109/SIS.2007.368035

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Casella, G., & Berger, R. L. (2002). *Statistical Inference* (2nd ed.). Duxbury.

Chen, M.-H., Shao, Q.-M., & Ibrahim, J. G. (2000). *Monte Carlo Methods in Bayesian Computation*. Springer New York. https://doi.org/10.1007/978-1-4612-1276-8

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric Statistics: A Step-by-Step Approach* (2nd ed.). Wiley.

Craiu, R. V., & Rosenthal, J. S. (2014). Bayesian Computation Via Markov Chain Monte Carlo. *Annual Review of Statistics and Its Application*, *1*(1), 179–201. https://doi.org/10.1146/annurev-statistics-022513-115540

das Chagas Moura, M., Azevedo, R. V., Droguett, E. L., Chaves, L. R., Lins, I. D., Vilela, R. F., & Filho, R. S. (2016). Estimation of expected number of accidents and workforce unavailability through Bayesian population variability analysis and Markov-based model. *Reliability Engineering & System Safety*, *150*, 136–146. https://doi.org/10.1016/j.ress.2016.01.017

das Chagas Moura, M., Droguett, E. L., Firmino, P. R. A., & Ferreira, R. J. (2014). A competing risk model for dependent and imperfect condition–based preventive and corrective maintenances. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *228*(6), 590–605. https://doi.org/10.1177/1748006X14540878

Droguett, E. L., Groen, F., & Mosleh, A. (2004). The combined use of data and expert estimates in population variability analysis. *Reliability Engineering & System Safety*, *83*(3), 311–321. https://doi.org/10.1016/j.ress.2003.10.007

Erto, P., & Giorgio, M. (2013). *A note on using Bayes priors for Weibull distribution*. 31. http://arxiv.org/abs/1310.7056

Frohner, F. H. (1985). Analytic Bayesian Solution of the Two-Stage Poisson-Type Problem in Probabilistic Risk Analysis. *Risk Analysis*, *5*(3).

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press.

Gelman, A., Simpson, D., & Betancourt, M. (2017). The Prior Can Often Only Be Understood in the Context of the Likelihood. *Entropy*, *19*(10), 555. https://doi.org/10.3390/e19100555

Gouet, R., López, F. J., Maldonado, L., & Sanz, G. (2019). Statistical Inference for the Weibull Distribution Based on δ-Record Data. *Symmetry*, *12*(1), 20. https://doi.org/10.3390/sym12010020

Grabski, F., & Sarhan, A. (1996). Empirical Bayes estimation in the case of exponential reliability. *Reliability Engineering and System Safety*, *53*, 105–113.

Gribok, A., Agarwal, V., & Yadav, V. (2020). Performance of empirical Bayes estimation techniques used in probabilistic risk assessment. *Reliability Engineering & System Safety*, *201*, 106805. https://doi.org/10.1016/j.ress.2020.106805

Guan, X., He, J., Jha, R., & Liu, Y. (2012). An efficient analytical Bayesian method for reliability and system response updating based on Laplace and inverse first-order reliability computations. *Reliability Engineering and System Safety*, *97*(1), 1–13. https://doi.org/10.1016/j.ress.2011.09.008

Guo, J., Li, Z., & Keyser, T. (2018). A Bayesian approach for integrating multilevel priors and data for aerospace system reliability assessment. *Chinese Journal of Aeronautics*, *31*(1), 41–53. https://doi.org/10.1016/j.cja.2017.08.020

Hamada, M. S., Wilson, A. G., Reese, C. S., & Martz, H. F. (2008). *Bayesian Reliability*. Springer New York. https://doi.org/10.1007/978-0-387-77950-8

Hamdan, S. a. (2008). Hybrid Particle Swarm Optimiser using multi-neighborhood topologies. *INFOCOMP Journal of Computer Science*, *7*(1), 36–43. http://177.105.60.18/index.php/infocomp/article/view/204

Hameed, Z., Vatn, J., & Heggset, J. (2011). Challenges in the reliability and maintainability data collection for offshore wind turbines. *Renewable Energy*, *36*(8), 2154–2165. https://doi.org/10.1016/j.renene.2011.01.008

Hartley, D., & French, S. (2021). A Bayesian method for calibration and aggregation of expert judgement. *International Journal of Approximate Reasoning*, *130*, 192–225. https://doi.org/10.1016/j.ijar.2020.12.007

Hoffman, M. D., & Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*, 1593–1623.

