UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Frailty model for multiple repairable systems hierarchically represented in serial/parallel structures under assumption of ARAm imperfect repairs

Oilson Alberto Gonzatto Junior

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGEs)

ICMC USP SÃO CARLOS

Oilson Alberto Gonzatto Junior

Frailty model for multiple repairable systems hierarchically represented in serial/parallel structures under assumption of ARAm imperfect repairs

Doctoral dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP and to the Department of Statistics – DEs-UFSCar, in partial fulfillment of the requirements for the degree of the Doctorate Interagency Program Graduate in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Francisco Louzada Neto
Co-advisor: Profa. Dra. Vera Lucia Damasceno Tomazella

USP – São Carlos
June 2021

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

G643f

Gonzatto Junior, Oilson Alberto
Frailty model for multiple repairable systems hierarchically represented in serial/parallel structures under assumption of ARAm imperfect repairs / Oilson Alberto Gonzatto Junior; orientador Francisco Louzada Neto; coorientador Vera Lucia Damasceno Tomazella. -- São Carlos, 2021.
173 p.

Tese (Doutorado - Programa Interinstitucional de Pós-graduação em Estatística) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2021.

1. Repairable systems. 2. power law process. 3. competing risks. 4. series-parallel structure. 5. frailty model. I. Louzada Neto, Francisco, orient. II. Tomazella, Vera Lucia Damasceno, coorient. III. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Oilson Alberto Gonzatto Junior

Modelo de fragilidade para múltiplos sistemas reparáveis hierarquicamente representados em estruturas série/paralelas sob a suposição de reparos imperfeitos ARAm

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Doutor em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística. *VERSÃO REVISADA*

Área de Concentração: Estatística

Orientador: Prof. Dr. Francisco Louzada Neto
Coorientador: Profa. Dra. Vera Lucia Damasceno Tomazella

USP – São Carlos
Junho de 2021

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Defesa de Tese de Doutorado do candidato Oilson Alberto Gonzatto Junior, realizada em 27/04/2021.

Comissão Julgadora:

Prof. Dr. Francisco Louzada Neto (USP)

Prof. Dr. Enrico Antonio Colosimo (UFMG)

Prof. Dr. Vinicius Fernando Calsavara (CIPE)

Prof. Dr. Pedro Luiz Ramos (-)

Prof. Dr. Jeremias da Silva Leão (UFAM)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.

ACKNOWLEDGEMENTS

I use this space to express my sincere thanks to everyone (and there were many!) That helped me in the elaboration of this research and that made this work possible.

Heartfelt thanks to my supervisor Vera Tomazella for all the guidance, for taking the lead in choosing the theme, about which I knew so little, and which I now know a little more. I appreciate all the care with the deadlines, the suggestions and, mainly, the joy with which you work with us!

Thank you Professor Francisco Louzada for the invitation to participate in the project that motivated this research, this was an unexpected but very timely change! I thank you for running the project as a whole, for fighting for our team and keeping us at the forefront of Petrobras' priorities! Furthermore, I am very grateful for the unique experience of having you, too, as a mentor. There were so many monitorings, courses, workshops and projects, and all this with so many different themes, and often simultaneously! hahahah After the despair is over, there is only thanks! hahahah

I also want to thank in a very special way the others teammates who helped me in the idealization of methods, obtaining data, writing texts and publishing the results. This thanks is yours (in alphabetical order not to commit to anyone) Alex Mota, Eder Milani, Felipe Alegria, José Cuminato, Gustavo Bochio, Ivan Pericini, Oscar Rodriguez, Paulo Henrique, Pedro Ramos, Reza Naki! Thank you very much indeed, this work was done with multiple hands!

Thank you very much to other friends, who gave me all the support to face this stage. Again, in alphabetical order, thank you Ana Paula Jorge, Anderson Ara, Caio Pena, Camila Ozelame, Carlos Pedraza, Claudia Montecino, Diego Nascimento, Gabriel Espírito Santo, Iman Ghodratitoostani, Javier (little son of Claudia and Caio), Josimara Tatiane, Marcos Jardel, Milton Miranda and Vitor Amorim! Thank you all for the many many many good times of busy work and many others of relaxation!

Thanks to all the professors and other employees of the PIPGEs, who keep everything running and allow us to maintain greater dedication to research. Finally, I would like to thank our social structure as a whole, which still contributes to the maintenance of research funding programs. In particular, mine received financial support from CAPES and Petrobras.

ABSTRACT

GONZATTO JUNIOR, O. A. **Frailty model for multiple repairable systems hierarchically represented in serial/parallel structures under assumption of ARAm imperfect repairs**. 2021. 173 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2021.

The main objective of this thesis is to extend the methodology used to treat failure time data. In particular, we wish to propose an appropriate modeling to a context of hierarchically represented repairable systems, subject to competitive risks and unobserved heterogeneity. To do that, we took one of the necessary steps, we propose modeling for a single repairable system with a hierarchical structure under the assumption that the failures follow a non-homogeneous Poisson process with a power-law intensity function under the frequentist and bayesian framework. In this context we used a corrective approach to remove bias with order $O(n^{-1})$, and the respective exact confidence intervals are proposed. We illustrate the use of both methods with an early-stage real project related to the traction system of an in-pipe robot. In the sequence, we introduced a framework to the reliability estimation in systems with serial structure and failure modes structured in a parallel way, we continued the studies of the robotic unit previously analyzed. Finally, we propose a statistical model to the reliability estimation of groups of repairable systems hierarchically represented, under a competing risks framework, with the consideration of the existence of unobserved heterogeneity that acts individually on the systems of each group, and also the possibility of imperfect repairs. To illustrate, we consider a database with the failures of agricultural machines categorized in different groups.

Keywords: Bias correction. Competing risks. Hierarchical systems. Maximum likelihood estimation. Power-law process..

RESUMO

GONZATTO JUNIOR, O. A. **Modelo de fragilidade para múltiplos sistemas reparáveis hierarquicamente representados em estruturas série/paralelas sob a suposição de reparos imperfeitos ARAm**. 2021. 173 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2021.

O principal objetivo desta tese é estender a metodologia utilizada para tratar dados de tempo de falha. Particularmente, buscamos propor uma modelagem apropriada a um contexto de sistemas reparáveis representados hierarquicamente, sujeitos a riscos competitivos e heterogeneidade não observada. Para isso, demos um passo necessário propondo a modelagem de um único sistema reparável com uma estrutura hierárquica, pressupondo que as falhas seguem um processo de Poisson não-homogêneo com uma função de intensidade lei de potência sob um panorama frequentista e bayesiano. Nesse contexto, nós utilizamos uma abordagem corretiva para remover o viés de ordem $O(n^{-1})$, e os respectivos intervalos de confiança exatos também foram propostos. Nós ilustramos o uso de ambos os métodos em dados vindos de um projeto real em estágio inicial relacionados ao sistema de tração de uma unidade robótica. Na sequência, introduzimos um panorama para a estimação da confiabilidade em sistemas seriados com modos de falha estruturados de forma paralela, nós demos continuidade aos estudos da unidade robótica previamente analisada. Por fim, nós propusemos um modelo estatístico para a estimação da confiabilidade de grupos de sistemas reparáveis hierarquicamente representados, sob a suposição de riscos competitivos e a consideração da existência de heterogeneidade não observada que atua individualmente nos sistemas dentro de cada grupo, e também a possibilidade de reparos imperfeitos. Para ilustrar, consideramos um conjunto de dados com o registro de falhas de máquinas agrícolas categorizadas em diferentes grupos.

Palavras-chave: Correção de Viés. Riscos Competitivos. Sistemas Hierárquicos. Estimação de Máxima Verossimilhança. Processo Lei de Potência..

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

MTBF_n	Mean Time Between the $(n-1)$-th and n-th Failure
ARA	Arithmetic Reduction of Age
ARI	Arithmetic Reduction of Intensity
BMAP	Bias-Corrected MAP
BOP	BlowOut Preventer
CMLE	Bias-Corrected MLE
CP	Coverage Probabilities
FTA	Fault Tree Analysis
HPP	Homogeneous Poisson Processes
IR	Imperfect Repair
MAP	Maximum *a posteriori* Probability
MLE	Maximum Likelihood Estimator
MR	Minimal Repair
MRE	Mean Relative Estimate
MSE	Mean Square Error
NHPP	Non-Homogeneous Poisson process
PLP	Power-Law Process
PR	Perfect Repair
RMSE	Root Mean Squared Error
RP	Renewal Processes

CONTENTS

CHAPTER

1

INTRODUCTION

The presence of repeated recurrences of an event of interest often arises in areas such as manufacturing, software development, medical applications, social sciences, and risk analysis, among others. Jiang and Liu (2017) exemplifies by saying that, in reliability engineering, when a complex system such as supercomputers, airplanes or cars is included in a study, several unexpected failures may be exposed by different defects or weaknesses in the products' design, manufacturing, operation, maintenance, and management. Models with this feature, as those exposed by Crowder (2001), are traditionally referred to as competing risks, or equivalently, a system with p components connected in series. A single component failure results in total system failure.

Recently, the availability evaluation of repairable systems with multiple failure modes is at the center of attention due to the broad application in engineering. According to the competing risks framework, a series system fails by the earliest occurrence of failure modes. Therefore, in this thesis, we utilized a model for components, whose failures happen due to one of the series competing failure mechanisms, whereby each of them acts related to the system independently.

A system can be broken down into several sub-systems, and sub-sub-systems compose the sub-systems in a hierarchical form until the elements cannot or are not worthy of being divided. The system's hierarchies can help engineers to better understand the relationships between components and their importance and functions. According to Lieping and Zhe (2009), they can further help engineers to determine the role and acceptable damaging degree of each part of the structure and their influences on the whole system under various external forces and effects.

Thus, structuring a problem according to a hierarchy can help to increase accuracy and facilitate useful analysis of failure factors. Note that the event of interest at the system level is expected to happen at its earliest occurrence. Therefore, a system can be anticipated to follow

a competing risks model. As an example given by Liu, Song and Zhang (2018), mechanical devices (e.g., gear pair and crank train) are always under multiple failure modes (including fracture, corrosion and wear), which compete with each other so that when one kind of failure happens, the device is invalid and other failure modes have no chance to occur anymore.

The components under consideration are repaired upon failure but are also preventively maintained. Thus, the excellent books by Crowder (2001) and Pintilie (2006), among others, motivate the need for accounting for competing risks in reliability and survival applications using several examples in industrial statistics and health sciences. More recently, Langseth and Lindqvist (2006) recorded cumulative service times of a component spanning over 1,600 time units, then marking each failure with its specific causing mode. In this case, the causes were categorized into two broad groups, each with several specified sub-causes. Tuli *et al.* (2000) analyzed repeated shunt failures in infants diagnosed with hydrocephalus, where the failures are known to occur due to a variety of causes.

The focus in this thesis is placed on failure data from repairable systems. Thus, solid modeling and analysis of this data provide equipment operators for better maintenance activities. In the repairable system literature, it is often assumed that failures occur following a Non-Homogeneous Poisson process (NHPP) with power-law intensity. The resulting process is usually referred to as the Power-Law Process (PLP). Proposed by Crow (1974), the PLP is convenient because it is easy to implement, flexible, and the parameters have valuable interpretation, as exemplified by Rigdon and Basu (1989) and Reis, Colosimo and Gilardoni (2019).

In the literature, the PLP has been widely used in modeling software reliability as in the work of Kyparisis and Singpurwalla (1985), reliability growth in the study of Crow (1982), repairable systems in Ascher and Feingold (1984), Engelhardt and Bain (1986), Rigdon and Basu (1989), and others. This assumption, however, implies the hypothesis of a Minimal Repair (MR) model (RAUSAND, 2003), which may not be the most realistic choice, in this sense, a class of Imperfect Repair (IR) models brings the possibility of considering a intermediate level of repair, somewhere between MR and Perfect Repair (PR). This idea is well exposed in works such as those of (KIJIMA; MORIMURA; SUZUKI, 1988; DOYEN; GAUDOIN, 2004; TOLEDO *et al.*, 2015).

Doyen and Gaudoin (2004) were responsible for proposing two of the main models for IR, the Arithmetic Reduction of Age (ARA) and Arithmetic Reduction of Intensity (ARI) model classes, whose repair efficiency is quantified by one of the model parameters. In addition, a constant characterizes the models in terms of the used memory order, which refers to the maximum number of previous information involved in the model's calculations.

Another common assumption is the absence of any unobserved heterogeneity among the causes that compete for the failure of the general system. An immediate consequence of this assumption is the disregard of effects that, although not observed, can have effects on the behavior of failure times (WIENKE, 2010). The literature that adequately addresses this issue

is related to frailty models. A review of these models can be found at (HOUGAARD, 1995; ANDERSEN *et al.*, 1996; WIENKE, 2010). From the perspective of multiplicative frailty, we introduce a multiplicative factor (a random variable) that inflates, deflates or preserves the failure intensity function.

The text here exposed has as its main objective to propose a hierarchical model for a repairable system subject to several failure modes (competing risks). Under MR, it is assumed that each failure mode has a power-law intensity. Hence, we develop a new PLP model with a MR under competing risks, which generalizes the model presented in Somboonsavatdee and Sen (2015). Furthermore, we discuss the inferential procedure for the parameters of the proposed model using the Maximum Likelihood Estimator (MLE), as well as the asymptotic confidence intervals based on the MLEs. Since the sample size is usually small, due to the problem of rare yet adverse failures in industrial scenarios (e.g., in aerospace, nuclear and petrochemical industries) that causes limited failure data availability, we may obtain biased estimators and unreliable asymptotic confidence intervals. To overcome this problem, we suggest a corrective approach to obtain unbiased estimators for the model parameters. Additionally, we discuss how to derive exact confidence intervals (CIs) based on these unbiased estimators.

1.1 Main practical motivation

The challenges in the production of offshore oil wells have been increasing over time, either due to the increase in technical difficulties because of the greater complexity of the areas to be explored, or due to improvements in the rules of the regulatory bodies in order to increase safety. There are two key pillars that should guide an oil well project: safety and productivity.

Petroleum industry loses billions of dollars yearly due to profit loss associated with production lines obstruction. Current flow assurance solutions are troublesome and cost hundreds of millions of dollars annually. Petrobras (abbreviation of *Petróleo Brasileiro S.A.*), which is the Brazil's largest oil and gas producer, has invested in technological innovation projects in order to minimize these losses and increase oil and gas production. Annelida is one of these Petrobras' innovation projects, which has been developed in partnership with the main Brazil's research centers. It regards an in-pipe robot that will be used at a near future to remove hydrates and paraffins that form in pipelines and can cause problems in oil and gas flow (see Figure 1). Several stages of the Annelida project have already been completed and many others are underway, generating important results for the development and improvement of its bases. Given the innovative nature of the project, the reliability modeling of the product has been one of the main objectives of the research centers.

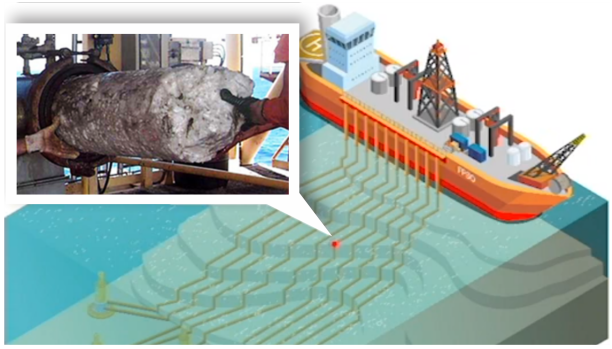


Figure 1 – Hydrates and paraffins are the most common duct obstruction causes.

In reliability engineering, it is well known that the reliability of a product can be assessed from the systems and subsystems that comprise it. Annelida is composed of several systems and subsystems each with well-defined objectives. Due to the high degree of criticality, in this work we consider the traction system of Annelida. In particular, we will study the return and forward locomotives subsystem (modules 11 and 25, in Figure 2), as well as the pressure vessel subsystem (modules 1 to 10 and 24, one of which is represented in module 24, in Figure 2).

The locomotive is responsible for conducting the robot inside the pipe, and once hydrate formation is identified, the robot will work on its safe removal for the oil to flow again. On the other hand, the pressure vessel set is the basic structural module for all the electrical and electronic components of Annelida, which contains 11 of these subsystems. The vessels' purpose is to withstand the forces, pressure and chemical conditions of the environment, safety containing and isolating the components in their interiors. The module also has the function to facilitate the heat exchange, allowing for suitable operational temperature of the electronic components.

A schematic of the studied systems is shown in Figure 2.

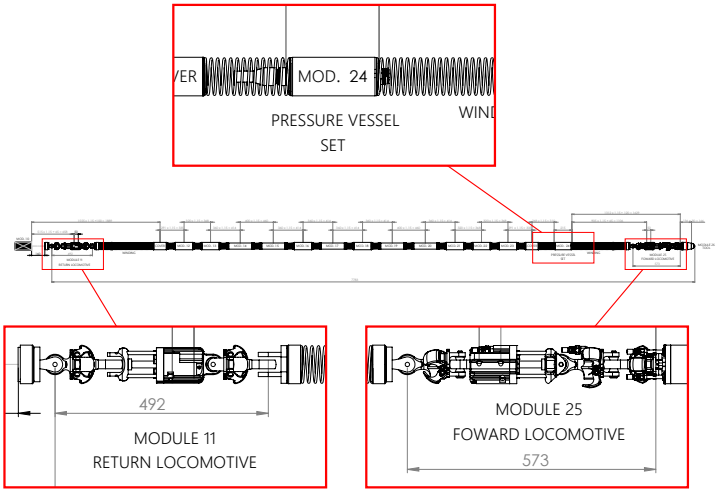


Figure 2 – Annelida's schematic diagram with the representation of one of the pressure vessel set (top), the others (modules 01 to 10) are not represented in this schematic, and traction systems (below).

1.2 Overview of the chapters

The text here exposed is linked in two main chapters.

The Chapter 2, theoretically supports the subject covered throughout the text. This is done by exposing the key concepts about the notion of Systems Repairability (in Section 2.1), where the definitions of Minimum, Perfect, and IR are exposed; of a Count Process (in Section 2.2), whose particular case, the NHPP, is the process used in this study; and Competing Risks (in Section 2.3), where the Latent Failure Times approach is that used in this study.

In the Chapter 3, we present a new statistical model to analyze single repairable systems with a hierarchical structure under the assumption that the repairs are minimal with a PLP intensity and also in the presence of competing risks. Moreover, we discuss classical inference for the model parameters through the MLEs and asymptotic confidence intervals and also perform a simulation study to investigate their properties. We develop improved estimators (bias-corrected MLEs), as well as exact confidence intervals for the model parameters, whose performances are again evaluated through a simulation study. We illustrate our proposed methodology using simulated reliability data of butterfly valves, reliability data based on an in-pipe robot traction system design information (real project in its early stage), and real reliability data of blowout preventer systems. Finally, we conclude the chapter with some final remarks and suggestions for work.

In the Chapter 4 following the previous study (Chapter 3) we describe how complex engineered systems may be modelled hierarchically by use of Bayesian methods. It is also assumed that repairs are minimal and each failure mode has a power-law intensity. Some properties of the new model are discussed. We conduct statistical inference under an objective Bayesian framework. A simulation study is carried out to investigate the efficiency of the proposed methods. Finally, our methodology is illustrated by two practical situations currently addressed in a project under development arising from a partnership between Petrobras and six research institutes.

In the Chapter 5, taking a step forward, in studies (Chapters 3 and 4), the purpose is a direct extension of Chapters 3 in the sense of, the extension comprises the representation of complex systems through a hierarchical structure in series and/or in parallel. For this, we deduce the general form of the model, the likelihood function. We present the mechanism for random numbers generating, which allows obtaining point and interval estimates (via parametric bootstrap) in a more convenient way for reliability curves at any level of the system hierarchy. We illustrate the application of the model with the sequence of application previously presented.

In the Chapter 6 we integrate the previous ideas, and extend into the context of multiple repairable systems with unobserved heterogeneity modeled by a frailty term and also the possibility of IRs whose initial failure rate is in the form of the power law. We did a

simulation study to assess the quality of the proposed estimators. To illustrate, we consider a database with the failures of 38 agricultural machines categorized in five different groups.

In the Chapter 7 we made some comments on the perspective of applying the developed methodology and also listed some possibilities for extending and deepening the analyzes.