ISO. (2016). *ISO 14224: Petroleum, petrochemical and natural gas industries - Collection and exchange of reliability and maintenance data for equipment.*

ISO. (2020). *ISO 25000: Software product Quality Requirements and Evaluation.* https://iso25000.com/index.php/en/iso-25000-standards/iso-25010/62

Jackson, C., & Mosleh, A. (2016). Bayesian inference with overlapping data: Reliability estimation of multi-state on-demand continuous life metric systems with uncertain evidence. *Reliability Engineering and System Safety*, *145*, 124–135. https://doi.org/10.1016/j.ress.2015.09.006

Jerez, D. J., Jensen, H. A., & Beer, M. (2022). An effective implementation of reliability methods for Bayesian model updating of structural dynamic models with multiple uncertain parameters. *Reliability Engineering and System Safety*, *225*. https://doi.org/10.1016/j.ress.2022.108634

Jia, X., & Guo, B. (2022). Reliability analysis for complex system with multi-source data integration and multi-level data transmission. *Reliability Engineering and System Safety*, *217*. https://doi.org/10.1016/j.ress.2021.108050

Kaminskiy, M. P., & Krivtsov, V. V. (2005). A Simple Procedure for Bayesian Estimation of the Weibull Distribution. *IEEE Transactions on Reliability*, *54*(4), 612–616. https://doi.org/10.1109/TR.2005.858093

Kaplan, S. (1983). On a "Two-Stage" Bayesian Procedure for Determining Failure Rates from Experiential Data. *IEEE Transactions on Power Apparatus and Systems*, *102*(1).

Karandikar, J. M., Schmitz, T. L., & Abbas, A. E. (2014). Application of Bayesian Inference to Milling Force Modeling. *Journal of Manufacturing Science and Engineering*, *136*(2). https://doi.org/10.1115/1.4026365

Kelly, D. L., & Smith, C. L. (2009). Bayesian inference in probabilistic risk assessment—The current state of the art. *Reliability Engineering & System Safety*, *94*(2), 628–643. https://doi.org/10.1016/j.ress.2008.07.002

Kodoth, M., Aoyama, S., Sakamoto, J., Kasai, N., Khalil, Y., Shibutani, T., & Miyake, A. (2020). Leak frequency analysis for hydrogen-based technology using bayesian

and frequentist methods. *Process Safety and Environmental Protection*, *136*, 148–156. https://doi.org/10.1016/j.psep.2020.01.025

Kumari, P., Halim, S. Z., Kwon, J. S.-I., & Quddus, N. (2022). An integrated risk prediction model for corrosion-induced pipeline incidents using artificial neural network and Bayesian analysis. *Process Safety and Environmental Protection*, *167*, 34–44. https://doi.org/10.1016/j.psep.2022.07.053

Kundu, D. (2008). Bayesian Inference and Life Testing Plan for the Weibull Distribution in Presence of Progressive Censoring. *Technometrics*, *50*(2), 144–154. https://doi.org/10.1198/004017008000000217

Kundu, D., & Gupta, R. D. (2007). A convenient way of generating gamma random variables using generalized exponential distribution. *Computational Statistics & Data Analysis*, *51*(6), 2796–2802. https://doi.org/10.1016/j.csda.2006.09.037

Langseth, H., Haugen, K., & Sandtorv, H. (1998). Analysis of OREDA data for maintenance optimisation. *Reliability Engineering & System Safety*, *60*, 103–110.

Langseth, H., & Lindqvist, B. H. (2006). Competing risks for repairable systems: A data study. *Journal of Statistical Planning and Inference*, *136*(5), 1687–1700. https://doi.org/10.1016/j.jspi.2004.10.032

Lee, C.-E., Kim, S. U., & Lee, S. (2014). Time-dependent reliability analysis using Bayesian MCMC on the reduction of reservoir storage by sedimentation. *Stochastic Environmental Research and Risk Assessment*, *28*(3), 639–654. https://doi.org/10.1007/s00477-013-0779-x

Lin, J. C. (2002). The use of indirect evidence for Bayesian reliability analysis. *2002 Proceedings of the Annual Reliability and Maintainability Symposium*, 429–433. https://doi.org/10.1109/rams.2002.981679

Ling, C., Lei, J., & Kuo, W. (2024). A sequential two-stage approach based on variational Bayesian inference for reliability-redundancy allocation. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *238*(1), 136–157. https://doi.org/10.1177/1748006X221130123

Liu, J. S. (2004). *Monte Carlo Strategies in Scientific Computing*. Springer New York. https://doi.org/10.1007/978-0-387-76371-2