1.3 Products of the thesis

The results of this thesis were organized in four scientific articles. The first one has already been published on IEEE Access journal, the second is in the process of peer review, while the third and fourth are being reviewed by the executing team. The following is a list of elaborated texts, associated with the next chapters.

Louzada, F.; Cuminato, J. A.; Rodriguez, O. M. H.; Tomazella, V. L. D.; Ferreira, P. H.; Ramos, P. L.; Niaki, S. R. A.; Gonzatto-Junior, O. A.; Perissini, I. C.; Alegría, L. F. A.; Colombo, D.; Martins, D. E. A.; Santos, H. F. L.. **A Repairable System Subjected to Hierarchical Competing Risks: Modeling and Applications**. (*This paper was published in IEEE Access journal, 2019*).

Louzada, F.; Cuminato, J. A.; Rodriguez, O. M. H.; Tomazella, V. L. D.; Ferreira, P. H.; Ramos, P. L.; Milani, E. A; Bochio, G.; Perissini, I. C.; Gonzatto Junior, O. A.; Mota, A. L.; Alegría, L. F. A.; Colombo, D.; Perondi, E. A.; Wentz A. V.; Silva Júnior, A. L. Barone, D. A. C.; Santos, H. F. L.; Magalhães M. V. C.. **Improved objective Bayesian estimator for a PLP model hierarchically represented subject to competing risks under minimal repair regime**. (*This paper was under review in PlosOne journal*).

Louzada, F.; Tomazella, V. L. D.; Bochio, G.; Gonzatto Junior, O. A.; Milani, E. A.. **Confidence estimation by adopting a PLP model hierarchically represented with series-parallel connections subject to competing risks under minimal repair regime.**. (*This paper is under internal review by the authors*).

Gonzatto Junior, O. A.; Fernandes, W. R.; Tomazella, V. L. D.; Louzada, F.. **Frailty Model for Multiple Repairable Systems Hierarchically Represented Subject to Competing Risks.**. (*This paper is under internal review by the authors*).

To illustrate the organization of the chapters in terms of the proposed models, look at Figure 3. On the left are all the assumptions associated with the more general model, presented in Chapter 6, described in the central block. On the right are the details with the indication of the associated chapter.

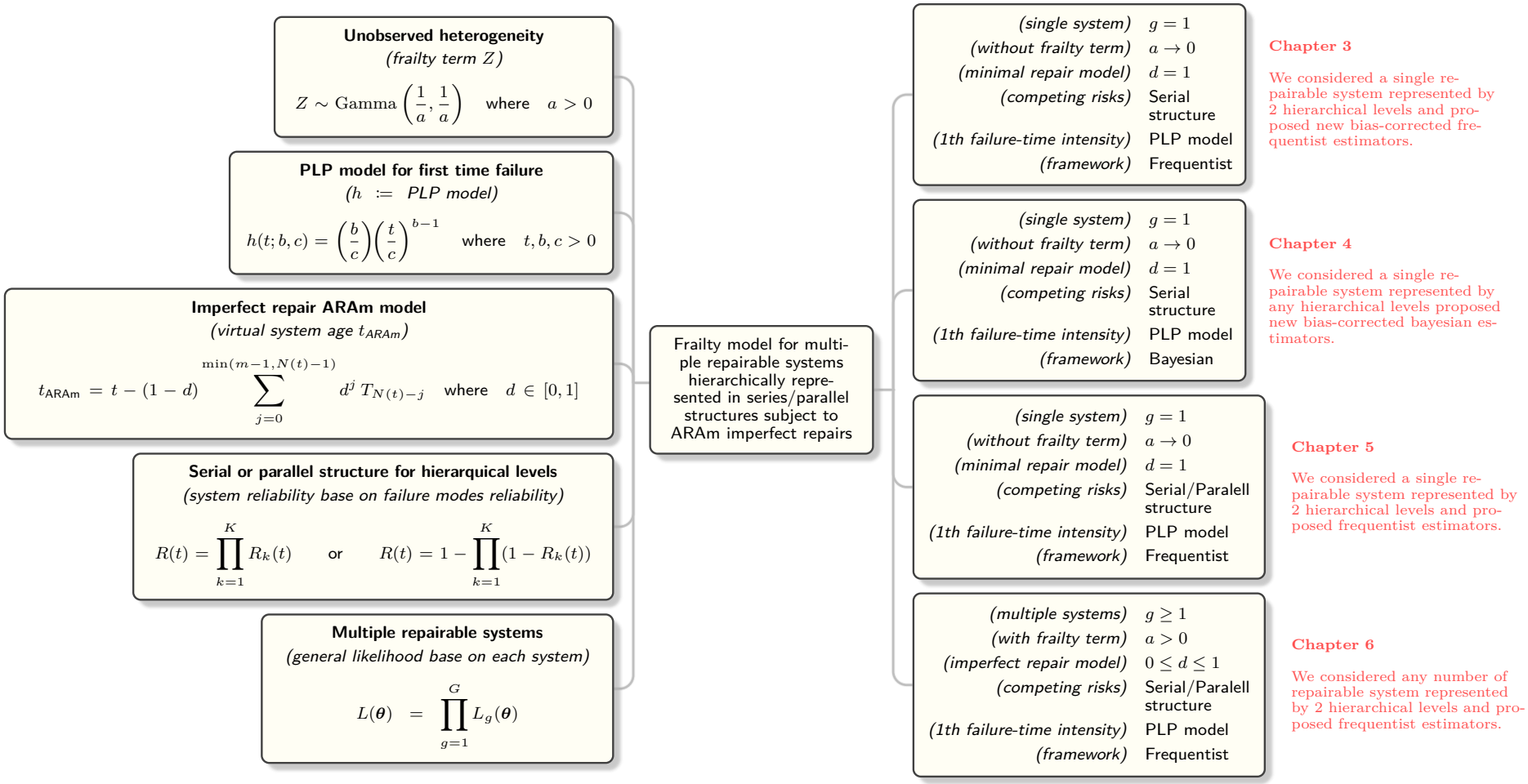


Figure 3 – General framework relating the proposed models and the thesis chapters.

CHAPTER

2

BACKGROUND

In this chapter we try to present all the preliminary concepts necessary for the development of subsequent studies. The definitions, examples and explanations follow very closely the bibliographic reference used for the research.

Contents of Chapter

2.1 Systems

A system can be understood as a collection of sub-systems and/or components organized in a specific way to perform an activity of interest. A system can be classified as to its repairability, in this sense, we say that a system can be non-repairable or repairable, the analysis methodology should take this into account.

According to Ascher and Feingold (1984), a repairable system can be defined as a

system that, after a failure, can return its activity satisfactorily due to a repair action without having to replace the system as a whole. The proper methodological approach to dealing, specifically, with repairable systems avoids a number of misconceptions arising from the use of a methodology that does not take this into account, which is often done in practice. In this sense, the text of Ascher and Feingold (1984) is one of the basic references, which is exclusively dedicated to describing the reliability of repairable systems.

Another background text that addresses the difference between the methodologies used in repairable and non-repairable systems is the text of Rigdon and Basu (2000). In this context, we define a repairable system as a system that can be restored after a failure, provided that its repair is, in some way, compensatory to the user. Examples include automobiles, refrigerators, televisions, etc. Non-repairable systems, on the other hand, are systems whose failure leads to system disposal, such as a light bulb, or even a system whose repair cost is close to the cost of a new equipment, in that sense the discard is more compensatory, as certain cell phone failures.

Systems repairability can be categorized into three repair actions: MR, PR, and IR. Each carries with it a particular understanding of the behavior of the failures observed in a given repairable system.

2.1.1 MR

The MR is the most explored repair action in the literature, its assumption is very common in many models. An MR aims to correct only the component that caused the system failure, leaving it in the same condition as the system was immediately before the failure, a condition known in the literature as "As Bad as Old" (ABAO).

A typical example of a repairable system is a car. Normally, the operating time of cars is expressed in terms of the indicated mileage. Repair actions will not normally result in extra mileage. Repair time is therefore negligible. Many repairs are performed with simple adjustments or single component replacements. The assumption of MR is thus often applicable, therefore, repair actions keep the system in pre-failure condition. According to Kijima (1989), this assumption is plausible for systems consisting of several components, each having its own failure mode.

Under the assumption of an MR action, the appropriate mathematical/statistical methodology, for understanding the behavior of system failures, is the theory of Stochastic Process (SP), particularly NHPP. As said by Rausand and Hoyland (2004), in such a process, the probability of failures is independent of the failure history, only the age of the system. The use of this methodology will be further described in Section 2.2.3.

2.1.2 PR

A PR is characterized by, after repair action, providing the system with a condition known in the literature as "As Good As New" (AGAN). According to Kijima (1989), it is reasonable to assume a PR condition when working with simple systems whose component replacements represent a substantial increase in system reliability, as to we understand that we are in a condition equivalent to that of a new system.

The counting process associated with a PR perspective is a MR model (see 2.2.2). The main reason why RP are considered is because in the modeling of more complicated processes, there are moments when the system returns to an AGAN state at random. The times observed in these renewals are incorporated via RP so that it is possible to distinguish long-term behavior from system failures.

2.1.3 IR

The action of an IR is responsible for providing the system with a condition that is not sufficiently similar to that of a PR, however, it is better than the immediately preceding the failure (MR). In practice, the repair actions are not limited to the extremes of an MR or PR, in most cases the repair reaches an intermediate condition between these two.

According to Kijima, Morimura and Suzuki (1988), in the situation where a system is composed of few vulnerable components, it is more appropriate to consider that a repair action takes the system condition to an intermediate level between MR and PR. In this context, repair actions are directed at maintaining the system as a whole and not directly at the subsystem or component that caused the failure.

Much research has been done around this problem, Brown and Proschan (1983) has considered a model where, at each failure time, a PR occurs with probability p, while a MR occurs with $1-p$, both independent of fault history. The notion of virtual age was introduced by Kijima (1989), which is a positive function of the actual system age and the history of failure times. This model considers a ρ parameter that represents the degree of repair efficiency, which has as its particular case the actions of MR and PR.

Subsequently, Doyen and Gaudoin (2004) proposed two classes of IR models: the model class of "Arithmetic Reduction of Intensity" (ARI) which considers an arbitrary reduction in system failure rate based on a memory of arbitrary degree; and the model class of ARA which, in this case, considers a reduction in system age, also taking into account arbitrary memory degree.

Some of these models will be further discussed in Section 2.2.5.

2.2 Models for counting processes

Since the modeling of a repairable system admits multiple occurrences of the event in a same system studied, it is important that we also know how to stochasticly model the behavior of occurrence counting. In this sense, the Stochastic Process we're interested in is a counting process, and that's what we'll talk about in the next section.

Stochastic Processes's theory formally describes numerous behaviors of practical interest, a theoretically deeper description can be seen, for example, in the work of Ross (1996). For the subject matter here, a general definition of a Stochastic Process in Time is given in the Definition 2.2.1.

Definition 2.2.1: Stochastic Process in Time

A Stochastic Process in Time is a collection of random variables, denoted by $\{X(t), t \in \Theta\}$, where Θ is the set that indexes (temporally) the process.

PS.: When the Θ set is an enumerable set, the $\{X(t), t \in \Theta\}$ process is said to be a Discrete Time Stochastic Process. On the other hand, when the Θ is a continuous set, the $\{X(t), t \in \Theta\}$ process is said to be a Continuous Time Stochastic Process.

In addition, the process of interest, in the present study, aims to understand *counts resulting from a random behavior*, and can be formally represented with the Definition 2.2.2.

Definition 2.2.2: counting process

A Stochastic Process $\{N(t), t \geq 0\}$ is said to be a counting process if it satisfies

1. $N(t) \geq 0$;
2. $N(t)$ is integer;
3. $s < t \Rightarrow N(s) \leq N(t)$;
4. For any $s, t \geq 0$ with $s < t$, the amount $N(t) - N(s)$ represents the number of occurrences of the event of interest, observed in the range $(s, t]$;

PS.: It is also possible to represent counting process $\{N(t), t \geq 0\}$ as the sequence of times until the event $T_1, T_2, \ldots$, or the sequence of times between events $S_1, S_2, \ldots$. Either of these representations carry the same information as the others with respect to counting process. The illustration of such observations can be seen in Figure 4.

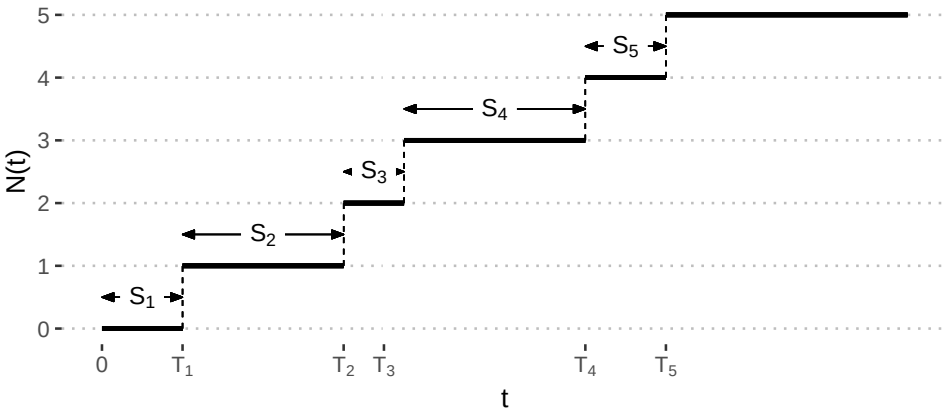


Figure 4 – Representation of the number of events $N(t)$, times to events T_i, and times between events S_i.

In the study of repairable systems, the behavior of the process $\{N(t), t \geq 0\}$ provides numerous indications of the appropriate methodology for addressing the problem. For example, when $N(t)$ is a nonlinear function of time, the assumption that the times between failures are independent and identically distributed is not true.

To better understand the behavior of such processes, some concepts must be kept in mind. The first one is the notion of **independent increments**, whose idea tells us that, if a high number of events were observed within a certain time interval, this will not influence the behavior of future failures. Formally, we consider the Definition 2.2.3.

Definition 2.2.3: Independent Increments

We say that a counting process $\{N(t), t \geq 0\}$ has independent increments if, for any $0 < t_1 < t_2 < \cdots < t_k$, with $k = 2, 3, \ldots$, the quantities $N(t_1) - 0, N(t_2) - N(t_1), \ldots, N(t_k) - N(t_{k-1})$ are independent random variables.

Another important concept is **stationary increments**, which tells us that the distribution of the number of occurrences in a given interval depends only on the length of the interval, not its position with respect to its origin. Formally, this can be set with the Definition 2.2.4.

Definition 2.2.4: Stationary Increments

We say that a counting process $\{N(t), t \geq 0\}$ has stationary increments if, for any t, s with $0 \leq s < t$ and any constant $c > 0$, the random quantities $N(t) - N(s)$ and $N(t + c) - N(s + c)$ are identically distributed.

PS. 1: We call a counting process as *Stationary (or Homogeneous) counting process*, if it can be characterized by Stationary Increments.
PS. 2: We call a counting process as *Nonstationary (or Non-Homogeneous) counting process*, if it is not stationary or does not become stationary after a given time.

It is also important to keep in mind what is a **regular process**. Such a process is

characterized by the impossibility of observing more than one occurrence at the same time instant. Formally, we have the Definition 2.2.5.

Definition 2.2.5: Regular Process

A counting process is said to be Regular (or Ordered) if

$$\Pr(N(t+\tau)-N(t)\geq 2)=o(\tau), \quad \text{for small values of} \quad \tau,$$

and where $o(\tau)$ is any function such that

$$\lim_{\tau\to 0}\frac{o(\tau)}{\tau}=0.$$

Another very important aspect for studying a counting process is the *Process Rate*, that is, the intensity with which events occur. The function that represents this feature is the *Intensity Function*, whose formal definition is in Definition 2.2.6.

Definition 2.2.6: Intensity Function

The intensity function of a counting process at time t is given by

$$\lambda(t)=\lim_{\tau\to 0}\frac{\Pr(N(t+\tau)-N(t)=1\mid \mathcal{H}_{t^-})}{\tau}, \qquad \forall t\geq 0,$$

where $\mathcal{H}_{t^-}$ denotes the history set of previous failure times.
In addition, the expected number of occurrences in the interval $(0,t]$ (or the mean cumulative function).

$$\Lambda(t)=\mathbb{E}(N(t))=\int_0^t \mathbb{E}(\lambda(s))\,\mathrm{d}s,$$

The counting processes commonly addressed in the literature are: Homogeneous Poisson Processes (HPP); Renewal Processes (RP); NHPP and IR Processes. Each will be further described in the following sections. An illustration of the counting processes associated with the respective repair models, described above, can be seen in Figure 5.

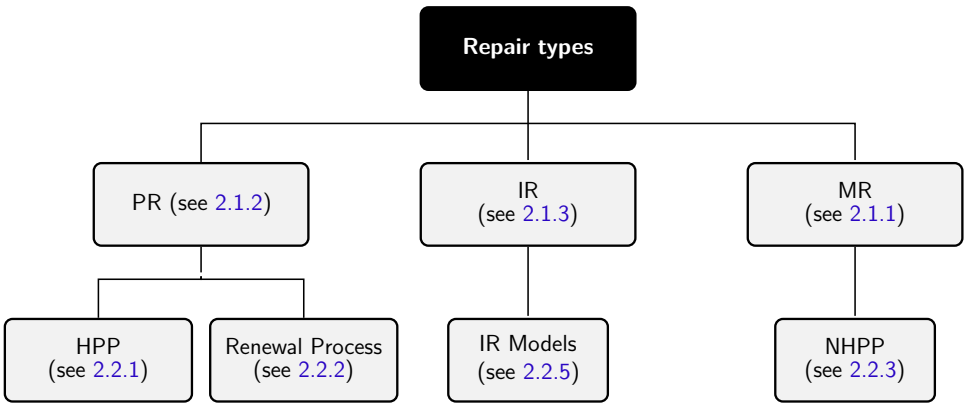


Figure 5 – Repair types and associated count processes.

2.2.1 HPP

There are some equivalent definitions for this counting process, one widely used is the Definition 2.2.7.

Definition 2.2.7: HPP

We say that counting process $\{N(t), t \geq 0\}$ is a HPP with a rate of λ, to $\lambda > 0$ if it meets the following conditions.

1. $N(0) = 0$;
2. The process has stationary (see 2.2.4) and independent increments (see 2.2.3);
3. $\Pr(N(\tau) = 1) = \lambda\,\tau + o(\tau)$;
4. $\Pr(N(\tau) \geq 2) = o(\tau)$

The definition of a HPP results in some immediate consequences, some of which will be listed here, others can be found in the texts of Ross (1996), Thompson (1988), and Ascher and Feingold (1984).

The first of these, is the fact that, in addition to having stationary (see 2.2.4) and independent (see 2.2.3) increments, the HPP is a regular counting process (see 2.2.5). The second consequence, is that HPP ROCOF is constant and independent of time, i.e.

$$\lambda(t) = \lambda, \qquad \text{for all} \qquad t \geq 0.$$

The third, shows us that the count of occurrences in a interval of the form $(\tau, t+\tau]$ has a Poisson distribution, with expected value $\lambda\,t$, that is

$$\Pr(N(t+\tau) - N(\tau) = n) = \frac{(\lambda\,\tau)^n}{n!} e^{-\lambda\,t}, \qquad \text{for all} \qquad t \geq 0, \tau > 0.$$

The fourth consequence shows that the average occurrence count in the interval $(\tau, t+\tau]$ is

$$\Lambda(t+\tau) - \Lambda(\tau) = \mathbb{E}(N(t+\tau) - N(\tau)) = \lambda\,t,$$

particularly, $\mathbb{E}(N(t)) = \lambda\,t$ e $\mathrm{Var}(N(t)) = \lambda\,t$;

As a fifth consequence, the times between occurrences $S_1, S_2, \ldots$, are independent and identically distributed random variables with exponential distribution of mean $1/\lambda$.

Finally, the sixth consequence listed here, tells us that time up to n-th occurrence, $T_n = \sum_{i=1}^{n} S_i$ has Gamma distribution with parameters vector (n, λ), whose density is represented by

$$f_{T_n}(t) = \frac{\lambda}{(n-1)!}(\lambda\,t)^{n-1} e^{-\lambda\,t} \qquad \text{for} \qquad t \geq 0.$$

2.2.2 RP

The Renewal Theory has its origin in the study of strategies for the replacement of technical components, but was later developed as a general theory within Stochastic Processes. This process is used to model equipment renovations or replacements.