Macedo, J. B., Maior, C. S., Lins, I. D., Azevedo, R., Moura, M. das C., da Silva, M. F., Silva Nóbrega, M. V. da, Vitale, G., & Vasques, R. R. (2023). Using experts' opinion for Bayesian prior reliability distribution of on-demand equipment: A case study of a novel sliding sleeve valve for open-hole wells. *Reliability Engineering & System Safety*, *238*, 109430. https://doi.org/10.1016/j.ress.2023.109430

Mahmoudi, J. (2021). SIL analysis of subsea control system components based on a typical OREDA database. *Quality and Reliability Engineering International*, *37*(8), 3297–3313. https://doi.org/10.1002/qre.2909

Maior, C. B. S., Macêdo, J. B., Lins, I. D., Moura, M. C., Azevedo, R. V., Santana, J. M. M., da Silva, M. F., & Magalhães, M. V. C. (2022). Bayesian prior distribution based on generic data and experts' opinion: A case study in the O&G industry. *Journal of Petroleum Science and Engineering*, *210*. https://doi.org/10.1016/j.petrol.2021.109891

Martinez, J.-M., Collette, Y., Baudin, M., Christopoulou, M., & Baudin, M. (2013). *pyDOE Design of Experiments for Python*. https://pythonhosted.org/pyDOE/

McShane, B., Adrian, M., Bradlow, E. T., & Fader, P. S. (2008). Count Models Based on Weibull Interarrival Times. *Journal of Business & Economic Statistics*, *26*(3), 369–378. https://doi.org/10.1198/073500107000000278

Montgomery, D. C., & Runger, G. C. (2013). *Applied Statistics and Probability for Engineers* (6th ed.). John Wiley & Sons.

Mosleh, A., Bier, V. M., & Apostolakis, G. (1988). A critique of current practice for the use of expert opinions in probabilistic risk assessment. *Reliability Engineering & System Safety*, *20*(1), 63–85. https://doi.org/10.1016/0951-8320(88)90006-3

Mundform, D. J., Schaffer, J., Kim, M.-J., Shaw, D., Thongteeraparp, A., & Supawan, P. (2011). Number of Replications Required in Monte Carlo Simulation Studies: A Synthesis of Four Studies. *Journal of Modern Applied Statistical Methods*, *10*(1), 19–28. https://doi.org/10.22237/jmasm/1304222580

Neal, R. M. (2011). MCMC Using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo* (pp. 113–162). Chapman and Hall/CRC. https://doi.org/10.1201/b10905-6

Nelson, W. (1982). *Applied Life Data Analysis*. JohnWiley&Sons.

Noel, M. M. (2012). A new gradient based particle swarm optimization algorithm for accurate computation of global minimum. *Applied Soft Computing*, *12*(1), 353–359. https://doi.org/10.1016/j.asoc.2011.08.037

Osheyor Gidiagba, J., Daraojimba, C., Ayo Ogunjobi, O., Anthony Ofonagoro, K., & Adepoju Fawole, A. (2023). ADVANCING OFFSHORE OIL AND GAS FACILITIES: A COMPREHENSIVE REVIEW OF INNOVATIVE MAINTENANCE STRATEGIES FOR ENHANCED RELIABILITY AND EFFICIENCY. *Economic Growth and Environment Sustainability*, *2*(2), 84–95. https://doi.org/10.26480/egnes.02.2023.84.95

Porn, K. (1996). The two-stage Bayesian method used for the T-Book application. *Reliability Engineering and System Safety*, *51*, 169–179.

Rausand, M., & Oien, K. (1996). The basic concepts of failure analysis. *Reliability Engineering and System Safety*, *53*, 73–83.

Ren, H., Guo, Q., Cao, Z., & Ren, H. (2023). Risk prediction for petroleum exploration based on Bayesian network classifier. *Geoenergy Science and Engineering*, *228*, 211924. https://doi.org/10.1016/j.geoen.2023.211924

René Van Dorp, J., & Mazzuchi, T. A. (2004). A general Bayes exponential inference model for accelerated life testing. *Journal of Statistical Planning and Inference*, *119*(1), 55–74. https://doi.org/10.1016/S0378-3758(02)00411-1

Riddell, A., Hartikainen, A., & Carter, M. (2021). *PyStan*. https://pypi.org/project/pystan

Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer New York. https://doi.org/10.1007/978-1-4757-4145-2

Roberts, G. O., & Rosenthal, J. S. (2004). General state space Markov chains and MCMC algorithms. *Probability Surveys*, *1*(none), 20–71. https://doi.org/10.1214/154957804100000024

Sales da Cunha, B., das Chagas Moura, M., Azevedo, R., Marques Santana, J. M., Bezerra Souto Maior, C., Didier Lins, I., Mendes, R., Nogueira Lima, E., Campos Lucas, T., Siqueira, P. G., & Souza Vidal de Negreiros, A. C. (2024). A Bayesian approach for reliability estimation for non-homogeneous and interval-censored

failure data. *Process Safety and Environmental Protection*, *182*, 775–788. https://doi.org/10.1016/j.psep.2023.11.080

Sandtorv, H. A., & Thompson, D. W. (1996). Practical experiences with a data collection project: the OREDA project. *Reliability Engineering and System Safety*, *51*, 159–167.

Santana, J. M., Cunha, B., Aichele, D., Azevedo, R., Moura, M. das C., Maior, C., Lins, I., Mendes, R., Lima, E., & . E. D. (2022). PetroBayes; An Effortless Software to Perform Bayesian Reliability Estimation. *32nd European Safety and ReliabilityConference (ESREL 2022)*, 1806–1812. https://doi.org/10.3850/978-981-18-5183-4_s01-10-543-cd

Sedehi, O., Papadimitriou, C., & Katafygiotis, L. S. (2019). Probabilistic hierarchical Bayesian framework for time-domain model updating and robust predictions. *Mechanical Systems and Signal Processing*, *123*, 648–673. https://doi.org/10.1016/j.ymssp.2018.09.041

Sheikholeslami, R., & Razavi, S. (2017). Progressive Latin Hypercube Sampling: An efficient approach for robust sampling-based analysis of environmental models. *Environmental Modelling & Software*, *93*(April 2018), 109–126. https://doi.org/10.1016/j.envsoft.2017.03.010

Shultis, J. K., Johnson, D. E., Miliken, G. A., & Eckhoff, N. D. (1981). *Use of non-conjugate prior distributions in compound failure models*.

SINTEF, & NTNU. (2015). *Offshore and Onshore Reliability Data, Volume 1 - Topside Equipment* (6th Editio).

Siqueira, P. G., Lucas, T., Cunha, B., Santana, J. M., Azevedo, R., Moura, M., Lins, I., Maior, C. S., & Lima, E. (2023). *PetroBayes' Modules for Reliability Assessment for Oil and Gas Industry*. 2236–2237. https://doi.org/10.3850/978-981-18-8071-1_p418-cd

Siqueira, P. G., Moura, M. das C., & Duarte, H. O. (2022). A Bayesian population variability based method for estimating frequency of maritime accidents. *Process Safety and Environmental Protection*, *163*, 308–320. https://doi.org/10.1016/j.psep.2022.05.035

Siu, N. O., & Kelly, D. L. (1998). Bayesian parameter estimation in probabilistic risk assessment. *Reliability Engineering and System Safety*, *62*, 89–116.

Smith, D. J. (2005). Understanding terms and jargon. In *Reliability, Maintainability and Risk* (pp. 11–23). Elsevier. https://doi.org/10.1016/B978-075066694-7/50003-X

Soland, R. M. (1969). Bayesian Analysis of the Weibull Process with Unknown Scale and Shape Parameters. *IEEE Transactions on Reliability*, *R-18*(4), 181–184. https://doi.org/10.1109/TR.1969.5216348

Straub, D., & Papaioannou, I. (2015). Bayesian Updating with Structural Reliability Methods. *Journal of Engineering Mechanics*, *141*(3). https://doi.org/10.1061/(asce)em.1943-7889.0000839

Tanizaki, H. (2008). A Simple Gamma Random Number Generator for Arbitrary Shape Parameters. *Economics Bulletin*, *3*(7), 1–10.

Taofeek Popoola, L., Shehu Grema, A., Kayode Latinwo, G., Gutti, B., & Saheed Balogun, A. (2013). Corrosion problems during oil and gas production and its mitigation. *International Journal of Industrial Chemistry*, *4*(35). http://www.industchem.com/content/4/1/35

Vaurio, J. K. (1987). On Analytic Empirical Bayes Estimation of Failure Rates. *Risk Analysis*, *7*(3).

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-Normalization, Folding, and Localization: An Improved R for Assessing Convergence of MCMC. *Bayesian Analysis*, *16*(2). https://doi.org/10.1214/20-BA1221

Wang, L., Pan, R., Wang, X., Fan, W., & Xuan, J. (2017). A Bayesian reliability evaluation method with different types of data from multiple sources. *Reliability Engineering and System Safety*, *167*, 128–135. https://doi.org/10.1016/j.ress.2017.05.039

Wang, Z. M., & Yang, J. G. (2012). Numerical method for Weibull generalized renewal process and its applications in reliability analysis of NC machine tools. *Computers and Industrial Engineering*, *63*(4), 1128–1134. https://doi.org/10.1016/j.cie.2012.06.019