In a RP, when a functional component is put into operation at time $t = 0$ and experiences a time failure T_1, it is replaced by a new equivalent component (or restored to a condition AGAN). This process is repeated with each new failure recorded. In this context, we understand that component replacement times are independent and identically distributed, the number of failures (and thus renewals) in the interval $(0,t]$ is denoted by $N(t)$.

An RP is a counting process with times between occurrences $S_1, S_2, \ldots$, considered independent and identically distributed. The observed occurrences are referred to, in this context, as *renewals*. The RP is a generalization of HPP, where the distribution of times between occurrences is exponential with parameter λ.

2.2.3 NHPP

The NHPP is a generalization of HPP in the sense that the failure rate is a function of time, i.e. failure rate may vary according to specific fault behavior criteria.

Definition 2.2.8: NHPP

A counting process $\{N(t), t \geq 0\}$ is a NHPP (or non-stationary) with intensity function $\lambda(t)$ to $t \geq 0$, if

1. $N(0) = 0$.
2. $\{N(t), t \geq 0\}$ has independent increments (see 2.2.3).
3. $\Pr(N(t+\tau) - N(t) \geq 2) = o(\tau)$, i.e., the system does not observe more than one failure at a time.
4. $\Pr(N(t+\tau) - N(t) = 1) = \lambda(t)\tau + o(\tau)$.

It is important to note that NHPP does not require stationary increments. This means that the occurrence of events may be more likely at certain time periods than others, so the times between occurrences are generally neither independent nor identically distributed. As a consequence, methodologies that consider independent and identically distributed data do not apply to an NHPP.

NHPP is usually used to understand possible trends in times between occurrences such as the improvement or deterioration of a system. In this sense, improving systems intuitively have decreasing ROCOF functions, while deteriorating systems have increasing ROCOF functions.

Due to the assumption of independent increments, the number of occurrences within a

specific interval $(t_1, t_2]$ will be independent of occurrences (and times between occurrences) prior to time t_1. When there is an occurrence in time t_1, the ROCOF function conditional on the occurrence history, $\lambda(t|\mathcal{H}_t)$, will be, in the next time interval, $\lambda(t)$ and therefore, regardless of the observed story $\mathcal{H}_{t_1}$ (until the time t_1). The practical implication of this assumption is that the history-dependent ROCOF function is the same immediately before and immediately after the repair performed. The terminology used because of this assumption is called MR; see Ascher and Feingold (1984).

When components (in operation for a long time) are replaced with new components, an NHPP is certainly not an appropriate model. For a NHPP model to be realistically suitable, the commissioned components should behave identically to those as old as them. In this sense, they should be aged outside the system for the same period of time and under identical conditions.

Considering a system composed of numerous components. Suppose a critical component fails, which in turn causes the system to fail. If such a component is immediately replaced by one of the same type (changeover time may be negligible). Since only a small fraction of the system has been replaced, it seems natural to assume that system reliability after repair is essentially the same as immediately before the failure. In other words, the assumption of a minimum repair is a realistic approximation. When an NHPP is used to model a repairable system, the system is treated as a *black box* where there is no knowledge of how the system *looks inside*.

Considering an NHPP with ROCOF $\lambda(t)$, and suppose that the failures occur at times $T_1, T_2, \ldots$. An illustration of $\lambda(t)$, in the context where a deteriorating system is monitored, can be seen in Figure 6.

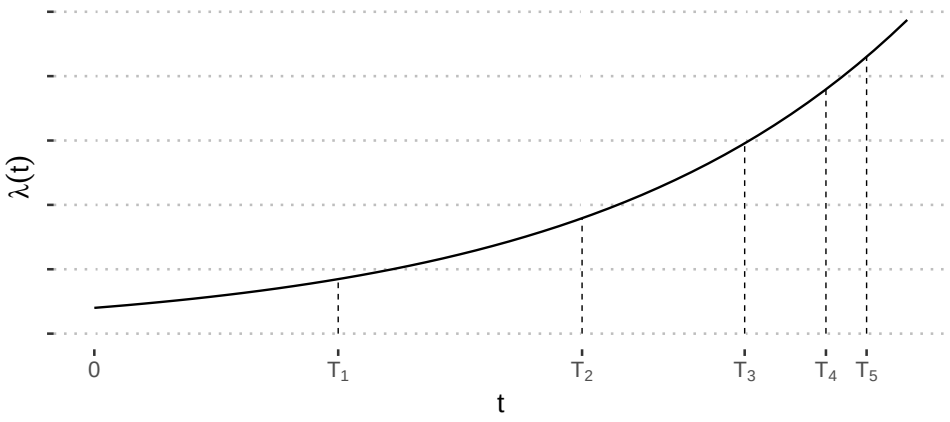


Figure 6 – ROCOF $\lambda(t)$ for the NHPP.

Similarly to what is identified about HPP, some immediate consequences are also observed from the definition of an NHPP (exposed in 2.2.8). Some of them tell us that the

number of occurrences in the interval $(0,t]$ has Poisson distribution, with parameter $\Lambda(t)$, i.e.,

$$\Pr(N(t)=n)=\frac{[\Lambda(t)]^n}{n!}e^{-\Lambda(t)} \qquad \text{para} \qquad n=0,1,2,\ldots,$$

therefore, the mean and variance for the number of occurences in $(0,t]$ is $\mathbb{E}(N(t))=\mathrm{Var}(N(t))=\Lambda(t)$.

The cumulative ROCOF $\Lambda(t)$ is also the mean number of occurrences in the interval $(0,t]$, and is sometimes also called *mean value function* of the process. When n is large, $\Pr(N(t)\leq n)$ can be determined by a normal approximation

$$\Pr(N(t)\leq n)=\Pr\left(\frac{N(t)-\Lambda(t)}{\sqrt{\Lambda(t)}}\leq\frac{n-\Lambda(t)}{\sqrt{\Lambda(t)}}\right)\approx\Phi\left(\frac{n-\Lambda(t)}{\sqrt{\Lambda(t)}}\right).$$

In addition, it follows that the number of occurrences in the interval $(s,t]$, with $s<t$, also has Poisson distribution.

$$\Pr(N(t)-N(s)=n)=\frac{[\Lambda(t)-\Lambda(s)]^n}{n!}e^{-[\Lambda(t)-\lambda(s)]} \qquad \text{for} \qquad n=0,1,2,\ldots$$

and the mean number of occurrences in the range $(s,t]$ is

$$\mathbb{E}(N(t)-N(s))=\Lambda(t)-\Lambda(s)=\int_s^t\lambda(u)\,\mathrm{d}u.$$

The probability of non-occurrences in the interval $(s,t]$, with $s<t$ is

$$\Pr(N(t)-N(s)=0)=\exp\left\{-\int_s^t\lambda(u)\,\mathrm{d}u\right\}.$$

Considering T_n the time to n-th occurrence, for $n=0,1,2,\ldots,$ where $T_0=0$. The distribution of T_n is given by

$$\Pr(T_n>t)=\Pr(N(t)\leq n-1)=\sum_{k=1}^{n-1}\frac{\Lambda(t)^k}{k!}e^{-\Lambda(t)}.$$

When $\Lambda(t)$ is large, this probability can be determined using a normal approximation.

$$\Pr(T_n>t)=\Pr(N(t)\leq n-1)\approx\Phi\left(\frac{n-1-\Lambda(t)}{\sqrt{\Lambda(t)}}\right).$$

2.2.3.1 Time to first occurrence

Assume that T_1 denotes the time of $t=0$ until the first failure. The reliability function of T_1 is given by

$$R_1(t)=\Pr(T_1>t)=\Pr(N(t)=0)=e^{-\Lambda(t)}=\exp\left\{-\int_0^t\lambda(u)\,\mathrm{d}u\right\}.$$

Thus, the failure rate $h(t)$ function of the first time between occurrences is equal to the ROCOF $\lambda(t)$ function of the process. However, there is a different meaning between the two expressions.

This result has a disconcerting implication, since the distribution of the first time between occurrences determines the ROCOF function of the entire process. As stated in Thompson (1981), under the assumption of an NHPP, if we are able to estimate the failure rate function for the first time between occurrences, then we will estimate the lifetime ROCOF function of the phenomenon studied, which is a counterintuitive fact.

2.2.3.2 Time between occurrences

Assume that the process is observed in time t_0 (regardless of whether it is a time of occurrence or any arbitrary point in time). Also assume that $Y(t_0)$ denotes the time until the next occurrence. In this context, the distribution of $Y(t_0)$ is determined by

$$\begin{aligned}\Pr(Y(t_0) > t) &= \Pr(N(t+t_0) - N(t_0) = 0) \\ &= e^{\Lambda(t+t_0)-\Lambda(t_0)} \\ &= \exp\left\{-\int_{t_0}^{t+t_0} \lambda(u)\,\mathrm{d}u\right\} \\ &= \exp\left\{-\int_0^{t} \lambda(u+t_0)\,\mathrm{d}u\right\}.\end{aligned}$$

If t_0 represents the time, T_{n-1}, of the occurrence $n-1$. In this case, $Y(t_0)$ denotes the time between $(n-1)$-th and n-th failure (that is, n-th time between occurrences $S_n = T_n - T_{n-1}$). The failure rate function (in this case, conditional to $T_{n-1} = t_0$) of n-th time between occurrences S_n is

$$h_{t_0}(t) = \lambda(t+t_0) \qquad \text{for} \qquad t \geq 0.$$

The Mean Time Between the $(n-1)$-th and n-th Failure (MTBF_n), at time t_0, is given by

$$\mathrm{MTBF}_n = \mathbb{E}(T_n) = \int_0^{\infty} \Pr(Y_{t_0} > t)\,\mathrm{d}t = \int_0^{\infty} \exp\left\{-\int_0^{t} \lambda(u+t_0)\,\mathrm{d}u\right\}\mathrm{d}t.$$

Example 2.2.9: [Adapted from (RAUSAND; HOYLAND, 2004)]

Given an NHPP with ROCOF

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

for $\beta, \eta > 0$ and $t \geq 0$. The mean number of occurrences in the interval $(0,t)$ is given by

$$\Lambda(t) = \mathbb{E}(N(t)) = \int_0^t \lambda(u)\,\mathrm{d}u = \left(\frac{t}{\eta}\right)^{\beta}.$$

The distribution of time to first failure, T_1, can be determined by the reliability function.

$$R_1(t) = e^{-\Lambda(t)} = e^{-(t/\eta)^{\beta}} \qquad \text{for} \qquad t \geq 0,$$

that is, a Weibull distribution with scale parameter η and form β.
If the process is observed at time t_0, the reliability of the time until next failure $Y(t_0)$ is given by

$$\Pr(Y(t_0) > t) = \exp\left\{-\int_0^t \lambda(u+t_0)\,\mathrm{d}u\right\} = \exp\left\{-\left[\left(\frac{t_0+t}{\eta}\right)^{\beta} - \left(\frac{t_0}{\eta}\right)^{\beta}\right]\right\}.$$

If t_0 denotes the $(n-1)$-th time of failure, the time to next failure $Y(t_0)$, is the time n-th time between occurrences S_n and the failure rate function of S_n is

$$h_{t_0}(t) = \left(\frac{\beta}{\eta}\right)\left(\frac{t+t_0}{\eta}\right)^{\beta-1},$$

This is a conditional failure rate, since the occurrence $n-1$ occurred at time $t_0 = T_{n-1}$.
The mean time between failures n and $n-1$ is given by

$$\mathrm{MTBF}_n = \int_0^{\infty} \exp\left\{-\left[\left(\frac{t_0+t}{\eta}\right)^{\beta} - \left(\frac{t_0}{\eta}\right)^{\beta}\right]\right\}\mathrm{d}t$$

2.2.3.3 Parametric Likelihood Function in an NHPP

In either of these contexts, choosing a convenient parametric form for $\lambda(t)$ allows us to obtain a sufficiently flexible model to the number of failures of a repairable system in the context of MR, i.e., when only a small proportion of the system's constituent parts are replaced at repair.

2.2.3.3.1 Truncation by time

Consider the situation in which the failures of a given repairable system are monitored until a predetermined time, T. In this context, the representation of the likelihood function for

the parameter vector, $\boldsymbol{\theta}$, considering the observed time failures $\boldsymbol{t} = (t_0, t_1, \ldots, t_n)$, where $t_n < T$ and $t_0 = 0$, should consider the random behavior of time as well as the number of failures $N(T)$. Like this

$$L(\boldsymbol{\theta};\boldsymbol{t}) = f(t_1 \mid t_0) f(t_2 \mid t_1) \cdots f(t_n \mid t_1, t_2, \ldots, t_{n-1}) \Pr(N(T) = n \mid t_1, t_2, \ldots, t_n).$$

To approximate a term of the form $f(t_i \mid t_0, \ldots, t_{i-1})$, where $i = 1, \ldots, n$, we must keep in mind that

$$\begin{aligned}
f(t_i \mid t_0, \ldots, t_{i-1}) &= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\, 1 - R(t_i \mid t_0, \ldots, t_{i-1}) \,\Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\, 1 - \Pr(T_i > t_i \mid t_0, \ldots, t_{i-1}) \,\Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\, 1 - \Pr(T_i > t_i \mid t_{i-1}) \,\Big] \qquad \left(\begin{array}{c} \text{Independent} \\ \text{Increments} \end{array} \right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\, 1 - \Pr(N(t_i) - N(t_{i-1}) = 0) \,\Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\, 1 - e^{-(\Lambda(t_i) - \Lambda(t_{i-1}))} \,\Big] \\
&= \lambda(t_i)\, e^{-(\Lambda(t_i) - \Lambda(t_{i-1}))}.
\end{aligned}$$

Similarly, for $\Pr(N(t) = n \mid t_1, t_2, \ldots, t_n)$, we have

$$\Pr(N(T) = n \mid t_n) = \Pr(N(T) - N(t_n) = 0) = e^{-(\Lambda(T) - \Lambda(t_n))}.$$

Thus, the likelihood function has the form

$$L(\boldsymbol{\theta};\boldsymbol{t}) = \underbrace{\lambda(t_1)\, e^{-(\cancel{\Lambda(t_1)} - \cancel{\Lambda(t_0)})}}_{f(t_1|t_0)} \overbrace{\underbrace{\lambda(t_2)\, e^{-(\cancel{\Lambda(t_2)} - \cancel{\Lambda(t_1)})}}_{f(t_2|t_0,t_1)}}^{0} \cdots \underbrace{\lambda(t_n)\, e^{-(\cancel{\Lambda(t_n)} - \cancel{\Lambda(t_{n-1})})}}_{f(t_n|t_0,\ldots,t_{n-1})} \underbrace{e^{-(\Lambda(T) - \cancel{\Lambda(t_n)})}}_{\Pr(N(T)|t_0,\ldots,t_n)}$$

$$= \prod_{i=1}^{n} \lambda(t_i) \times e^{-\Lambda(T)}$$

and the log likelihood function is given by

$$\ell(\boldsymbol{\theta};\boldsymbol{t}) = \sum_{i=1}^{n} \log \lambda(t_i) - \Lambda(T).$$

2.2.3.3.2 Failure truncation

In this context, the system is supposed to be watched until the n-th failure occurs. The previous expressions apply normally, replacing the time T with t_n. That is, the likelihood function is given by

$$L(\boldsymbol{\theta};\boldsymbol{t}) = \prod_{i=1}^{n} \lambda(t_i) \times e^{-\Lambda(t_n)},$$

and the log-likelihood

$$\ell(\boldsymbol{\theta};\boldsymbol{t}) = \sum_{i=1}^{n} \log \lambda(t_i) - \Lambda(t_n).$$

2.2.3.4 Parametric Models for an NHPP

From a parametric perspective, it is possible to define convenient features to describe the ROCOF function of an NHPP. In this context, three models will be briefly presented: the Power Law Model, the Linear Model, and the Log-linear Model. These templates can be written in general.

$$\lambda(t) = \gamma g(t;\nu),$$

where γ is a common multiplier, and $g(t;\nu)$ determines the form of the ROCOF function, $\lambda(t)$. All models can be parameterized in a number of ways.

2.2.3.4.1 Power Law Model

In this model, the NHPP ROCOF function is defined as

$$\lambda(t) = \left(\frac{\beta}{\eta}\right)\left(\frac{t}{\eta}\right)^{\beta-1} \qquad \text{for} \qquad t,\eta,\beta > 0.$$

This NHPP is sometimes referred to as the "Weibull process". This is because its ROCOF function has the same analytical expression as the failure rate function of the Weibull distribution. In addition, the first instance T_1 of this process has Weibull distribution with shape parameter β and scale η. However, it is important to avoid mentioning a "Weibull process" since, as stated by Ascher and Feingold (1984) this terminology gives the mistaken impression that the Weibull distribution can be used to model the behavior of times between occurrences in a repairable system.

A repairable system modeled by an NHPP with power law intensity function can be interpreted as an enhancing system if $0 < \beta < 1$, and deteriorating if $\beta > 1$. In the situation where $\beta = 1$ the model particularizes in an HPP.

2.2.3.4.2 Linear Model

In the linear model, the NHPP ROCOF function is defined by

$$\lambda(t) = (\beta t + 1)\alpha \qquad \text{for} \qquad \alpha > 0 \text{ and } t \geq 0.$$

A repairable system modeled by the linear model, as show in Vesely (1991) and Atwood (1992), is interpreted as deteriorating if $\beta > 0$, and improving when $\beta < 0$. When $\beta < 0$, then $\lambda(t)$ will become less than zero at some point. In this sense, the model should be used only at time intervals where $\lambda(t) > 0$.

2.2.3.4.3 Log-linear Model

The log-linear model was proposed by Cox and Lewis (1966), in this model the NHPP ROCOF function is defined by

$$\lambda(t) = e^{\alpha+\beta t} \qquad \text{for} \qquad -\infty < \alpha, \beta < \infty \text{ and } t \geq 0.$$

A repairable system modeled by a log-linear model is improving if $\beta < 0$, and deteriorating if $\beta > 0$. When $\beta = 0$ the log-linear model becomes particularized in an HPP with $\lambda(t) = e^{\alpha}$.

2.2.4 Statistical trend graphics

Some graphs can help us to see if there is a failure trend behavior. Guthrie (2020) highlights some of them:

2.2.4.1 Accumulated Failures vs. Age of the System

This graph is constructed in the form of a increasing in time step function, where each step is a new failure. This graph expresses three main behaviors (Figure 7):

1. **Behavior close to linear**: In this case we are not able to perceive any specific trend;
2. **Behavior close to logarithmic**: The concavity downwards exposes that the times between failure are becoming more and more *long*, indicating an improvement in the system;
3. **Behavior close to exponential**: The concavity upwards exposes that the times between failure are becoming more and more *short*, indicating a degradation in the system.

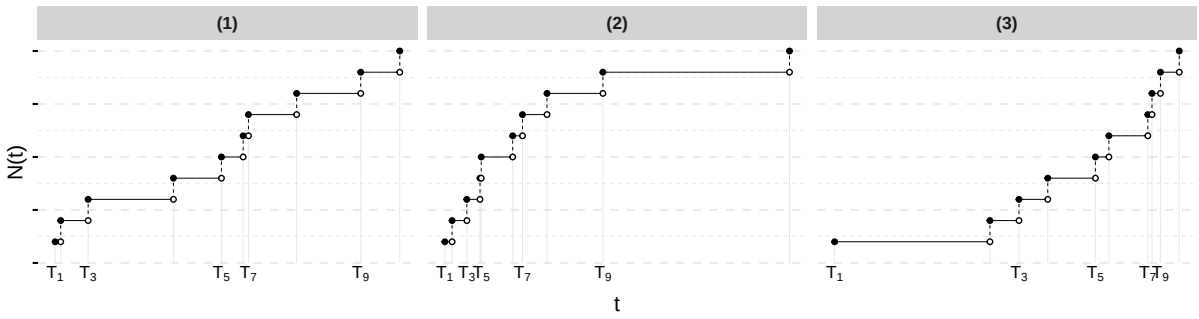


Figure 7 – Trend graphics. Accumulated Failures vs. Age of the System

2.2.4.2 Interarrival times vs. Number of failures

This graph is constructed in the form of a scatter or line plot. This graph also expresses three main behaviors (Figure 8):

1. **Constant behavior**: In this case we are not able to perceive any specific trend;

2. **Increasing behavior**: The times between failure are becoming more and more *long*, indicating an improvement in the system;

3. **Decreasing behavior**: The times between failure are becoming more and more *short*, indicating a degradation in the system.