Wilson, A. G., & Fronczyk, K. M. (2016). Bayesian Reliability: Combining Information. *Quality Engineering*, 0–0. https://doi.org/10.1080/08982112.2016.1211889

Wu, S. J., Chen, D. H., & Chen, S. T. (2006). Bayesian inference for Rayleigh distribution under progressive censored sample. *Applied Stochastic Models in Business and Industry*, *22*(3), 269–279. https://doi.org/10.1002/asmb.615

Xi, B., Ming Tan, K., & Liu, C. (2013). Logarithmic Transformation-Based Gamma Random Number Generators. *Journal of Statistical Software*, *55*(4). http://www.jstatsoft.org/

Yang, L., Guo, Y., & Kong, Z. (2019). On the performance evaluation of a hierarchical-structure prototype product using inconsistent prior information and limited test data. *Information Sciences*, *485*, 362–375. https://doi.org/10.1016/j.ins.2019.02.018

Yang, Q., Hong, Y., Chen, Y., & Shi, J. (2012). Failure Profile Analysis of Complex Repairable Systems With Multiple Failure Modes. *IEEE Transactions on Reliability*, *61*(1), 180–191. https://doi.org/10.1109/TR.2011.2182225

Zhou, T., Modarres, M., & Droguett, E. L. (2019). Multi-unit risk aggregation with consideration of uncertainty and bias in risk metrics. *Reliability Engineering & System Safety*, *188*, 473–482. https://doi.org/10.1016/j.ress.2019.04.001

APPENDIX A – PRIOR DISTRIBUTION ESTIMATION RESULTS BY VARYING THE PSO SEARCH SPACE

In this section, the results of the prior distribution estimation are presented for different values of the constants $\{cs_{min}, cs_{max}\}$ that determine the search space for $S[\alpha]$. The values considered are: $\{10,55\}$, $\{20,65\}$, $\{30,75\}$, $\{40,95\}$. For each search space, different test sizes are considered, and 5 independent prior estimations are executed. The results are presented in terms of the hyperparameters of the prior distribution and are analyzed in terms of the log-likelihood (log-L), p-value and NRMSE.

Results for $cs_{min} = 10$ and $cs_{max} = 55$.

Test	a_α	b_α	a_β	b_β	log-L	p-value	NRMSE
	6.65	0.0017	11.17	6.5456	-4132.05	0.8693	1.12684
	7.08	0.0019	10.65	6.5216	-4131.84	0.7166	1.09151
50x50	6.87	0.0018	11.24	6.645	-4131.73	0.8693	1.15553
	6.78	0.0018	9.74	5.7703	-4132.13	0.2719	1.25332
	5.96	0.0015	10.86	6.5689	-4132.18	0.7166	1.04458
	6.54	0.0017	9.23	5.7739	-2059.33	0.0678	1.01936
	6.96	0.0019	9.96	6.2453	-2058.81	0.8693	1.09193
50x25	8.20	0.0022	10.05	6.328	-2058.7	0.2719	0.94752
	7.07	0.0019	8.56	5.2501	-2059.5	0.5487	1.16867
	6.83	0.0018	10.03	6.288	-2058.81	0.8693	0.99731
	8.26	0.002	8.98	5.7798	-342.515	0.7166	0.99054
50x05							
	9.07	0.0022	9	5.8127	-342.612	0.9667	0.96044

Test	a_α	b_α	a_β	b_β	log-L	p-value	NRMSE
	7.83	0.0019	9.25	6.0202	-342.499	0.8693	0.9847
	6.98	0.0017	9.25	6.0662	-342.71	0.7166	1.01354
	9.58	0.0024	9.67	6.2094	-342.642	0.8693	1.0553
	6.36	0.0017	9.21	6.041	-2088.23	0.915	0.97036
	7.08	0.0019	10.23	6.3569	-2087.99	0.7102	1.06394
25x50	6.91	0.0017	10.7	6.4757	-2088.43	0.0779	0.91335
	6.52	0.0016	10.62	6.514	-2088.42	0.4755	0.90234
	7.11	0.0019	10.2	6.3176	-2088.08	0.915	1.0652
	7.09	0.002	8.84	5.8279	-1054.35	0.4755	1.05274
	7.1	0.002	9.25	6.0674	-1054.36	0.9955	1.05104
25x25	5.91	0.0017	9.41	6.0587	-1054.6	0.4755	1.28981
	6.3	0.0017	8.84	5.878	-1054.5	0.4755	0.97068
	5.87	0.0016	9.67	6.2224	-1054.51	0.7102	1.11754
	9.24	0.0025	11.06	6.6399	-185.314	0.4755	1.14086
	7.83	0.002	10.79	6.5321	-185.468	0.9955	1.05185
25x05	8.61	0.0023	11.03	6.5958	-185.347	0.285	1.146
	8.3	0.0021	10.55	6.3088	-185.456	0.9955	1.04
	9.78	0.0026	9.54	5.6037	-185.521	0.7102	1.17551

Test	a_α	b_α	a_β	b_β	log-L	p-value	NRMSE
	7.09	0.0015	6.33	4.5146	-389.937	1	0.62921
	5.45	0.0011	6.78	4.9356	-389.973	0.0794	0.60621
05x50	5.68	0.0011	5.65	4.0804	-390.114	0.873	0.57253
	4.72	0.001	4.19	2.7623	-390.101	0.873	1.08212
	4.21	0.0009	3.93	2.6603	-390.169	0.0794	1.16261
	4.67	0.0009	7.43	5.2428	-192.123	1	0.66619
	5.48	0.0011	8.09	5.3859	-192.153	1	0.72542
05x25	4.75	0.0009	6.53	4.6507	-192.142	0.3571	0.63942
	4.88	0.0009	6.9	4.7605	-192.124	0.873	0.62863
	5.4	0.001	6.46	4.4112	-192.148	0.873	0.61616
	3.5	0.0007	8.93	5.8715	-32.4853	0.3571	0.97192
	3.64	0.0007	8.95	5.9831	-32.4695	0.0794	0.84332
05x5	3.93	0.0007	8.87	5.8841	-32.5051	0.3571	0.74146
	3.69	0.0007	7.27	5.2857	-32.5512	0.3571	0.74565
	4.04	0.0008	8.9	5.9672	-32.4869	0.873	0.82401

Results for $cs_{min} = 20$ and $cs_{max} = 65$.

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
50x50	5.1	0.0021	0.95	0.6409	-4143.66	0.1786	5698.24

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	10.59	0.0027	4.48	2.6659	-4140.66	0.3959	1.34218
	8.61	0.0023	10.96	6.6442	-4132.52	0.5487	0.99986
	6.13	0.0016	10.61	6.4675	-4132.17	0.7166	1.11822
	6.66	0.0018	9.57	5.6952	-4132.58	0.5487	1.32317
	8.02	0.0021	9.9	6.1589	-2059.05	0.8693	0.93136
	7.93	0.0021	9.75	6.2449	-2058.76	0.1786	0.89054
50x25	7.65	0.002	9.95	6.2822	-2059.05	0.5487	0.90671
	8.3	0.0022	9.77	6.2017	-2058.83	0.1124	0.89768
	8.38	0.0022	9.9	6.1773	-2058.95	0.3959	0.91051
	9.51	0.0023	8.57	5.5994	-342.643	0.5487	0.84759
	9.7	0.0024	9.71	6.2305	-342.629	0.2719	0.97845
50x05	10.59	0.0026	7.82	5.0527	-343.037	0.9667	0.97123
	9.49	0.0023	8.33	5.4052	-342.684	0.0678	0.95148
	8.95	0.0022	9.15	6.0292	-342.635	0.1786	0.94905
	6.29	0.0018	9.71	5.8557	-2088.94	0.9955	1.48197
	7.42	0.0019	10.12	6.2778	-2088.72	0.285	0.92184
25x50	6.06	0.0016	10.14	6.2719	-2088.68	0.9955	1.13267
	5.51	0.0014	10.5	6.3827	-2089.04	0.4755	1.14049

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	6.51	0.0018	9.81	6.1285	-2088.82	0.285	1.20353
	7.41	0.002	8.65	5.5971	-1054.04	0.915	0.97835
	5.76	0.0016	8.87	5.7913	-1054.37	0.9955	1.16061
25x25	7.14	0.0019	8.97	5.833	-1054.03	0.4755	0.93579
	7.01	0.0019	9.94	6.2954	-1054.3	0.9955	1.04483
	6.99	0.002	9.12	6.022	-1053.88	0.7102	1.08596
	8.88	0.0023	11.34	6.6844	-185.409	0.4755	1.08997
	11.38	0.0031	10.07	6.3118	-185.197	0.4755	1.01012
25x05	10.04	0.0026	10.69	6.5167	-185.198	0.9955	0.98756
	10.18	0.0027	10.29	6.3891	-185.186	0.4755	1.00116
	10.48	0.0028	10.58	6.5033	-185.16	0.0779	1.02318
	4.09	0.0008	6.71	4.4985	-390.104	0.3571	0.81133
	4.13	0.0009	6.75	4.5549	-390.091	0.873	1.01428
05x50	3.91	0.0008	6.7	4.6366	-390.086	1	0.86019
	4.89	0.001	6.86	4.3956	-390.129	0.3571	0.87709
	5.74	0.0012	4.74	3.2721	-390.117	0.3571	0.77609
05x25	4.9	0.0009	5.77	3.7929	-191.976	0.873	0.69579
	5.1	0.001	5.97	3.8997	-192.006	1	0.78709