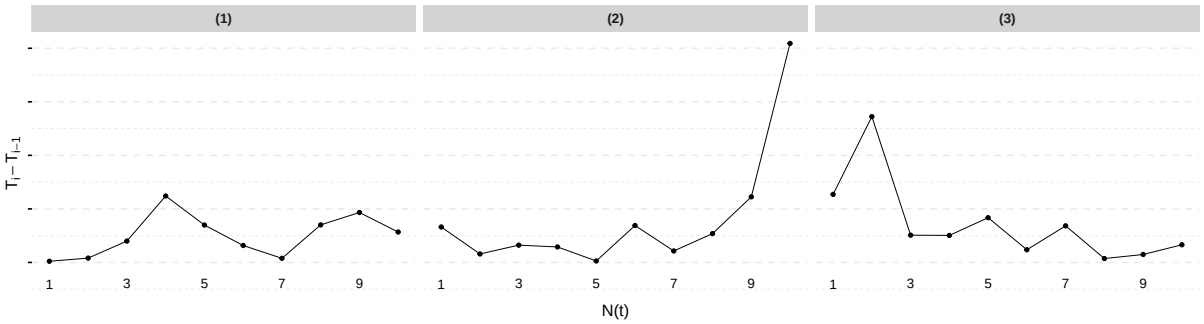


Figure 8 – Trend graphics. Interarrival times vs. Number of failures

2.2.4.3 Reciprocals of the Interarrival times vs. Number of failures

This graph is also constructed in the form of a scatter or line plot and represents an estimate of the failure rate based only on the waiting time since the last failure. It also expresses three main behaviors (Figure 9):

1. **Constant behavior**: We are not able to perceive any specific trend;

2. **Decreasing behavior**: Indicates a improvement system.

3. **Increasing behavior**: Indicates an degradation in the system;

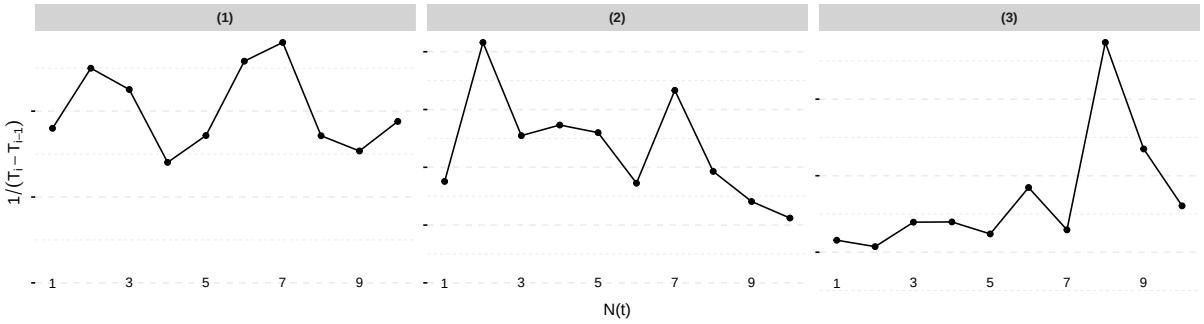


Figure 9 – Trend graphics. Reciprocals of the Interarrival times vs. Number of failures

2.2.4.4 Duane plot

A graph that is also widely used in reliability analysis is the Duane graph. It was proposed by Duane (1964) and consists of plotting the accumulated MTBF versus failure times on the log-scale. If the observed result has a behavior close to linear, it is an indication that the intensity function with the power law form can be applied for the modeling of the failure process.

In practice to execute Duane's initial proposal, if the ith failure occurred at time t_i, we plot t_i/i vs. t_i, on the log-scale. This is justified, therefore, if the process $\{N(t), t > 0\}$ is governed by a power law intensity, then the expected value of $N(t)$ takes the form $\mathbb{E}[N(t)] = (t/\mu)^{\beta}$, thus $t/\mathbb{E}[N(t)] = \mu^{\beta} t^{1-\beta}$. So, in the log-scale, we will obtain a straight line with slope equal to $1-\beta$ and intercept equal to $\beta \log(\mu)$. In summary, a straight line on a duane chart is equivalent to an NHPP Power Law Model.

Note that the same idea can be applied by plotting the expected cumulative number of failures vs. time on the log scale.

2.2.5 IR Processes

When we use a renewal process, the repair action is considered perfect, because when the repair action is complete, we understand that the system is "as good as new". On the other hand, when using an NHPP, we assume that the repair action is minimal, which means that the reliability of the system immediately after the repair is equal to immediately before the failure, which means that the system is "as bad as old". Such processes are extreme cases of repair, and it is reasonable to think that there is a whole gradient of possibilities between these two extremes. In this sense, some models suggest a situation of *normal repair* or even IR. This repair action is between MR and a renewal.

IR models can be categorized into two main groups: the first one concentrates the effect of repair actions on reducing failure rates, i.e., direct influence on ROCOF function; The second of them considers the idea of virtual age, where the age of the system is, somehow, updated by the effect of the repair action. Further details of these models can be found in the texts of Pham and Wang (1996) and Hokstad (1997), for example.

2.2.5.1 ARI Model

This type of modeling assumes that each repair action results in a reduction in the ROCOF conditional function, $\lambda(t|\mathcal{H}_t)$, where $\mathcal{H}_t$ represents system history up to, but not including, the time t. The degree of reduction can be a fixed amount, a percentage of the current failure rate value, both proposed by Chan and Shaw (1993), or even a function of the process history.

Given the failure times $\boldsymbol{t} = (t_1, \ldots, t_n)$ and considering t_{i-} and t_{i+} the times immediately before and after t_i with $i = 1, \ldots, n$, respectively. The models proposed by Chan and Shaw (1993) can be expressed by the conditional ROCOF function as follows.

$$\lambda(t_{i+}|\mathcal{H}_{t_{i+}}) = \begin{cases} \lambda(t_{i-}|\mathcal{H}_{t_{i-}}) - \delta & \text{for a fixed reduction } \delta, \\ \lambda(t_{i-}|\mathcal{H}_{t_{i-}})(1-\rho) & \text{for a proportional reduction } 0 \leq \rho \leq 1. \end{cases}$$

This model expresses that, between two failures, the conditional ROCOF function is assumed to be parallel to the ROCOF $\lambda_1(t)$ function.

The ρ parameter represents the efficiency of the repair action, note that when $\rho = 0$, there is MR (that is, the Chan and Shaw proportional reduction model is particularized in a NHPP). On the other hand, when $\rho = 1$, the repair action will make the conditional ROCOF function equal to zero, which does not represent a renewal process, since times between occurrences will not be identically distributed except for in which case $\lambda_1(t)$ is a linear function. The illustration of a possible behavior of $\lambda(t|\mathcal{H}_t)$ can be seen in Figure 10.

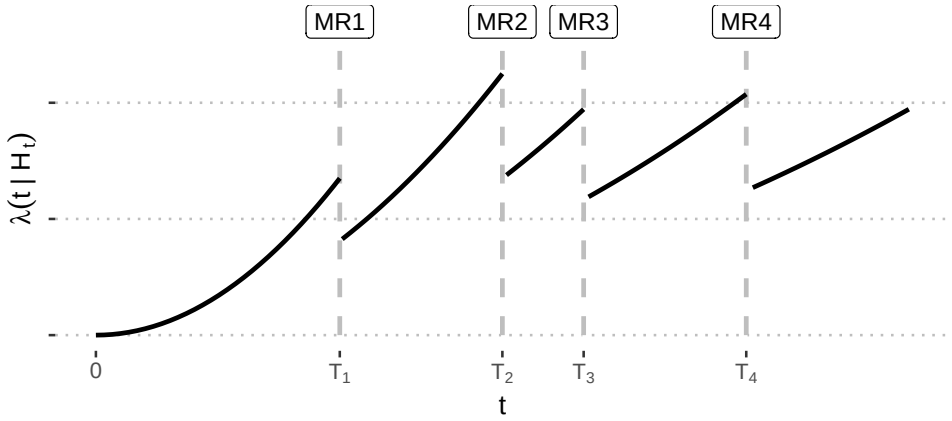


Figure 10 – Conditional ROCOF for Chan and Shaw's proportional reduction model for time failures T_1, T_2, T_3, T_4 and $\rho = \rho_0$.

In another moment, Doyen and Gaudoin (2004) generalized the Chan and Shaw's model, proposing a set of models whose proportionality factor, ρ, depend on the process history, $\mathcal{H}_t$. In this context, the conditional ROCOF function is given by

$$\lambda(t_{i+}|\mathcal{H}_{t_{i+}}) = \lambda(t_{i-}|\mathcal{H}_{t_{i-}}) - \varphi(i, t_1, t_2, \ldots, t_i),$$

where $\varphi(i, t_1, t_2, \ldots, t_i)$ denotes the reduction factor over the conditional ROCOF function due to the repair action. Between two failures, it is assumed that the conditional ROCOF function is parallel to the initial ROCOF function $\lambda_1(t)$. These assumptions lead to the conclusion that

$$\lambda(t|\mathcal{H}_{t_{i+}}) = \lambda_1(t) - \sum_{i=1}^{N(t)} \varphi(i, t_1, t_2, \ldots, t_i).$$

If a proportional reduction is assumed after each repair action (as proposed in the Chan and Shaw's model), the conditional ROCOF function in the interval $(0,t_1)$ will be represented by $\lambda(t|\mathcal{H}_t) = \lambda_1(t)$ and, in the interval $[t_1,t_2)$, will be $\lambda(t|\mathcal{H}\mathcal{H}) = \lambda_1(t) - (1-\rho)\lambda_1(t_1)$, if the process is continued, the resulting expression defines the model named by Doyen and Gaudoin (2004) as ARI with *infinite memory* (ARI_∞), whose conditional ROCOF function has the form

$$\lambda(t|\mathcal{H}_t) = \lambda_1(t) - (1-\rho)\sum_{i=0}^{N(t)} \rho^i \lambda_1(t_{N(t)-i}).$$

Another possible approach assumes that the repair action can only reduce a proportion of the accumulated wear since the previous repair action. In this case, the model is mathematically formulated by

$$\lambda(t_{i+}|\mathcal{H}_{t_{i+}}) = \lambda(t_{i-}|\mathcal{H}_{t_{i-}}) - (1-\rho)\,[\lambda(t_{i-}|\mathcal{H}_{t_{i-}}) - \lambda(t_{(i-1)+}|\mathcal{H}_{t_{(i-1)+}})],$$

so, the resulting conditional ROCOF function of this model is

$$\lambda(t|\mathcal{H}_t) = \lambda_1(t) - (1-\rho)\lambda_1(t_{N(t)}), \tag{2.1}$$

which defines the model, called by Doyen and Gaudoin (2004), as the ARI with *memory one* (ARI_1). Thus, if $\rho = 1$, the system is "as bad as old" after the repair action, which characterizes an NHPP as a particular case. On the other hand, if $\rho = 0$, the conditional ROCOF function is readjusted by the repair action, but the process is not a renewal process since the time between occurrences is not identically distributed. An illustration of the intensity function behavior under this approach can be seen in Figure 11.

The ARI_1 model has a $\lambda_{\min}(t)$ deterministic function that is always smaller than the conditional ROCOF function so that there is no null probability that the ROCOF function will be excessively close to $\lambda_{\min}(t)$. This intensity is the minimum wear intensity, or even a lower upper limit for the conditional ROCOF function, and is given by

$$\lambda_{\min}(t) = \rho\,\lambda_1(t).$$

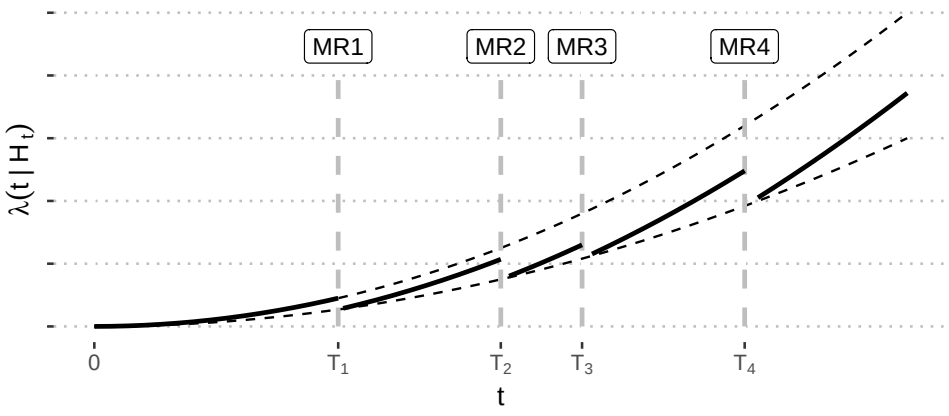


Figure 11 – Conditional ROCOF for ARI_1 model for time failures T_1, T_2, T_3, T_4 and $\lambda = \lambda_0$. In upper dashed line, the function $\lambda_1(t)$ and in lower dashed line, the function $\lambda_{\min}(t)$.

The ARI_∞ and ARI_1 models are two extreme cases, in this same context, a model that considers the last m failure moments can be defined by giving an ARI with *memory* m (ARI_m) model, whose conditional ROCOF function is given by

$$\lambda(t|\mathcal{H}_t) = \lambda_1(t) - (1-\rho)\sum_{i=0}^{\min(m-1,N(t))} \rho^i \lambda(t_{N(t)-i}|\mathcal{H}_{t_{N(t)-i}}).$$

The minimum wear intensity in this context is given by

$$\lambda_{\min}(t) = \rho^m \lambda_1(t).$$

In any model of type ARI, the ρ parameter denotes an efficiency index on the repair action.

- $0 < \rho < 1$: The repair action is efficient;
- $\rho = 1$: The repair action is optimal. The conditional ROCOF function is set to zero;
- $\rho = 0$: The repair action has no effect on system wear. In other words, the system after the repair action is brought to the condition "as bad as old";
- $\rho < 0$: The repair action is harmful to the system, it may be responsible for causing new problems.

2.2.5.1.1 Parametric Likelihood Function in ARI Model

Determination of the likelihood function is obtained analogously to the Section 2.2.3.3, however, the new definition of the intensity function should be considered.

- *Truncation by time*

  In this situation, the failures are monitored until the time T. Thus, the representation of the likelihood function for the parameter vector, $\boldsymbol{\theta}$, considering the observed time failures $\boldsymbol{t} = (t_0, t_1, \ldots, t_n)$, where $t_n < T$ and $t_0 = 0$, must be consider the random behavior of time failures as well as the number $N(T)$. Like this

$$L(\boldsymbol{\theta};\boldsymbol{t}) = f(t_1 \mid t_0) f(t_2 \mid t_1) \cdots f(t_n \mid t_1, t_2, \ldots, t_{n-1}, n) \Pr(N(T) = n).$$

To approximate a term of the form $f(t_i \mid t_0,\ldots,t_{i-1})$, where $i=1,\ldots,n$, we must keep in mind that

$$\begin{aligned}
f(t_i \mid t_0,\ldots,t_{i-1}) &= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1-R(t_i \mid t_0,\ldots,t_{i-1})\ \Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1-\Pr(T_i > t_i \mid t_0,\ldots,t_{i-1})\ \Big] \\
&\overset{*}{=} \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1-\Pr(T_i > t_i \mid t_{i-1})\ \Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1-\Pr(N(t_i)-N(t_{i-1})=0)\ \Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1-\exp\Big\{-\Big(\Lambda(t_i)-\Lambda(t_{i-1})\Big)\Big\}\ \Big] \\
&\overset{**}{=} \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1-\exp\Big\{-\Big(\Lambda_1(t_i)-\Lambda_1(t_{i-1})-(1-\rho)(t_i-t_{i-1})\lambda_1(t_{i-1})\Big)\Big\}\ \Big] \\
&= \Big[\ \lambda_1(t_i)-(1-\rho)\,\lambda_1(t_{i-1})\ \Big]\times \\
&\quad \times\exp\Big\{-\Big(\Lambda_1(t_i)-\Lambda_1(t_{i-1})-(1-\rho)(t_i-t_{i-1})\lambda_1(t_{i-1})\Big)\Big\},
\end{aligned}$$

where the step marked with $*$ uses the assumption of independent increments, and the step highlighted with $**$ uses the definition of the intensity function of the ARI_1 model, show in (2.1).

Similarly, for the term $\Pr(N(t)=n \mid t_1,t_2,\ldots,t_n)$, we have

$$\begin{aligned}
\Pr(N(T)=n \mid t_n) &= \Pr(N(T)-N(t_n)=0) \\
&= \exp\Big\{-\Big(\Lambda(T)-\Lambda(t_n)\Big)\Big\} \\
&= \exp\Big\{-\Big(\Lambda_1(T)-\Lambda_1(t_n)-(1-\rho)(T-t_n)\lambda_1(t_n)\Big)\Big\}.
\end{aligned}$$

Thus, the likelihood function considering the intensity function of a ARI_1 model has the form

$$\begin{aligned}
L(\boldsymbol{\theta};\boldsymbol{t}) = \prod_{i=1}^{n}\Big[\ \lambda_1(t_i)-(1-\rho)\,\lambda_1(t_{i-1})\ \Big]\times \\
\times\exp\Big\{-\Big(\Lambda_1(t_i)-\Lambda_1(t_{i-1})-(1-\rho)(t_i-t_{i-1})\lambda_1(t_{i-1})\Big)\Big\}\times \\
\times\exp\Big\{-\Big(\Lambda_1(T)-\Lambda_1(t_n)-(1-\rho)(T-t_n)\lambda_1(t_n)\Big)\Big\},
\end{aligned}$$

and the log likelihood function is given by

$$\begin{aligned}
\ell(\boldsymbol{\theta};\boldsymbol{t}) = \sum_{i=1}^{n}\Big[\ \log\Big(\lambda_1(t_i)-(1-\rho)\,\lambda_1(t_{i-1})\Big)+ \\
-\Lambda_1(t_i)+\Lambda_1(t_{i-1})+(1-\rho)(t_i-t_{i-1})\lambda_1(t_{i-1})\ \Big]+ \\
-\Lambda_1(T)+\Lambda_1(t_n)+(1-\rho)(T-t_n)\lambda_1(t_n),
\end{aligned}$$

- *Failure truncation*

  Considering a failure truncation approach, the likelihood function representation for the parameter vector, $\boldsymbol{\theta}$, given the observed failure times $\boldsymbol{t} = (t_0, t_1, \ldots, t_n)$, with $t_0 = 0$, should consider the random pattern of failure times, and the number of failures as a fixed amount. Like this,

$$L(\boldsymbol{\theta}; \boldsymbol{t}) = f(t_1 \mid t_0) f(t_2 \mid t_1) \cdots f(t_n \mid t_1, t_2, \ldots, t_{n-1}, n).$$

  The terms $f(t_i \mid t_0, \ldots, t_{i-1})$, with $i = 1, \ldots, n$, are obtained analogously to the situation built in the context of time truncation. Thus, the likelihood function, considering the intensity function of the ARI_1 model, in a failure truncation context, is given by

$$L(\boldsymbol{\theta}; \boldsymbol{t}) = \prod_{i=1}^{n} \Big[\lambda_1(t_i) - (1-\rho)\lambda_1(t_{i-1}) \Big] \times \\ \times \exp\Big\{ - \Big(\Lambda_1(t_i) - \Lambda_1(t_{i-1}) - (1-\rho)(t_i - t_{i-1})\lambda_1(t_{i-1}) \Big) \Big\},$$

  and the log likelihood function is given by

$$\ell(\boldsymbol{\theta}; \boldsymbol{t}) = \sum_{i=1}^{n} \log \Big[\lambda_1(t_i) - (1-\rho)\lambda_1(t_{i-1}) \Big] + \\ - \Lambda_1(t_i) + \Lambda_1(t_{i-1}) - (1-\rho)(t_i - t_{i-1})\lambda_1(t_{i-1}),$$

2.2.5.2 ARA Model

Another approach to addressing the effect of an IR action is to reduce the system's age by a time proportional to the operating time until the immediately preceding repair action. In this context, the system *age* is determined based on the concept of *virtual age*.