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	5.09	0.001	6.34	4.0651	-192.003	0.873	0.8158
	4.8	0.0009	5.47	3.6565	-191.984	0.873	0.71599
	4.91	0.0009	5.76	3.8496	-191.988	0.873	0.6726
	2.89	0.0005	8.73	5.8973	-32.5589	0.873	0.84753
	2.69	0.0005	9.27	6.0708	-32.6125	0.873	1.14079
05x05	3.47	0.0007	8.46	5.6922	-32.5769	1	0.9495
	3.14	0.0006	8.52	5.7754	-32.5613	0.873	0.92426
	3.39	0.0006	8.12	5.692	-32.536	0.873	0.74383

Results for $cs_{min} = 30$ and $cs_{max} = 75$.

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	6.69	0.0017	10.1	6.3058	-4131.65	0.0058	0.91611
	16.28	0.0051	325.95	191.7807	-4135.45	0.1124	1.06509
50x50	7.05	0.0018	10.38	6.3994	-4131.22	0.7166	0.93166
	7.4	0.0019	10.42	6.4457	-4131.14	0.5487	0.91352
	7.78	0.002	11.19	6.6761	-4131.39	0.1786	0.98525
	7.61	0.002	9.98	6.3217	-2058.58	0.1786	0.91223
50x25	8.58	0.0023	9.65	6.2044	-2058.93	0.5487	0.8834
	7.59	0.002	9.73	6.1572	-2058.68	0.8693	0.92584

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	7.15	0.0019	9.85	6.2739	-2058.67	0.5487	0.95411
	7.54	0.002	9.14	5.8533	-2058.88	0.5487	0.92985
	7.87	0.0019	9.05	5.9851	-342.762	0.2719	0.94852
	6.28	0.0015	8.31	5.498	-343.457	0.8693	1.01559
50x05	8.47	0.002	7.48	4.8923	-343.167	0.5487	0.94236
	8.57	0.0021	8.46	5.5167	-342.815	0.8693	0.9853
	8.88	0.0022	9.22	5.7203	-343.034	0.9667	1.08445
	5.9	0.0015	9.67	6.1225	-2088.83	0.4755	0.87057
	6.11	0.0017	10.7	6.4678	-2088.8	0.7102	1.38392
25x50	5.89	0.0016	10.7	6.5234	-2088.75	0.4755	1.2844
	6.05	0.0015	9.42	5.9188	-2088.83	0.0356	1.03369
	6.98	0.0019	10.52	6.4705	-2088.7	0.7102	1.14249
	4.45	0.0018	25.04	16.2288	-1056.37	0.1558	2.76767
	7.56	0.0021	8.69	5.7792	-1053.98	0.4755	0.96489
25x25	7.82	0.0021	8.99	5.8309	-1054.1	0.4755	0.92474
	7.4	0.0021	7.82	5.2866	-1054.19	0.7102	1.01239
	6.19	0.0017	8.96	5.9878	-1054.22	0.4755	1.00266
25x5	8.72	0.0023	11.06	6.571	-185.274	0.4755	1.12795

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	8.5	0.0023	11.1	6.6604	-185.264	0.285	1.17176
	9	0.0024	10.68	6.4678	-185.272	0.7102	1.0975
	8.35	0.0023	10.3	6.2709	-185.289	0.7102	1.20183
	8.5	0.0023	10.61	6.445	-185.267	0.7102	1.14715
	5.77	0.0012	4.91	3.3299	-389.927	0.873	0.7841
	5.42	0.0011	5.42	3.6671	-389.987	0.873	0.7499
05x50	5.85	0.0012	4.79	3.2341	-389.95	0.873	0.76728
	3.88	0.0008	4.46	2.886	-390.07	0.873	1.22967
	5.74	0.0012	5.18	3.5411	-390.937	0.873	0.77088
	5.31	0.001	6.97	4.5324	-192.149	0.873	0.69513
	4.92	0.0009	6.19	4.2437	-192.043	0.873	0.63293
05x25	4.95	0.0009	5.99	4.1004	-192.12	0.873	0.62774
	4.57	0.0009	6.69	4.7411	-192.055	0.873	0.70256
	4.87	0.0009	6.61	4.5498	-192.038	0.3571	0.63589
	3.3	0.0006	8.23	5.7363	-32.4835	0.873	0.78156
	3.11	0.0006	8.09	5.5977	-32.5141	0.873	0.91296
05x05	3.09	0.0006	7.75	5.5386	-32.4982	0.873	0.87192
	3.5	0.0007	7.51	5.4542	-32.5639	0.873	0.8148

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	3.32	0.0006	7.82	5.5898	-32.5027	0.873	0.75394

Results for $cs_{min} = 40$ and $cs_{max} = 95$.

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	4.74	0.0012	10.26	6.3256	-4133.44	0.1786	1.20945
	553.62	0.1378	4.87	2.9383	-4140.41	0.1786	0.74433
50x50	5.34	0.002	1.69	0.9315	-4147.08	0.0058	477.535
	5.1	0.0013	10.53	6.4207	-4133.7	0.7166	1.19745
	5.46	0.0014	10.47	6.4523	-4133.27	0.1786	1.11448
	4.83	0.0013	9.38	6.0157	-2060.66	0.5487	1.25578
	4.96	0.0013	9.74	6.1121	-2060.49	0.2719	1.21992
50x25	5.12	0.0013	9.23	5.6049	-2060.94	0.9667	1.23988
	4.81	0.0013	7.95	4.9995	-2061.31	0.8693	1.45527
	4.94	0.0013	9.65	6.0646	-2060.51	0.0678	1.23351
	5.29	0.0013	8.76	5.9053	-344.01	0.0678	1.09989
	5.56	0.0013	8.35	5.6606	-344.112	0.1786	0.97846
50x05	5.28	0.0013	8.45	5.7647	-344.137	0.0217	1.08392
	5.09	0.0012	7.43	4.9843	-344.378	0.9977	1.08205
	5.5	0.0013	7.84	5.3193	-344.253	0.1786	1.01162

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	7.6	0.0041	12.07	7.2701	-2091.73	0	7.001
	4.6	0.0012	10.04	6.3362	-2088.28	0.285	1.25708
25x50	5.35	0.0014	10.51	6.314	-2088.27	0.4755	1.31539
	4.93	0.0013	10.27	6.2382	-2088.21	0.915	1.38875
	4.25	0.0012	9.03	5.7672	-2088.79	0.4755	1.64148
	4.53	0.0013	8.19	5.5817	-1055.15	0.7102	1.32354
	4.38	0.0012	8.48	5.8042	-1055.14	0.915	1.4854
25x25	4.41	0.0012	8.42	5.5948	-1055.13	0.4755	1.27471
	4.5	0.0012	9.16	5.9959	-1055.05	0.7102	1.22063
	4.55	0.0013	9.07	5.9062	-1054.97	0.1558	1.47929
	6.14	0.0015	11.09	6.5617	-185.999	0.915	1.11563
	6.01	0.0015	11.13	6.4703	-186.02	0.285	1.23798
25x05	5.91	0.0015	10.13	6.0695	-186.013	0.7102	1.2163
	5.9	0.0015	10.78	6.4772	-185.956	0.7102	1.19905
	5.99	0.0015	11.07	6.5659	-185.939	0.285	1.18538
	5.69	0.0012	6.23	4.8094	-389.92	0.3571	0.58447
05x50	5.64	0.0011	7.36	5.3049	-389.941	0.873	0.56265
	5.28	0.0011	4.5	3.2781	-389.998	0.873	0.71929

Test	a_α	b_α	a_β	b_β	log-L	p-valor	NRMSE
	5.63	0.0012	5.24	3.7106	-390.02	1	0.74622
	5.64	0.0012	5.79	3.8935	-389.99	0.3571	0.81648
	5.15	0.001	5.92	4.2893	-192.003	0.3571	0.63258
	5.11	0.001	5.25	3.5488	-192.031	0.873	0.74853
05x25	5.25	0.001	6.03	4.0791	-192.017	0.873	0.67877
	4.39	0.0008	5.04	3.4951	-192.005	0.873	0.67669
	4.86	0.0009	5.05	3.4643	-191.99	0.3571	0.67328
	2.98	0.0006	8.34	5.7283	-32.5393	0.873	1.04689
	3.16	0.0006	8.87	5.9535	-32.5034	0.3571	0.9219
05x05	3.02	0.0006	8.78	5.7586	-32.6015	0.873	1.13286
	3.02	0.0006	7.68	5.532	-32.5055	0.3571	0.91019
	2.83	0.0006	8.74	5.9043	-32.532	0.873	1.29858