The idea behind this model, established by Malik (1979), considers that the conditional ROCOF function right after the first repair action is given by

$$\lambda(t_{1+} | \mathcal{H}_{t_{1+}}) = \lambda_1(t_1 - \nu),$$

where $t_1 - \nu$ denotes the new age (the virtual age) of the system. At the time of the second failure, the conditional ROCOF function is given by $\lambda(t_{2+} | \mathcal{H}_{t_{2+}}) = \lambda_1(t_2 - 2\nu)$. At a given time t, the function is given by $\lambda(t | \mathcal{H}_t) = \lambda_1(t - N(t)\nu)$. It is also possible to consider that ν is a function of the process history, so

$$\lambda(t | \mathcal{H}_t) = \lambda_1 \left(t - \sum_{i=1}^{N(t)} \nu(i, t_1, \ldots, t_i) \right),$$

Some time later, Doyen and Gaudoin (2004) presented the idea of a proportional reduction in age of the system, based on age immediately before the repair action. Thus, we

define the ARA with *infinite memory* model (ARA_∞), whose conditional ROCOF function is given by

$$\lambda(t|\mathcal{H}_t) = \lambda_1\left(t - (1-\rho)\sum_{i=1}^{N(t)} \rho^i t_{N(t)-i}\right).$$

In this model, when $\rho = 1$, one has that $\lambda(t|\mathcal{H}_t) = \lambda_1(t)$, that is, the model particularizes in an NHPP. On the other hand, when $\rho = 0$, one has to $\lambda(t|\mathcal{H}_t) = \lambda_1(t - t_{N(t)})$, which represents a repair action that leads to a condition of "as good as new", that is, a PR . Therefore, the NHPP and the renewal process are particular cases of the ARA_∞ model.

The model introduced by Malik (1979) assumes that the repair action performed at time t_i reduces the last operating time from $t_i - t_{i-1}$ to a proportional amount of that time, that is, $(1-\rho)(t_i - t_{i-1})$, where $0 \leq \rho \leq 1$. The conditional ROCOF function associated with this approach is given by

$$\lambda(t|\mathcal{H}_t) = \lambda_1(t - (1-\rho)t_{N(t)}), \tag{2.2}$$

and the minimum wear intensity is equal to $\lambda_1(\rho t)$. This model is called ARA with *memory one* model (ARA_1).

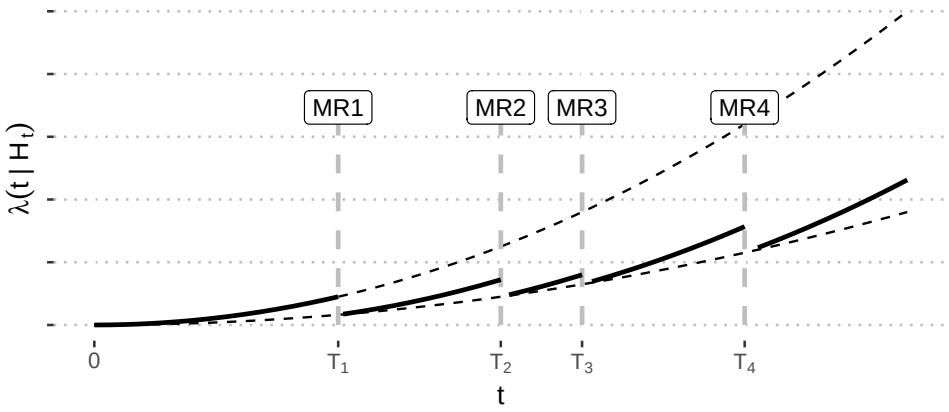


Figure 12 – Conditional ROCOF for ARA_1 model for time failures T_1, T_2, T_3, T_4 and $\rho = \rho_0$. In upper dashed line, the function $\lambda_1(t)$ and in lower dashed line, the function $\lambda_{\min}(t)$.

Similarly to the model ARA_m, it is also possible to define the ARA with *memory* m (ARA_m), in this context, the conditional ROCOF function is given by

$$\lambda(t|\mathcal{H}_t) = \lambda_1\left(t - (1-\rho)\sum_{i=1}^{\min(m-1,N(t))} \rho^i t_{N(t)-i}\right),$$

and the minimum wear intensity is of the form

$$\lambda_{\min}(t) = \lambda_1(\rho^m t).$$

2.2.5.2.1 Parametric Likelihood Function in ARA Model

Determination of the likelihood function is obtained analogously to the Section 2.2.3.3, however, the new definition of the intensity function should be considered.

- *Truncation by time*

  Consider the situation in which the failures of a given repairable system are monitored until a predetermined time T. In this context, the representation of the likelihood function for the parameter vector, $\boldsymbol{\theta}$, considering the observed time failures $\boldsymbol{t} = (t_0, t_1, \ldots, t_n)$, where $t_n < T$ and $t_0 = 0$, should consider the random behavior of time as well as the number of failures $N(T)$. Like this

$$L(\boldsymbol{\theta}; \boldsymbol{t}) = f(t_1 \mid t_0) f(t_2 \mid t_1) \cdots f(t_n \mid t_1, t_2, \ldots, t_{n-1}, n) \Pr(N(T) = n).$$

  To approximate a term of the form $f(t_i \mid t_0, \ldots, t_{i-1})$, where $i = 1, \ldots, n$, we must keep in mind that

$$\begin{aligned}
f(t_i \mid t_0, \ldots, t_{i-1}) &= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1 - R(t_i \mid t_0, \ldots, t_{i-1})\ \Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1 - \Pr(T_i > t_i \mid t_0, \ldots, t_{i-1})\ \Big] \\
&\overset{*}{=} \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1 - \Pr(T_i > t_i \mid t_{i-1})\ \Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1 - \Pr(N(t_i) - N(t_{i-1}) = 0)\ \Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1 - \exp\Big\{ - \Big(\Lambda(t_i) - \Lambda(t_{i-1})\Big)\Big\}\ \Big] \\
&\overset{**}{=} \frac{\mathrm{d}}{\mathrm{d}t_i}\Big[\ 1 - \exp\Big\{ - \Big(\Lambda_1(t_i - (1-\rho)\, t_{i-1}) - \Lambda_1(\rho\, t_{i-1})\Big)\Big\}\ \Big] \\
&= \lambda_1(t_i - (1-\rho)\, t_{i-1}) \exp\Big\{ - \Big(\Lambda_1(t_i - (1-\rho)\, t_{i-1}) - \Lambda_1(\rho\, t_{i-1})\Big)\Big\},
\end{aligned}$$

  where the step marked with $*$ uses the assumption of independent increments, and the step highlighted with $**$ uses the definition of the intensity function of the ARA_1 model, show in (2.2).

  Similarly, for the term $\Pr(N(t) = n \mid t_1, t_2, \ldots, t_n)$, we have

$$\begin{aligned}
\Pr(N(T) = n \mid t_n) &= \Pr(N(T) - N(t_n) = 0) \\
&= \exp\Big\{ - \Big(\Lambda(T) - \Lambda(t_n)\Big)\Big\} \\
&= \exp\Big\{ - \Big(\Lambda_1(T - (1-\rho)\, t_n) - \Lambda_1(\rho\, t_n)\Big)\Big\}.
\end{aligned}$$

Thus, the likelihood function considering the intensity function of a ARA_1 model has the form

$$L(\boldsymbol{\theta};\boldsymbol{t}) = \prod_{i=1}^{n} \lambda_1(t_i - (1-\rho)t_{i-1}) \exp\Big\{ - \Big(\Lambda_1(t_i - (1-\rho)t_{i-1}) - \Lambda_1(\rho\, t_{i-1})\Big)\Big\} \times$$
$$\times \exp\Big\{ - \Big(\Lambda_1(T - (1-\rho)t_n) - \Lambda_1(\rho\, t_n)\Big)\Big\},$$

and the log likelihood function is given by

$$\ell(\boldsymbol{\theta};\boldsymbol{t}) = \sum_{i=1}^{n} \Big[\log\Big(\lambda_1(t_i - (1-\rho)t_{i-1})\Big) - \Lambda_1(t_i - (1-\rho)t_{i-1}) + \Lambda_1(\rho\, t_{i-1}) \Big] +$$
$$- \Lambda_1(T - (1-\rho)t_n) + \Lambda_1(\rho\, t_n),$$

- *Failure truncation*

  In this context, the system is monitored until a fixed number of failures. Thus, the likelihood function representation for the parameter vector $\boldsymbol{\theta}$, considering the observed failure times $\boldsymbol{t} = (t_0, t_1, \ldots, t_n)$, where $t_0 = 0$, should consider the random behavior of times to failure, however, the number of failures is a fixed amount. Thus,

  $$L(\boldsymbol{\theta};\boldsymbol{t}) = f(t_1 \mid t_0) f(t_2 \mid t_1) \cdots f(t_n \mid t_1, t_2, \ldots, t_{n-1}, n).$$

  The approximation of the term $f(t_i \mid t_0, \ldots, t_{i-1})$, where $i = 1, \ldots, n$, is analogous to that given in the situation of a time truncation. Thus, the likelihood function considering the intensity function of a ARA_1 model, and a failure truncation context, has the form

  $$L(\boldsymbol{\theta};\boldsymbol{t}) = \prod_{i=1}^{n} \lambda_1(t_i - (1-\rho)t_{i-1}) \exp\Big\{ - \Big(\Lambda_1(t_i - (1-\rho)t_{i-1}) - \Lambda_1(\rho\, t_{i-1})\Big)\Big\},$$

  and the log likelihood function is given by

  $$\ell(\boldsymbol{\theta};\boldsymbol{t}) = \sum_{i=1}^{n} \Big[\log \lambda_1(t_i - (1-\rho)t_{i-1}) - \Lambda_1(t_i - (1-\rho)t_{i-1}) + \Lambda_1(\rho\, t_{i-1}) \Big],$$

2.3 Competing Risks Model

Practical situations often admit numerous causes that compete to originate a particular event under study. Such a context is formalized mathematically and statistically with *Competing Risk Theory*.

A common strategy is to assume p failure modes, which at each failure are denoted by $\delta(t) = j$, for $j = 1, \ldots, p$ (we suppress the explicit dependence call $\delta(t) = \delta$ on failure time t for brevity). Thus, if n failures are observed in the time interval $(0, T]$, then we have the data $(t_1, \delta_1), \ldots, (t_n, \delta_n)$, where $0 < t_1 < \cdots < t_n < T$ are the system failure times, and the δ_i's indicate the j-th failure mode associated with the i-th failure time, for $i = 1, \ldots, n$.

The counting process $N_j(t)$ with behavior according to the cause-specific intensity function

$$\lambda_j(t;\delta) = \lim_{\tau \to 0} \frac{\Pr(\delta = j, N_j(t+\tau) - N_j(t) = 1 | \mathcal{H}_{jt})}{\tau}.$$

Consequently, $N(t) = \sum_{j=1}^{p} N_j(t)$, which is the global system failure counting process, can be seen as a superposition process whose intensity function is given by

$$\lambda(t;\delta) = \sum_{j=1}^{p} \lambda_j(t;\delta).$$

The corresponding cause-specific and overall cumulative intensities are given, respectively, by

$$\Lambda_j(T) = \int_0^T \lambda_j(u;\delta)\, \mathrm{d}u \qquad \text{and} \qquad \Lambda(T) = \sum_{j=1}^{p} \Lambda_j(T).$$

2.4 Frailty Model

A well-known approach with fragility models was introduced by Vaupel, Manton and Stallard (1979), it is the direct extension of the Cox models introduced by Cox (1972). For a long time, research in fragility turned to the medical and reliability area, which had unobserved heterogeneity that could not be explained by Cox's model. A frailty term in the model can denote the influence of any ill-specified effect or whose verification could not be carried out, in other words, it is a latent effect. Commonly, frailty models are presented in multiplicative form, as a positive value that inflates, deflates or preserves the intensity function that describes the model (WIENKE, 2010; ANDERSEN *et al.*, 1996; HOUGAARD, 1995).

In the perspective of a study with several repairable systems, although the reliability of the systems have the same default behavior, there may be unobservable factors that all systems are subject to, but which influence their respective behaviors individually. So, to considere this unobserved heterogeneity, we consider a frailty model. This model is defined by the conditioned versions of the intensity, accumulated intensity and reliability functions, generically and respectively expressed by

$$\lambda(t\,|\,z) = z\underline{\lambda}(t), \qquad \Lambda(t\,|\,z) = z\underline{\Lambda}(t) \qquad \text{and} \qquad R(t\,|\,z) = e^{-\underline{\Lambda}(t)z},$$

where z denotes the frailty term, that inflates (when $z > 1$), deflates ($z < 1$) or preserves ($z = 1$) the failure intensity function.

Commonly, the probability model adopted for the random variable Z, whose realizations are the frailty terms, is the Gamma model, although this is an arbitrary choice as stated by Slimacek and Lindqvist (2016). In particular, the parameterization $\mathrm{Gamma}(1/\alpha, 1/\alpha)$ is used, whose probability density function is expressed by

$$f(z) = \frac{(1/\alpha)^{1/\alpha}}{\Gamma(1/\alpha)} z^{1/\alpha - 1} e^{-(1/\alpha)z},$$

where $\mathbb{E}(z) = 1$ and $\mathrm{Var}(z) = \alpha$. This avoids the inclusion of an identifiability problem in the model and also brings the possibility of interpreting unobserved heterogeneity based on the parameter α (ELBERS; RIDDER, 1982).

Based on this construction, we can find the marginal reliability function for failure times using the Laplace transform, since the reliability $R(t)$ is expressed by

$$R(t) = \int_0^\infty R(t|z) f(z)\,\mathrm{d}z = \int_0^\infty e^{-\underline{\Lambda}(t)z} f(z)\,\mathrm{d}z,$$

and integrals written in this way have a well-known Laplace transform, given by $Q(\underline{\Lambda}(t)) = (1+\alpha\underline{\Lambda}(t))^{-1/\alpha}$ and can be used to solve this problem (WIENKE, 2010). In this sense

$$\lambda(t) = \frac{1}{Q(\underline{\Lambda}(t))}\frac{\mathrm{d}}{\mathrm{d}t}\Big[Q(\underline{\Lambda}(t))\Big], \qquad \Lambda(t) = -\log Q(\underline{\Lambda}(t)) \qquad \text{and} \qquad R(t) = Q(\underline{\Lambda}(t)),$$

and so,

$$\lambda(t) = -\frac{\underline{\lambda}(t)}{1+\alpha\underline{\Lambda}(t)}, \qquad \Lambda(t) = -\log\left[\left(1+\alpha\underline{\Lambda}(t)\right)^{-1/\alpha}\right] \qquad \text{and} \qquad R(t) = (1+\alpha\underline{\Lambda}(t))^{-1/\alpha},$$

The value of α represents the variance of the random variable Z and, therefore, brings indications of the existence of unobserved effects that could significantly affect the intensity of failures, making some systems more susceptible to failures than others.

CHAPTER

3

A REPAIRABLE SYSTEM SUBJECTED TO HIERARCHICAL COMPETING RISKS: MODELING AND APPLICATIONS

In this chapter, we propose modeling for a single repairable system with a hierarchical structure under the assumption that the failures follow a NHPP (see 2.2.3) with a power-law intensity function (see 2.2.3.4.1). The properties of the new model are discussed in detail. The parameter estimators are obtained using the maximum likelihood method. A corrective approach is used to remove bias with order $O(n^{-1})$, and the respective exact confidence intervals are proposed. A simulation study is conducted to show that our estimators are bias-free. The proposed modeling is illustrated via a toy example on a butterfly valve system, an example of an early-stage real project related to the traction system of an in-pipe robot, and also a real example on a blowout preventer system.

Contents of Chapter

3.1 The Proposed Framework

A major challenge when modeling repairable system data is how to consider the effect of a repair action taken immediately after a failure has occurred. It is usually supposed, for the sake of simplicity, that the repair actions are instantaneous. However, it is not suitable for many real systems. Therefore, the most investigated assumptions are either minimal or perfect repair at failures. In the former, it is assumed that the repair action after a failure restores the system (i.e., the intensity) to the same state as it was before the failure, e.g., by replacing a failed minor component (flat tire) of a large composite system (car); while in the latter, the repair action leaves the system as if it was new, e.g., via replacement of a failed system (an engine with a broken connecting rod) by a brand new one (WANG; PHAM, 2006). According to the engineering literature, these repair or maintenance actions are often called ABAO and AGAN, respectively (Aven (1983), Barlow and Hunter (1960), Aven and Jensen (2000), Finkelstein (2004), Mazzuchi and Soyer (1996)). However, more complex models suppose that the repair effect lies between ABAO and AGAN (i.e., the failure intensity is reduced to a level between ABAO and AGAN). These models are known as IR models, but they are not considered here (see, e.g., Doyen and Gaudoin (2004)).

3.1.1 MR Model

The repairable system model for the failure data will be implemented according to NHPP under the assumption of MR. Furthermore, based on the time truncation design, the likelihood and corresponding log-likelihood function for a collection of failure data up to time T, are expressed as

$$L(\beta,\mu \mid n,\boldsymbol{t}) = \frac{\beta^n}{\mu^{n\beta}}\left(\prod_{i=1}^{n} t_i\right)^{\beta-1} \exp\left\{-\left(\frac{T}{\mu}\right)^{\beta}\right\},$$

and

$$\ell(\beta,\mu \mid n,\boldsymbol{t}) = n\log(\beta) + (\beta-1)\sum_{i=1}^{n}\log(t_i) - \left(\frac{T}{\mu}\right)^{\beta} - n\beta\log(\mu),$$

respectively, where we assume that for $n \geq 1$, failures are observed at times $t_1 < t_2 < \cdots < t_n < T$ (see, e.g., Rigdon and Basu (2000)). The MLE of β and μ, which are both biased, can be written as

$$\widehat{\beta} = \frac{n}{\sum_{i=1}^{n}\log\left(\frac{T}{t_i}\right)} \qquad \text{and} \qquad \widehat{\mu} = \frac{T}{n^{1/\widehat{\beta}}}. \tag{3.1}$$

Since the MLE (3.1) suffer from bias, and inadequate confidence intervals for small samples, several studies have been performed to overcome these drawbacks. Some further discussions are given in Section 3.4.

There is always a concern about how to determine the confidence intervals under the classical inference. For the sake of illustration, Rigdon and Basu (2000) present the confidence interval for the scale parameter. The results showed that such an interval has no simple interpretation. Moreover, the authors found that the usual methodologies result in very long intervals. In turn, in some cases, the pivotal quantity, which is used to derive the aforementioned classical intervals, does not exist, or it is difficult to be obtained.

Bain (2017) extensively investigated confidence intervals for the scale parameter. The outcome of their research has shown that due to the non-existence of the pivotal quantity in the setting of time truncated data, finding confidence intervals for the scale parameter becomes difficult. Despite the extensive efforts, in most cases, the approaches still have limitations. For instance, Gaudoin, Yang and Xie (2006) studied the interval estimation for the scale parameter according to the PLP model. They used the Fisher information matrix to derive asymptotic confidence intervals, while several constraints have been reported on their results. Wang, Xie and Zhou (2013) considered a more sophisticated approach to obtain a generalized confidence interval for the scale parameter under some usual assumptions. Furthermore, Somboonsavatdee and Sen (2015) have shown methods to obtain the frequentist confidence intervals for the scale parameter under competing risks.

Oliveira, Colosimo and Gilardoni (2012) suggested reparametrizing the PLP intensity in terms of β and α, where

$$\alpha = \mathbb{E}(N(T)) = \left(\frac{T}{\mu}\right)^{\beta}.$$

In this case, the likelihood function is given by

$$L(\beta,\alpha \mid n,\boldsymbol{t}) = \beta^n e^{-n\beta/\widehat{\beta}}\alpha^n e^{-\alpha}\prod_{i=1}^{n}\frac{1}{t_i} \propto \gamma\left(\beta \mid n+1,\frac{n}{\widehat{\beta}}\right)\gamma(\alpha \mid n+1,1),$$

where $\gamma(x|a,b) = b^a x^{a-1}e^{-bx}/\Gamma(a)$, for $x,a,b>0$, is the probability density function of a gamma distribution with shape parameter a and scale parameter b. It is worth mentioning that β and α are orthogonal parameters, which play an important role for Bayesian inference (see, e.g., Cox and Reid (1987)).

3.1.2 Competing Risks

In reliability theory, the most commonly used system configurations are series, parallel, and series-parallel. Particularly, components in a series system are connected so that the failure of one of all components results in the system failure. For example, Figure 13 illustrates the Fault Tree Analysis (FTA) of the system $1,\ldots,p$. A series system is known as a competing

risks model because its failure can be classified as one of the p possible risks (or failure modes), which compete with each other to occur first and cause the system failure. Competing risks can provide a complete analysis of the probabilistic behavior of failures as many other methodologies presented in the literature. However, competing risks has an additional feature addressing not only failure times but also their causes through a pair of observations. Furthermore, the competing risks model involves the pair of observations (t, δ), where $t > 0$ denotes the failure time, while δ is the indicator of the component that failed.

In order to understand the competing risks framework for investigating repairable systems, a single system would be to consider successive failures at calendar time $0 < t_1 < \ldots < t_n < T$. Let us suppose that failures happen from an underlying competing risks structure, meaning that the system fails by the earliest occurrence of one of p exclusive failure modes. In this case, it is generally possible to observe the failure mode $\delta(t_i)$ at the failure time t_i. And for the system level, let us denote $\{N(t), t > 0\}$ the cumulative failure counter. In fact, if $N_j(t)$ represents the counting process corresponding to the j-th failure mode, it is easy to demonstrate that $N(t) = \sum_{j=1}^{p} N_j(t)$. The cause-specific intensity function of this process is

$$\lambda_j(t; \delta(t)) = \lim_{\Delta t \to 0} \frac{P(\delta(t) = j, N(t+\Delta t) - N(t) = 1 \mid N(s), 0 \leq s \leq t)}{\Delta t}, \tag{3.2}$$

for $j = 1, \ldots, p$.

According to equation (3.2), the time and the failure mode are stochastically independent if and only if $\lambda_1(t), \ldots, \lambda_p(t)$ are proportional to each other, giving a simple extension of a similar result from the competing risks literature in failure time modeling of non-repairable systems.

As pointed out by many works in the literature, complex repairable systems are mostly considered under the assumption of stochastic independence, which is based on the physically independent functioning of components (see, e.g., Yusof and Abdullah (2016), Crowder (2001), Høyland and Rausand (2009) and Wu and Scarf (2017)). It is essential to mention that the results exposed here is also based on this common assumption of independent risks, or equivalently, independent failure modes.

3.1.3 Modeling MR under Competing Risks

Let us assume that the system is observed up to time T, and that the adopted model is reparametrized in terms of β_j and

$$\alpha_j = \mathbb{E}[N_j(T)] = \left(\frac{T}{\mu_j}\right)^{\beta_j},$$

where $N_j(.)$ is the j-th cause-specific counting process, for $j = 1, \ldots, p$. This implies that β_j and α_j are orthogonal parameters.

A common strategy is to assume that the components of the repairable system under investigation can implement different operations, which are subject to different kinds of failure. Let us consider p failure modes, which at each failure are denoted by $\delta(t) = j$, for $j = 1, \ldots, p$ (in the sequel, we will suppress the explicit dependence of δ on failure time t for brevity). Thus, if n failures are observed in the time interval $(0, T]$, then we have the data $(t_1, \delta_1), \ldots, (t_n, \delta_n)$, where $0 < t_1 < \cdots < t_n < T$ are the system failure times, and the δ_i's indicate the j-th failure mode associated with the i-th failure time, for $i = 1, \ldots, n$.

Let us consider again the counting process $N_j(t)$ with behavior according to the cause-specific intensity function

$$\lambda_j(t;\delta) = \lim_{\Delta t \to 0} \frac{P(\delta = j, N(t, t+\Delta t] \geq 1)}{\Delta t}.$$

Consequently, $N(t) = \sum_{j=1}^{p} N_j(t)$, which is the global system failure counting process, can be seen as a superposition of NHPPs whose intensity function is given by

$$\lambda(t;\delta) = \sum_{j=1}^{p} \lambda_j(t;\delta).$$

The corresponding cause-specific and overall cumulative intensities are given, respectively, by

$$\Lambda_j(T) = \int_0^T \lambda_j(u;\delta)\, \mathrm{d}u \quad \text{and} \quad \Lambda(T) = \sum_{j=1}^{p} \Lambda_j(T).$$

Under the assumption that the failures from the j-th cause follow a NHPP with power-law intensity function, we can write the cause-specific intensities as

$$\lambda_j(t;\delta) = \left(\frac{\beta_j}{\mu_j}\right)\left(\frac{t}{\mu_j}\right)^{\beta_j - 1},$$

and

$$\Lambda_j(T) = \left(\frac{T}{\mu_j}\right)^{\beta_j} = \mathbb{E}(N_j(T)),$$

for $j = 1, \ldots, p$.

3.2 The Proposed Model – Hierarchical Competing Risks Model

In this section, we propose to analyze failure data representing events from a single repairable system studied under the parametric framework of a PLP that is subject to hierarchical competing risks. It consists of a generalization of the work done by Somboonsavatdee and Sen (2015) for the cases where there is the presence of secondary failure causes (sub-systems or sub-trees' branches), as illustrated in Figure 13. The hierarchical competing risks problem/structure

can also be represented by a block diagram showing a nested series system, where the sub-systems (and also the systems) are connected in such a way that the failure of a single sub-system (or component) results in the corresponding system failure and, consequently, in the whole system failure (see Figure 14).

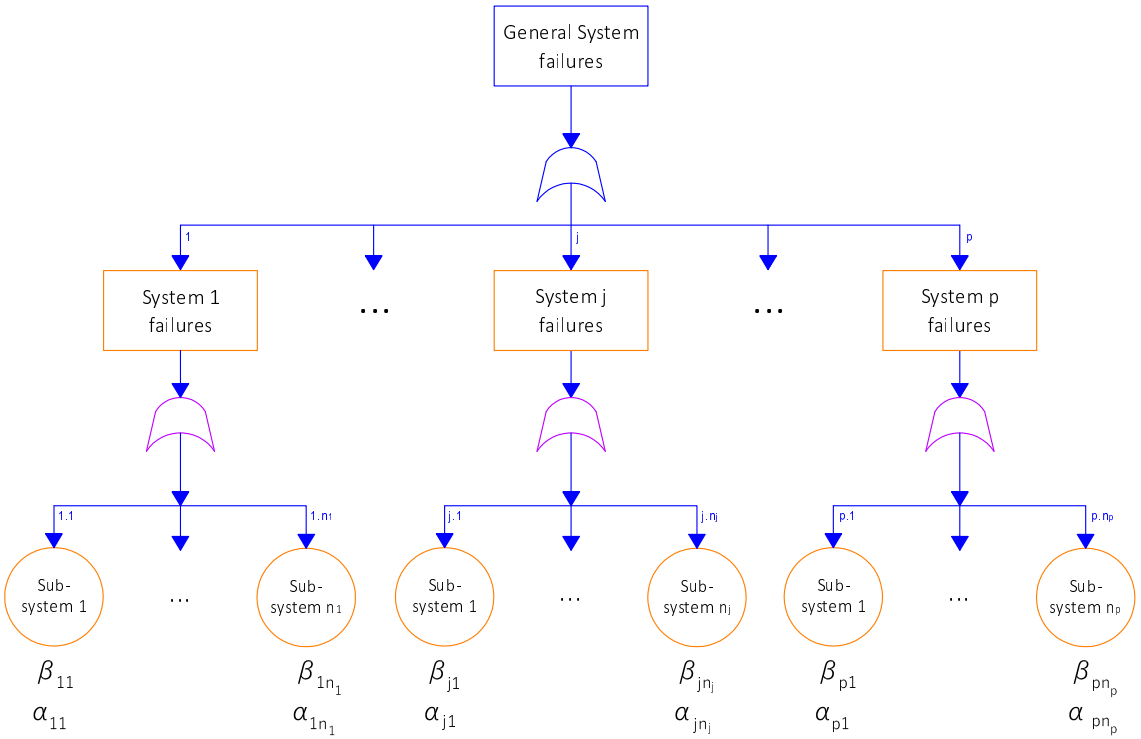


Figure 13 – The general system structure (FTA) considering our proposed hierarchical competing risks model.

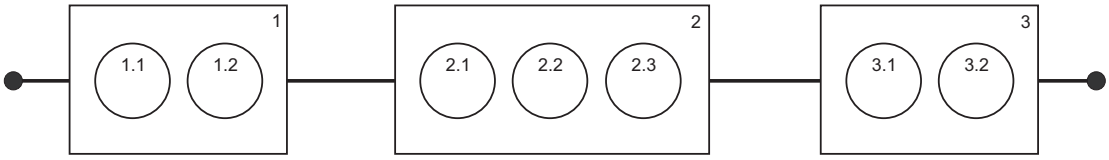


Figure 14 – Block diagram for a nested series system with 2, 3 and 2 series sub-systems (or components) within series systems 1, 2 and 3, respectively.

The hierarchical competing risks model's data consist of 3-tuples (t, δ, ψ), where $t > 0$ denotes the failure time, δ is the indicator of the leading failure cause (system), and ψ is the indicator of the sub-cause (sub-system).

Then, our proposed model for failure analysis can be formulated as follows. First, we assume that the failures from a sub-system k of a system j follow an NHPP with intensity function given by

$$\lambda_{jk}(t;\delta,\psi) = \left(\frac{\beta_{jk}}{\mu_{jk}}\right)\left(\frac{t}{\mu_{jk}}\right)^{\beta_{jk}-1}, \tag{3.3}$$

for $j = 1, \ldots, p$, and $k = 1, \ldots, n_j$, with n_j being the number of sub-systems for the j-th system; $\mu_{jk} > 0$ and $\beta_{jk} > 0$ are, respectively, the scale and shape parameters.

It follows that

$$\lambda(t) = \sum_{j=1}^{p} \sum_{k=1}^{n_j} \lambda_{jk}(t; \delta, \psi), \tag{3.4}$$

is the hazard function at time t. The sub-system-specific cumulative intensity is

$$\Lambda_{jk}(T) = \left(\frac{T}{\mu_{jk}}\right)^{\beta_{jk}}. \tag{3.5}$$

Here, we assumed that the failure causes related to the sub-systems are independent, therefore we expect that the failures may occur at different times. Nevertheless, if the failure of two or more sub-systems happens occasionally at the same time, the sub-system-specific cumulative intensity for each sub-system can be calculated from (3.5), hence the intensity function (3.4) can be computed in the presence of multiple failures at the same time.

It is seen from (3.4) that $\Lambda(T) = \sum_{j=1}^{p} \sum_{k=1}^{n_j} \Lambda_{jk}(T)$ is the cumulative hazard function at time T. Thus, we have that the reliability function is given by

$$R(T) = \exp\{-\Lambda(T)\} = \exp\left\{-\sum_{j=1}^{p} \sum_{k=1}^{n_j} \Lambda_{jk}(T)\right\}, \tag{3.6}$$

while the sub-system-specific reliability function is

$$R_{jk}(T) = \exp\left\{-\Lambda_{jk}(T)\right\}. \tag{3.7}$$

Then, similarly as in Section 3.1.3, we consider that the sub-system's lifetime is observed up to time T and we reparametrize our model in terms of β_{jk} and

$$\alpha_{jk} = \mathbb{E}[N_{jk}(T)] = \left(\frac{T}{\mu_{jk}}\right)^{\beta_{jk}}, \tag{3.8}$$

where $N_{jk}(.)$ is the j-th system and k-th sub-system-specific counting process.

3.3 Inference

In this section, we describe classical inference for the model that we introduced in Section 3.2. The MLE and Fisher information matrix, which is used for estimating the asymptotic variances of the MLE, are presented here.

Given the common (but sometimes unrealistic[1]) assumption that the failure modes act independently and are mutually exclusive, the classical inference for the proposed model is

[1] As pointed out by Meeker and Escobar (2014), it is possible that the failure of one component may either degrade or improve the reliability of other components, thus leading to either a positive or negative correlation between failure times in different system's components. Moreover, when this dependence exists, it is usually positive, since short (long) failure times of one mode tend to go with short (long) failure times of another.

conducted using the likelihood function, or equivalently, the log-likelihood function, which are defined, respectively, as follows:

$$L(\boldsymbol{\phi} \mid \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi}) = \prod_{i=1}^{n} \prod_{j=1}^{p} \prod_{k=1}^{n_j} [\lambda_{jk}(t_i; \delta_i, \psi_i)]^{\mathbb{I}(\delta_i = j, \psi_i = k)} \exp\left\{ - \sum_{j=1}^{p} \sum_{k=1}^{n_j} \Lambda_{jk}(T) \right\}, \tag{3.9}$$

and

$$\ell(\boldsymbol{\phi} \mid \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi}) = \sum_{i=1}^{n} \sum_{j=1}^{p} \sum_{k=1}^{n_j} \mathbb{I}(\delta_i = j, \psi_i = k) \log\left(\lambda_{jk}(t_i; \delta_i, \psi_i)\right) - \sum_{j=1}^{p} \sum_{k=1}^{n_j} \Lambda_{jk}(T), \tag{3.10}$$

where $\boldsymbol{\phi} = \left(\beta_{11}, \ldots, \beta_{pn_p}, \alpha_{11}, \ldots, \alpha_{pn_p}\right)$ denotes the general parameter vector; $\lambda_{jk}(t_i; \delta_i, \psi_i)$ and $\Lambda_{jk}(T)$ are given by (3.3) and (3.5), respectively; and $\mathbb{I}(\delta_i = j, \psi_i = k)$ is an indicator function. Before going further, it is important to mention again that our model is a generalization of the work by Somboonsavatdee and Sen (2015), which estimates the PLP in the presence of competing risks. However, our model has an additional hierarchical structure, and thus the proposed estimators require a comprehensive investigation to be performed.

The MLE can be obtained by maximizing the log-likelihood function (3.10). After some algebraic manipulation, such estimators can be written as

$$\widehat{\beta}_{jk} = \frac{n_{jk}}{\sum_{i=1}^{n} \log\left(\frac{T}{t_i}\right) \mathbb{I}(\delta_i = j, \psi_i = k)}, \tag{3.11}$$

and

$$\widehat{\alpha}_{jk} = n_{jk}, \tag{3.12}$$

where n_{jk} denotes the total number of failures due to the subcause k of the major cause j.

Since from (3.8), $\mathbb{E}(N_{jk}(T)) = \alpha_{jk}$, we have that the Fisher information matrix can be expressed as

$$I(\boldsymbol{\phi}) = \left(\begin{array}{ccc|ccc} I_{\boldsymbol{\beta}_1}(\boldsymbol{\phi}) & & 0 & & & \\ & \ddots & & & \mathbf{0} & \\ 0 & & I_{\boldsymbol{\beta}_p}(\boldsymbol{\phi}) & & & \\ \hline & & & I_{\boldsymbol{\alpha}_1}(\boldsymbol{\phi}) & & 0 \\ & \mathbf{0} & & & \ddots & \\ & & & 0 & & I_{\boldsymbol{\alpha}_p}(\boldsymbol{\phi}) \end{array}\right),$$

where, for each $j = 1, \ldots, p$,

$$I_{\boldsymbol{\beta}_j}(\boldsymbol{\phi}) = \begin{pmatrix} \frac{\alpha_{j1}}{\beta_{j1}^2} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{\alpha_{jn_j}}{\beta_{jn_j}^2} \end{pmatrix} \quad \text{and} \quad I_{\boldsymbol{\alpha}_j}(\boldsymbol{\phi}) = \begin{pmatrix} \frac{1}{\alpha_{j1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\alpha_{jn_j}} \end{pmatrix}.$$

$I(\boldsymbol{\phi})$ is a $\kappa \times \kappa$ diagonal matrix, with $\kappa = 2\sum_{j=1}^{p} j \times n_j$. The MLE have a closed-form expression and unique solution, consequently from the Central Limit Theorem, they are asymptotically normally distributed with a multivariate normal distribution, which can be given by

$$\widehat{\boldsymbol{\phi}} \sim N_{\kappa}(\boldsymbol{\phi}, I^{-1}(\boldsymbol{\phi})) \quad \text{as} \quad n_{jk} \to \infty. \tag{3.13}$$

3.3.1 Simulation Study

We conducted a simulation study to investigate the consistency and efficiency of the MLE presented in equations (3.11) and (3.12). To that end, we used two criteria: the Bias and Mean Square Error (MSE), which are given, respectively, by

$$\text{Bias}\,(\widehat{\phi}_w) = \frac{1}{M}\sum_{m=1}^{M}\left(\widehat{\phi}_w^{(m)} - \phi_w\right),$$

and

$$\text{MSE}\,(\widehat{\phi}_w) = \frac{1}{M}\sum_{m=1}^{M}\left(\widehat{\phi}_w^{(m)} - \phi_w\right)^2,$$

for $w = 1, \ldots, \kappa$, where $M = 50{,}000$ is the number of Monte Carlo replications and

$$\boldsymbol{\phi} = (\phi_1, \ldots, \phi_\kappa) = (\beta_{11}, \ldots, \beta_{pn_p}, \alpha_{11}, \ldots, \alpha_{pn_p})$$

represents the parameter vector. Moreover, $\widehat{\phi}_w^{(m)}$ denotes the MLE of ϕ_w obtained from sample m, for $m = 1, \ldots, M$.

By this approach, it is expected that good estimators have both bias and MSE close to zero. In turn, reasonable confidence intervals, which are produced here using the asymptotic normality of the MLE (as given in equation (3.13)), are expected to be short with Coverage Probabilities (CP) close to the nominal value of 95%. In this work, all computations and simulations were performed using the R software (R Core Team, 2019).

In what follows, we present the results for the simulation study. Due to space constraints, the results are reported only for these six scenarios (Table 1). However, similar findings are obtained for other parameter choices.

Table 1 – Six different scenarios with different parameter values in order to yield distinct sample sizes. A single system subject to 2 failure causes each with 3 sub-causes (Scenarios 1, 2 and 3), or 3 failure causes with 2, 3 and 2 sub-causes, respectively (Scenarios 4, 5 and 6), both under the assumption that the component system is observed in the fixed time interval $(0,T]$, where $T = 20$.

Cause and Sub-cause	Scenario 1		Scenario 2		Scenario 3		Scenario 4		Scenario 5		Scenario 6	
	α	β	α	β	α	β	α	β	α	β	α	β
1.1	3.30	0.80	14.56	1.20	10.92	1.20	8.00	2.00	4.39	0.80	8.00	2.00
1.2	2.23	0.50	10.37	0.90	13.49	1.10	4.97	0.40	8.24	0.80	40.48	1.10
1.3	7.00	1.00	12.00	1.00	12.74	1.20	—	—	—	—	—	—
2.1	4.00	1.00	4.00	1.00	17.89	1.50	12.07	0.60	9.05	0.60	72.82	1.20
2.2	2.00	1.00	20.00	2.00	21.97	1.80	12.21	0.70	4.67	0.30	40.48	1.10
2.3	2.00	1.10	10.92	1.20	12.07	1.60	8.94	1.50	8.10	1.10	8.94	1.50
3.1	—	—	—	—	—	—	9.94	0.40	8.94	0.50	80.96	1.10
3.2	—	—	—	—	—	—	7.69	0.80	7.37	1.30	25.49	1.20

By considering the well-known results regarding NHPPs (RIGDON; BASU, 2000), and also from the assumption that the failure modes are independent, we can generate the failure times, for each Monte Carlo replication, according to the following steps.

Data: p
$n_1, \ldots, n_p$
$\boldsymbol{\alpha} = (\alpha_{11}, \ldots, \alpha_{1n_1}, \ldots, \alpha_{jn_j}, \ldots, \alpha_{p1}, \ldots, \alpha_{pn_p})$
$\boldsymbol{\beta} = (\beta_{11}, \ldots, \beta_{1n_1}, \ldots, \beta_{jn_j}, \ldots, \beta_{p1}, \ldots, \beta_{pn_p})$
Result: $\{(\boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi})\} = \{(t_{111}, 1, 1), \ldots, (t_{n_{ijk}}, j, k), \ldots, (t_{n_{pn_p}pn_p}, p, n_p)\}$
for $j := 1$ **to** p **do**
  for $k := 1$ **to** n_j **do**
    $n_{jk} \sim \text{Poisson}(\alpha_{jk})$
    for $i := 1$ **to** n_{jk} **do**
      $U_{ijk} \sim \text{Uniform}(0,1)$
      $t_{ijk} = T\,U_{ijk}^{1/\beta_{jk}}$
      $\delta_{ijk} = j$
      $\psi_{ijk} = k$
    end
  end
end

Algorithm 1: How to generate the failure times.

As shown in Tables 2 and 3, the bias of the MLE varies depending on the α_{jk} parameter values, i.e., the mean number of failures. If the values of α_{jk} are small, the bias of $\widehat{\beta}_{jk}$ is considerably higher than expected, as well as the MSEs. This result is due to the systematic bias that the MLE of β_{jk} possesses. On the other hand, the maximum likelihood estimates of

the α_{jk}s are close to the true values, which is expected since $\widehat{\alpha}_{jk}$ is an unbiased estimator of α_{jk}. Note also that while the CP of the nominally 95% confidence intervals ($CP_{95\%}$) for the β_{jk}s seem to be satisfactory (i.e., they are close to 0.95), the $CP_{95\%}$ for the α_{jk}s are far from the assumed levels. This difference may occur because we are considering that the asymptotic normality of the MLE was achieved. However, this may not be true, returning inadequate confidence intervals. In order to overcome such a problem, in the next section, we will discuss an improved estimator for β_{jk}, as well as exact confidence intervals for both α_{jk} and β_{jk}.

Table 2 – Bias, MSE and $CP_{95\%}$ from the MLE, considering different parameter values (Scenarios 1, 2 and 3) and $M = 50{,}000$ simulated samples.

Parameter	Scenario 1			Scenario 2			Scenario 3		
	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$
α_{11}	0.485	1.642	0.998	−0.015	3.815	0.945	−0.007	3.302	0.908
β_{11}	0.408	2.430	0.955	0.099	0.406	0.955	0.144	0.579	0.955
α_{12}	0.813	1.425	0.997	0.022	3.220	0.936	−0.004	3.663	0.948
β_{12}	0.327	1.349	0.953	0.112	0.409	0.955	0.098	0.394	0.955
α_{13}	0.053	2.611	0.919	0.008	3.450	0.944	−0.003	3.557	0.927
β_{13}	0.214	0.841	0.956	0.102	0.407	0.954	0.115	0.459	0.954
α_{21}	0.325	1.829	0.997	0.325	1.822	0.997	−0.033	4.217	0.932
β_{21}	0.420	1.608	0.955	0.436	1.874	0.955	0.095	0.431	0.954
α_{22}	0.900	1.404	0.998	0.005	4.494	0.946	0.024	4.692	0.942
β_{22}	0.692	3.120	0.953	0.109	0.527	0.952	0.091	0.452	0.950
α_{23}	0.659	1.509	0.997	−0.012	3.284	0.910	−0.022	3.467	0.948
β_{23}	0.625	2.596	0.952	0.138	0.531	0.953	0.165	0.650	0.954

Table 3 – Bias, MSE and $CP_{95\%}$ from the MLE, considering different parameter values (Scenarios 4, 5 and 6) and $M = 50{,}000$ simulated samples.

Parameter	Scenario 4			Scenario 5			Scenario 6		
	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$
α_{11}	0.019	2.799	0.894	0.265	1.953	0.993	0.011	2.808	0.893
β_{11}	0.349	1.316	0.955	0.308	1.193	0.956	0.346	1.263	0.954
α_{12}	0.172	2.109	0.904	0.016	2.854	0.909	0.008	6.325	0.953
β_{12}	0.135	0.550	0.955	0.137	0.537	0.955	0.029	0.188	0.952
α_{21}	−0.020	3.480	0.948	−0.004	3.005	0.941	0.058	8.484	0.951
β_{21}	0.061	0.239	0.955	0.095	1.111	0.956	0.017	0.146	0.951
α_{22}	−0.019	3.488	0.910	0.217	2.019	0.996	0.017	6.396	0.950
β_{22}	0.071	0.279	0.956	0.105	0.377	0.954	0.028	0.186	0.951
α_{23}	0.033	2.958	0.942	0.024	2.828	0.900	−0.006	2.980	0.938
β_{23}	0.225	0.863	0.956	0.190	0.868	0.956	0.228	0.879	0.956
α_{31}	−0.014	3.146	0.923	−0.015	2.981	0.939	−0.009	8.978	0.944
β_{31}	0.053	0.205	0.955	0.079	0.432	0.957	0.013	0.126	0.952
α_{32}	0.019	2.747	0.945	0.042	2.683	0.931	−0.005	5.028	0.938
β_{32}	0.153	0.705	0.956	0.262	0.942	0.957	0.052	0.270	0.951

3.4 Bias Correction and Improved Confidence Intervals

Cox and Snell (1968) showed that, when the sample data are independent (although not necessarily identically distributed), the bias of $\widehat{\phi}_w$, for $w = 1, \ldots, \kappa$, can be written as

$$\text{Bias}(\widehat{\phi}_w) = \sum_{i=1}^{\kappa}\sum_{j=1}^{\kappa}\sum_{l=1}^{\kappa} s_{wi}(\boldsymbol{\phi}) s_{jl}(\boldsymbol{\phi}) \left(h_{ij,l}(\boldsymbol{\phi}) + 0.5 h_{ijl}(\boldsymbol{\phi})\right) + O(n^{-2}), \tag{3.14}$$

where s_{ij} is the (i,j)-th element of the variance-covariance matrix of $\widehat{\boldsymbol{\phi}}$,

$$h_{ijl}(\boldsymbol{\phi}) = \mathbb{E}\left[\frac{\partial^3 \log L(\boldsymbol{\phi})}{\partial\phi_i \partial\phi_j \partial\phi_l}\right],$$

and

$$h_{ij,l}(\boldsymbol{\phi}) = \mathbb{E}\left[\frac{\partial^2 \log L(\boldsymbol{\phi})}{\partial\phi_i \partial\phi_j} \cdot \frac{\partial \log L(\boldsymbol{\phi})}{\partial\phi_l}\right],$$

for $i, j, l = 1, \ldots, \kappa$.

Cordeiro and Klein (1994) proved that, even when the data are dependent, the bias expression (3.14) can be rewritten as

$$\text{Bias}(\widehat{\phi}_w) = \sum_{i=1}^{\kappa} s_{wi}(\boldsymbol{\phi}) \sum_{j=1}^{\kappa}\sum_{l=1}^{\kappa} s_{jl}(\boldsymbol{\phi}) \left(h_{ij}^{(l)}(\boldsymbol{\phi}) - 0.5 h_{ijl}(\boldsymbol{\phi})\right) + O(n^{-2}),$$

where

$$h_{ij}^{(l)}(\boldsymbol{\phi}) = \frac{\partial h_{ij}(\boldsymbol{\phi})}{\partial\phi_l}, \quad \text{for} \quad i, j, l = 1, \ldots, \kappa.$$

Firth (1993) showed that the first-order term is removed from the asymptotic bias of the MLE by considering the Jeffreys prior (JEFFREYS, 1946) as a penalty function in the likelihood equation for the exponential family of distributions. The Jeffreys prior can be obtained as the square root of the determinant of the expected Fisher information matrix $I(\boldsymbol{\phi})$. Thus, it follows from equation (**??**) that

$$\begin{aligned}
\pi^{\text{J}}(\boldsymbol{\phi}) &\propto |I(\boldsymbol{\phi})|^{1/2} \\
&= \left|\text{Diag}\left(\frac{\alpha_{11}}{\beta_{11}^2}, \ldots, \frac{\alpha_{1n_1}}{\beta_{1n_1}^2}, \ldots, \frac{1}{\alpha_{p1}}, \ldots, \frac{1}{\alpha_{pn_p}}\right)\right|^{1/2} \\
&= \prod_{j=1}^{p}\prod_{k=1}^{n_j} \frac{1}{\beta_{jk}}.
\end{aligned} \tag{3.15}$$

Note that after some algebraic manipulation, the likelihood function (3.9) can be rewritten as

$$L(\boldsymbol{\phi} \mid \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi}) \propto \prod_{j=1}^{p}\prod_{k=1}^{n_j} \gamma(\beta_{jk} \mid n_{jk}+1, n_{jk}/\widehat{\beta}_{jk}) \gamma(\alpha_{jk} \mid n_{jk}+1, 1).$$

The marginal distribution for each parameter is independent of the other parameters. Moreover, since the marginals follow a gamma distribution, they belong to the exponential family of distributions. Hence, the approach proposed by Firth (1993) is valid for our hierarchical competing risks model. The penalized log-likelihood function using the Jeffreys prior (3.15) as a penalized criterion can be written as

$$L_{\mathrm{P}}(\boldsymbol{\phi} \mid \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi}) \propto \prod_{j=1}^{p} \prod_{k=1}^{n_j} \gamma(\beta_{jk} \mid n_{jk}, n_{jk}/\widehat{\beta}_{jk}) \gamma(\alpha_{jk} \mid n_{jk}+1, 1),$$

Then, with some algebraic manipulation, we obtain the Bias-Corrected MLE (CMLE) given by

$$\tilde{\beta}_{jk} = \frac{n_{jk}-1}{n_{jk}} \widehat{\beta}_{jk}, \tag{3.16}$$

and

$$\tilde{\alpha}_{jk} = \widehat{\alpha}_{jk}, \tag{3.17}$$

which are unbiased to $O(n_{jk}^{-1})$. Although the penalized likelihood method introduced by Firth (1993) only ensures that the first-order term is removed from the asymptotic bias, we have that

$$\mathbb{E}\left[\tilde{\beta}_{jk} \mid \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi}\right] = \beta_{jk},$$

and

$$\mathbb{E}\left[\tilde{\alpha}_{jk} \mid \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi}\right] = \alpha_{jk}.$$

Therefore, the obtained CMLE are unbiased for $n_{jk} > 1$.

As we observed from the simulation results presented in Section 4.4, the asymptotic confidence intervals are not satisfactory for small samples. Using the improved estimates in the estimators of the asymptotic variance, which are used to obtain the confidence intervals, will return the worst results in terms of CPs than obtained with the MLE. On the other hand, by observing that

$$L_{\mathrm{P}}(\alpha_{jk} \mid \boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{\psi}) = \frac{\alpha_{jk}^{n_{jk}} e^{-\alpha_{jk}}}{n_{jk}!},$$

i.e., $\alpha_{jk} \sim \text{Erlang}(n_{jk}+1, 1)$, then $2\alpha_{jk} \sim \chi^2_{2(n_{jk}+1)}$. Therefore, the $100(1-\xi)\%$ confidence interval for α_{jk} can be calculated as

$$\left[\frac{1}{2}\chi^2_{2n_{jk}+2;\,\xi/2}; \frac{1}{2}\chi^2_{2n_{jk}+2;\,1-\xi/2} \right], \tag{3.18}$$

where $\chi^2_{a;\,\upsilon}$ represents the 100υ-th percentile of the chi-square distribution with a degrees of freedom.

Furthermore we have that the $100(1-\xi)\%$ confidence interval for β_{jk}. Since

$$L_{\mathrm{P}}(\beta_{jk}\,|\,\boldsymbol{t},\boldsymbol{\delta},\boldsymbol{\psi}) = \gamma\left(\beta_{jk}\;\middle|\; n_{jk}, \frac{n_{jk}}{\widehat{\beta}_{jk}}\right),$$

can be obtained directly from the quantile function of the gamma distribution, that is,

$$\left[\gamma_{\mathrm{Q}}\left(n_{jk}, \frac{n_{jk}}{\widehat{\beta}_{jk}}; \frac{\xi}{2}\right); \gamma_{\mathrm{Q}}\left(n_{jk}, \frac{n_{jk}}{\widehat{\beta}_{jk}}; 1-\frac{\xi}{2}\right)\right], \tag{3.19}$$

where $\gamma_{\mathrm{Q}}(a,b;\upsilon)$ is the quantile function of the gamma distribution with shape parameter a and scale parameter b, and $0 \leq \upsilon \leq 1$. This quantile function is available in most of the standard statistical softwares. For instance, in R it can be computed by using the *qgamma(.)* function. Thus, the exact confidence intervals for the model parameters can be obtained without the use of intensive computation.

3.4.1 Simulation Study

In this section, we perform a second simulation study with the same general specifications (i.e., same scenarios, number of Monte Carlo replications and evaluation criteria) of the first one shown in Section 4.4. However, the main goal now is to assess the performance (i.e., the consistency and efficiency) of the CMLE for the model parameters presented in equations (3.16) and (3.17), as well as of the exact confidence intervals given in equations (3.18) and (3.19). It is worthwhile mentioning that the generated samples are the same as those of Section 4.4, in order to achieve a fair comparison of the different approaches.

Tables 2, 3, 4 and 5 summarize the results. The CMLE of the β_{jk}s are more adequate, since their bias were successfully removed compared with their corresponding MLE. Moreover, the $\mathrm{CP}_{95\%}$ for the α_{jk}s using the exact confidence intervals, rather than the asymptotic confidence intervals, are in general higher and closer to the nominal value (0.95).

Table 4 – Bias, MSE and $CP_{95\%}$ from the CMLE, considering different parameter values (Scenarios 1, 2 and 3) and $M = 50{,}000$ simulated samples.

Parameter	Scenario 1			Scenario 2			Scenario 3		
	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$
α_{11}	0.485	1.642	0.940	−0.015	3.815	0.935	−0.007	3.302	0.954
β_{11}	0.004	1.262	0.950	0.001	0.358	0.951	0.003	0.459	0.949
α_{12}	0.813	1.425	0.960	0.022	3.220	0.911	−0.004	3.663	0.924
β_{12}	0.003	0.691	0.949	0.001	0.337	0.951	0.001	0.344	0.952
α_{13}	0.053	2.611	0.973	0.008	3.450	0.942	−0.003	3.557	0.952
β_{13}	−0.001	0.564	0.950	−0.001	0.347	0.950	0.001	0.396	0.950
α_{21}	0.325	1.829	0.944	0.325	1.822	0.945	−0.033	4.217	0.958
β_{21}	−0.001	0.896	0.950	0.006	1.012	0.950	−0.001	0.392	0.951
α_{22}	0.900	1.404	0.914	0.005	4.494	0.944	0.024	4.692	0.958
β_{22}	0.005	1.568	0.951	−0.003	0.486	0.951	0.000	0.419	0.947
α_{23}	0.659	1.509	0.914	−0.012	3.284	0.944	−0.022	3.467	0.958
β_{23}	−0.007	1.347	0.951	−0.001	0.441	0.949	0.001	0.550	0.949

Table 5 – Bias, MSE and $CP_{95\%}$ from the CMLE, considering different parameter values (Scenarios 4, 5 and 6) and $M = 50{,}000$ simulated samples.

Parameter	Scenario 4			Scenario 5			Scenario 6		
	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$	Bias	MSE	$CP_{95\%}$
α_{11}	0.019	2.799	0.957	0.265	1.953	0.961	0.011	2.808	0.954
β_{11}	−0.003	0.960	0.950	0.001	0.678	0.950	−0.008	0.939	0.950
α_{12}	0.172	2.109	0.968	0.016	2.854	0.949	0.008	6.325	0.944
β_{12}	0.000	0.314	0.951	0.000	0.380	0.949	0.001	0.181	0.950
α_{21}	−0.020	3.480	0.940	−0.004	3.005	0.937	0.058	8.484	0.949
β_{21}	0.000	0.205	0.951	0.002	0.588	0.950	0.000	0.143	0.950
α_{22}	−0.019	3.488	0.940	0.217	2.019	0.949	0.017	6.396	0.941
β_{22}	0.000	0.236	0.952	−0.002	0.222	0.951	−0.001	0.179	0.950
α_{23}	0.033	2.958	0.942	0.024	2.828	0.952	−0.006	2.980	0.939
β_{23}	−0.001	0.658	0.950	−0.002	0.581	0.950	−0.001	0.658	0.950
α_{31}	−0.014	3.146	0.965	−0.015	2.981	0.940	−0.009	8.978	0.950
β_{31}	0.000	0.163	0.950	0.002	0.271	0.952	−0.001	0.124	0.952
α_{32}	0.019	2.747	0.960	0.042	2.683	0.944	−0.005	5.028	0.955
β_{32}	0.002	0.476	0.950	0.002	0.671	0.951	0.000	0.254	0.949

3.5 Applications

In this section, we illustrate the usefulness of the new methodology considering three data sets: an artificial data set for a butterfly valve system (Section 4.5), a data set from a real early-stage project of an in-pipe robot traction system (Section 4.5.2), and a real data set consisting of failures of a blowout preventer system (Section 3.5.3).

3.5.1 Butterfly Valve System: A Toy Example

To illustrate the inference process in hierarchical competing risks model, we start with a toy example based on a butterfly valve system. Butterfly valves are exact, low cost and a lightweight valve with excellent capability and durability, consisting of fewer parts, which makes butterfly valves easy to maintain, repair and less structural support for productive use (YUSOF; ABDULLAH, 2016). They contain a disc, which is positioned in the center that can be rotated a quarter of a turn through a shaft running. For this reason, this kind of valve is known as quarter-turn valves (KIMURA *et al.*, 1995). The rotation of the disc determines the flow passing a pipe, whose maximum occurs when the disc is positioned parallel to the stream and minimum when perpendicular to it. The relative position between the geometric center of the disc and the shaft defines if the valve is namely symmetrical, eccentric or double eccentric (CORBERA; OLAZAGOITIA; LOZANO, 2016).

Butterfly valves include a wide range of applications with excellent isolation, throttling as well as on-off service and flow regulation (YUSOF; ABDULLAH, 2016). They provide reliable, long-term performance that satisfies a wide range of industrial applications such as oil and gas. For instance, the applications involve isolation or regulating of oil and gas equipment, fill and drain or bypass systems and other similar applications where the principal function for the control of the flow or pressure can be satisfied whether on or off (SONG *et al.*, 2009).

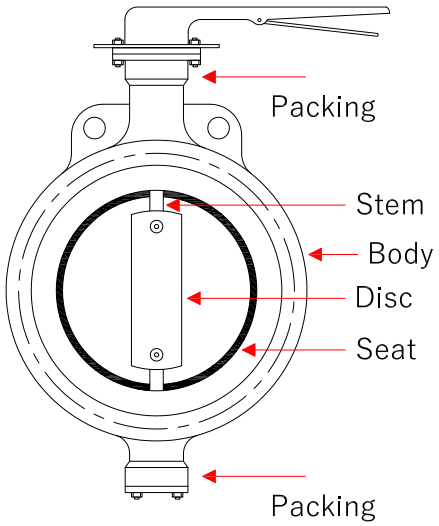


Figure 15 – Schematic diagram of a butterfly valve.

As shown in Figure 15, the butterfly valve includes a disc valve, placed inside one configured valve body that rotates about its axis separate from the axis of rotation of the stems that support the disc valve in position for a turn between opening and closure. A packing part is located connecting the valve body and the stem to prevent any leakage happening when the flow passes into the pipe. Furthermore, a ring seal acts as a seal between the metal disc and body to avoid any leakage when the valve is in the fully closed position.

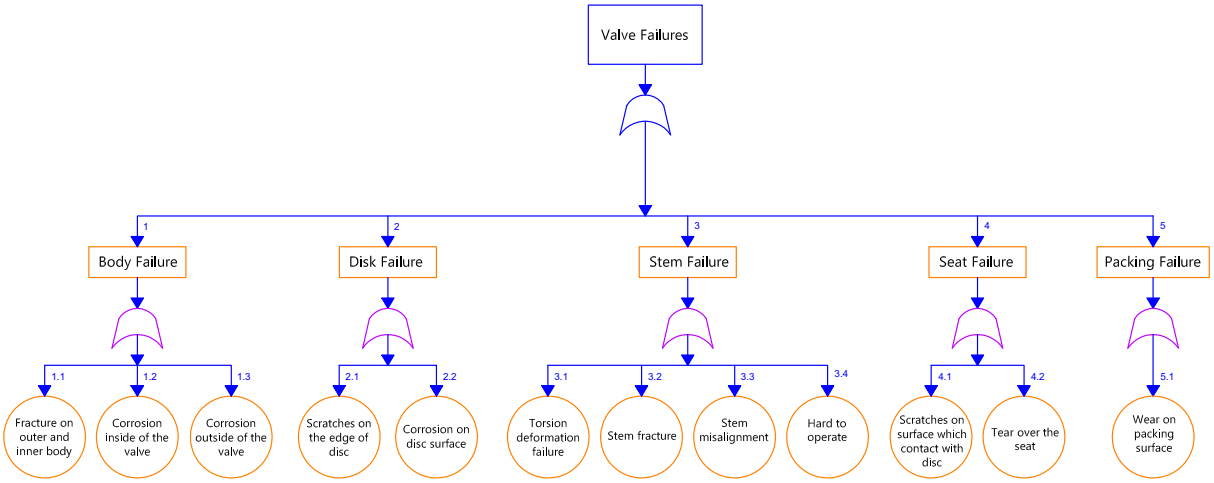


Figure 16 – FTA of butterfly valve failure.

Because of the critical impact of the butterfly valve in the industry, in this work, the attention is focused on carrying an FTA to increase the performance of this type of valve. The main goal is to know the failures and with new maintenance limits or avoid different risks within the valve performance. Hence, the FTA moves towards higher reliability, higher quality, and improved safety. As can be seen in Figure 16, we created the FTA based on the Failure Mode and Effects Analysis (FMEA) available in Bin and Abdullah (YUSOF; ABDULLAH, 2016), with reviewing primary components of the butterfly valve, which consists of a body, metal, disc, stem, seat and packing with several failure modes and their causes. It is worthwhile mentioning that these failures happen due to one of the series competing failure mechanisms, whereby each of them act related to the system independently. Based on the information provided in the FMEA by Bin and Abdullah (YUSOF; ABDULLAH, 2016), we were able to generate the data set, as shown in Table 6, which is representative of a butterfly valve system. Two numbers represent the failure modes, say 1.1, the first one stands for the system, in our example, system 1, and the second number stands for the sub-system, in our example, sub-system 1. Furthermore, we can evaluate the proportion of the PLP for each cause of failure by employing a graphical tool, which is known as the Duane plot (RIGDON; BASU, 2000). As can be seen in Figure 17, the values of the sub-systems are close to the line. This means that the obtained data set comes from a PLP, and our approach can be adequately used.

Table 6 – Failure data for a butterfly valve.

Failure Time	Failure Mode	Failure Time	Failure Mode	Failure Time	Failure Mode
32.03	1.1	10.38	1.3	46.54	4.1
7.37	1.1	57.37	2.1	33.63	4.1
4.38	1.1	2.68	2.1	25.58	4.1
44.35	1.1	13.24	2.1	20.84	4.1
30.00	1.1	15.45	2.1	39.87	4.1
0.41	1.2	12.42	2.2	22.17	4.1
1.71	1.2	50.39	2.2	4.91	4.1
45.86	1.2	56.48	2.2	57.63	4.1
0.13	1.2	47.89	2.2	26.00	4.2
16.41	1.2	33.93	2.2	31.30	4.2
17.40	1.2	1.52	3.1	51.34	4.2
36.61	1.2	48.39	3.1	8.38	4.2
0.22	1.2	3.55	3.2	44.89	5.1
47.98	1.3	43.97	3.2	35.09	5.1
4.98	1.3	10.46	3.3	48.25	5.1
8.71	1.3	37.14	3.3	11.46	5.1
59.46	1.3	24.93	3.4	9.22	5.1
7.87	1.3	4.68	3.4	48.59	5.1
41.67	1.3	33.72	3.4	35.95	5.1

Table 7 displays the bias-corrected maximum likelihood (CML) estimates, along with the corresponding 95% exact confidence intervals (CI 95%) for the model parameters. The results shown in this table suggest that the reliability of the body and stem components improve over time since the corresponding $\widehat{\beta}_{1k}$s and $\widehat{\beta}_{3k}$s are less than one. Moreover, observe that the reliability of the disk components may decrease over time due to corrosion on the disk surface ($\widehat{\beta}_{22} = 1.535 > 1$), while the reliability of the seat and packing components show an intermediate behavior since their CML estimates are close to one. Note, however, that almost all the CI 95% include the one. Therefore, we can not say that the cause-specific intensity functions of some components increase or decrease over time.

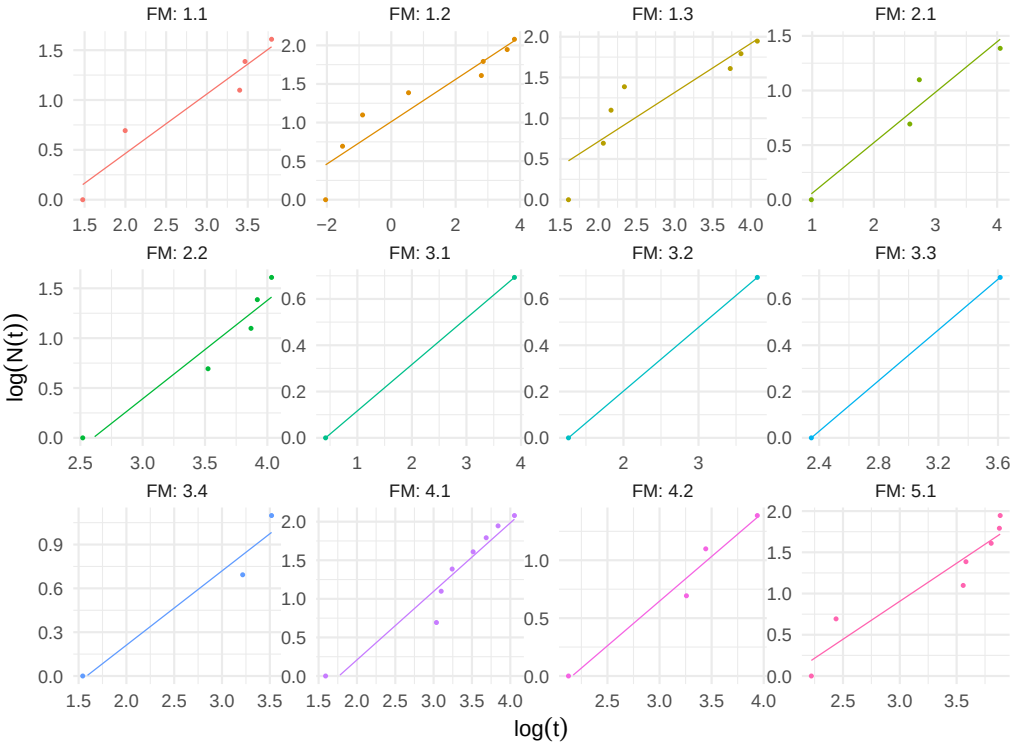


Figure 17 – Duane plots for the failure modes of a butterfly valve.

Table 7 – CML estimates and CI 95% of parameters β_{jk} and α_{jk}, considering the butterfly valve failure data.

Parameter	Estimate	95% Confidence Interval	
		Lower	Upper
β_{11}	0.631	0.256	1.616
α_{11}	5.000	2.202	11.668
β_{12}	0.297	0.146	0.612
α_{12}	8.000	4.115	15.763
β_{13}	0.682	0.320	1.484
α_{13}	7.000	3.454	14.423
β_{21}	0.498	0.181	1.456
α_{21}	4.000	1.623	10.242
β_{22}	1.535	0.623	3.931
α_{22}	5.000	2.202	11.668
β_{31}	0.257	0.062	1.432
α_{31}	2.000	0.619	7.225
β_{32}	0.319	0.077	1.775
α_{32}	2.000	0.619	7.225
β_{33}	0.449	0.109	2.502
α_{33}	2.000	0.619	7.225
β_{34}	0.499	0.154	1.804
α_{34}	3.000	1.090	8.767
β_{41}	1.046	0.516	2.156
α_{41}	8.000	4.115	15.763
β_{42}	0.831	0.302	2.428
α_{42}	4.000	1.623	10.242
β_{51}	1.133	0.531	2.466
α_{51}	7.000	3.454	14.423

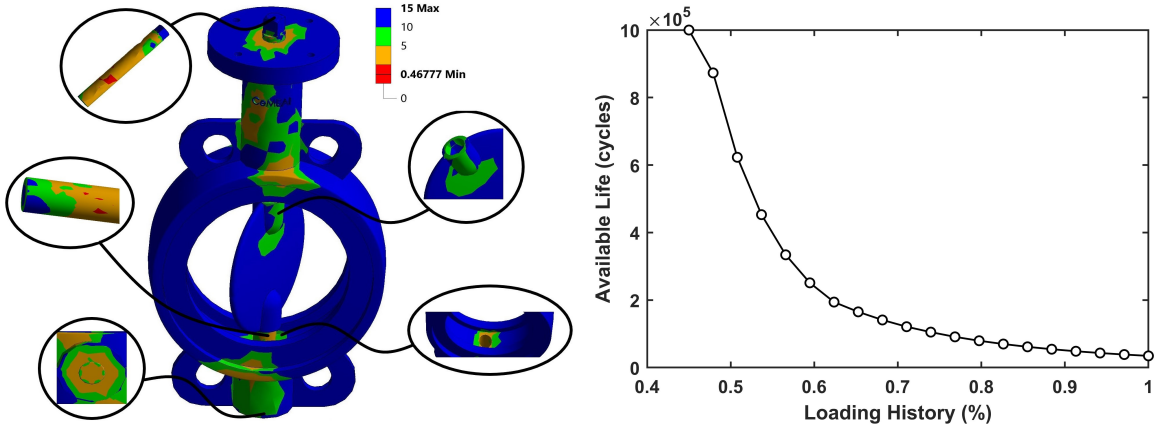
Table 8 – Estimates of the subsystem-specific and overall intensity functions at different times, considering the butterfly valve failure data.

Intensity	Time (months)				
	5	15	25	40	55
λ_{11}	0.132	0.088	0.073	0.061	0.054
λ_{12}	0.227	0.105	0.073	0.053	0.042
λ_{13}	0.175	0.124	0.105	0.091	0.082
λ_{21}	0.116	0.067	0.052	0.041	0.035
λ_{22}	0.034	0.061	0.080	0.103	0.122
λ_{31}	0.054	0.024	0.016	0.012	0.009
λ_{32}	0.058	0.027	0.019	0.014	0.011
λ_{33}	0.059	0.032	0.024	0.019	0.016
λ_{34}	0.087	0.050	0.039	0.031	0.026
λ_{41}	0.124	0.131	0.134	0.137	0.139
λ_{42}	0.084	0.070	0.064	0.059	0.056
λ_{51}	0.095	0.110	0.118	0.125	0.131
λ	1.245	0.888	0.797	0.744	0.723

It is essential to point out that such results can provide valuable insights to the maintenance crew. They also allow us to estimate the intensity function of each system or sub-system and the hazard function of the overall system. The estimated intensity functions can be obtained from (3.3), while the overall hazard function can be obtained from (3.4), with the parameters substituted by their estimates. Table 8 presents the estimated intensity functions for each sub-system and the estimated hazard function over some fixed failure times. Observe that the results shown in this table are in agreement with the ones presented in Table 7, that is, for the cases where $\widehat{\beta}_{jk} > 1$ the intensity function increases over time, while for $\widehat{\beta}_{jk} < 1$ the intensity function decreases over time. We also see that the overall hazard function decreases over time, which may be due to the repair and maintenance effects.

In order to provide a better understanding of the effect of fatigue damage on any point of the butterfly valve components, we then created a fatigue simulation, which is given as follows. Fatigue design of the butterfly valve is done using design fatigue curves, which are created based on the relationship between fatigue life and stress or strain. Because in the real structure of the valve, its components are constantly subjected to the high cycle fatigue stress and therefore, cracks begin from regions of concentrated stress resulting from this cycle fatigue and corresponding fatigue safety factor (SURESH, 1998). It is thus essential to determine the safety factor of fatigue failure, which indicates the ability of damage in the critical area in the valve components. The safety factor for this valve is determined by evaluating the effects of the loading history due to the fluid-structure interaction on fatigue life (HASUNUMA *et al.*, 2019). All computations and simulations required to build Figures 18a and 18b were performed with coupling CFD (fluids) and FEM (mechanics) models, which were prepared using the commercial pieces of software FLUENT and ANSYS 2019 R1. Figure 18a shows the safety

factor of fatigue life associated with any point of the butterfly valve components. As can be seen, the safety factor values are presented in terms of change between 0 and 15, in which the lower value identifies the critical damage locations in the component. Analysis over the safety factor of fatigue life results revealed that fatigue could be expected to initiate near the stems, due to their role as stress concentration points. It can be observed that the butterfly valve has several components connected in series. Therefore, a single component failure results in total system failure. The available fatigue life curve, in cycles, for the estimation of a finite lifetime of the butterfly valve under 50% to 100% of the fatigue loading history, is presented in Figure 18b. Analysis of the outcomes shows how the fatigue results change as a function of the loading at the critical location on the model. For instance, the results from this figure verified that the minimum value of fatigue life appeared at the maximum fatigue loading of 100%. Therefore, damage starts from the points related to the component with the lowest fatigue safety factor due to the significant stress concentration. Finally, as a conclusion, a good comparison was observed between the simulated fatigue damage results and statistical analysis. However, the present safety factor fatigue simulation indicates some disagreement, which is possibly related to differences in conditions between this simulation and statistical analysis.



(a) Fatigue safety factors of the butterfly valve components.

(b) Available fatigue life of the butterfly valve components.

Figure 18 – Fatigue analysis of the butterfly valve.

3.5.2 *In-Pipe Robot Traction System: Example on Early-Stage Innovative Project*

In this section, we consider another example based on a real problem we are working on in a partnership with Petrobras (abbreviation of *Petróleo Brasileiro S.A.*), which is the Brazil's largest oil and gas producer. The problem is related to the traction system of an in-pipe robot that was developed, though still in its early stage of development, to be used at a future time to remove hydrates that form in pipelines and can cause problems in oil and gas flow. In this case, a locomotive is responsible for conducting the robot inside the pipe, and once hydrate

formation is identified, the robot will work on its safe removal for the oil to flow again. A schematic of the studied system is shown in Figure 19.

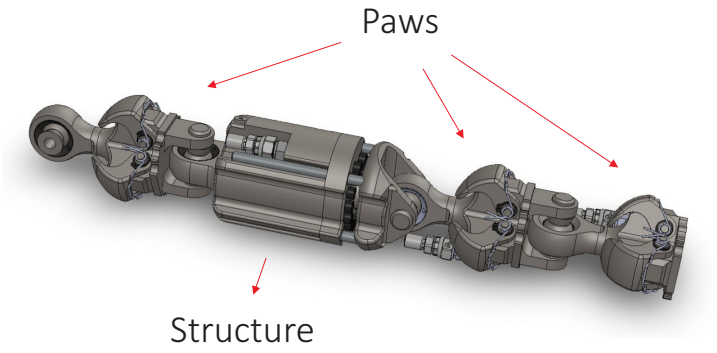


Figure 19 – Schematic diagram of an in-pipe robot traction system.

We obtained a suitable data set for this problem using a similar approach as proposed in the previous section, i.e., based on the limited but available information provided by FMEA and FTA tools. Due to the criticality of the traction system, we will focus on it instead of the overall locomotive system. An excerpt from the FMEA devised by the project executing team is shown in Table 9. On the other hand, the hierarchy of failure modes that compete with each other to cause a general system failure can be seen in Figure 20. Thus, these two tools supported the generation of the data set shown in Table 10, whose parameter representativeness tries to express the degree of severity and occurrence of the FMEA used.

Table 9 – FMEA for the in-pipe robot traction system. S = Severity, O = Occurrence, D = Detection.

General System	System	Failure Mode	S	O	D
Mechanical components	Paws	Compromised paw lining adhesive	7	3	3
		No arms retraction	9	3	9
		Paw slip	9	3	9
		Riser deformation	9	3	9
		Riser rupture	9	1	9
		Rubber coating degradation	7	3	7
	Structure	Cracking by atomic hydrogen permeation	9	9	9
		Stress concentration	9	5	9

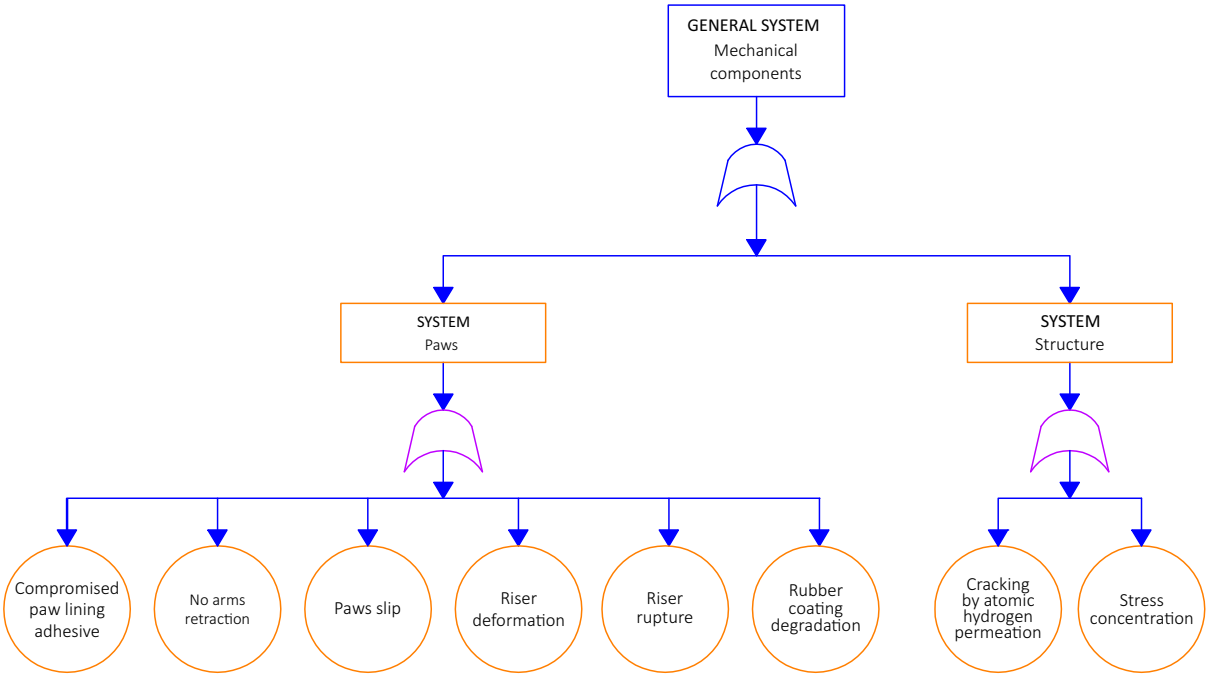


Figure 20 – FTA of in-pipe robot traction system failure.

Table 10 – Failure data for the in-pipe robot traction system. FT = Failure Time, FM = Failure Mode.

FT	FM	FT	FM	FT	FM	FT	FM	FT	FM	FT	FM
36.0	1.1	112.6	1.3	150.2	1.6	119.2	2.1	90.1	2.1	38.3	2.2
53.2	1.1	60.6	1.3	26.1	1.6	31.9	2.1	69.6	2.1	152.2	2.2
87.9	1.1	127.1	1.3	148.3	1.6	58.4	2.1	119.1	2.1	159.4	2.2
150.2	1.1	79.8	1.3	157.2	1.6	78.4	2.1	56.8	2.1	114.6	2.2
26.1	1.1	117.9	1.3	103.8	1.6	5.3	2.1	93.1	2.1	109.5	2.2
148.3	1.1	166.6	1.3	98.0	1.6	77.8	2.1	135.8	2.1	12.8	2.2
157.2	1.1	59.9	1.3	6.6	1.6	150.2	2.1	23.2	2.1	39.0	2.2
103.8	1.1	128.5	1.3	26.8	1.6	70.9	2.1	151.0	2.1	33.9	2.2
98.0	1.1	40.9	1.4	22.4	1.6	93.7	2.1	70.7	2.1	118.8	2.2
6.6	1.1	58.6	1.4	108.6	1.6	111.6	2.1	146.0	2.1	69.4	2.2
26.8	1.1	92.8	1.4	55.2	1.6	95.5	2.1	72.0	2.1	131.9	2.2
40.9	1.2	151.6	1.4	124.0	1.6	43.8	2.1	69.8	2.1	88.2	2.2
58.6	1.2	30.5	1.4	74.7	1.6	144.4	2.1	92.8	2.1	123.7	2.2
92.8	1.2	149.9	1.4	114.2	1.6	121.7	2.1	153.3	2.1	166.7	2.2
151.6	1.2	158.1	1.4	166.4	1.6	139.7	2.1	149.5	2.1	68.8	2.2
30.5	1.2	108.0	1.4	54.6	1.6	28.3	2.1	79.1	2.1	133.2	2.2
149.9	1.2	102.5	1.4	125.4	1.6	129.7	2.1	137.3	2.1	157.8	2.2
158.1	1.2	8.6	1.4	155.3	1.6	82.5	2.1	162.7	2.1	40.1	2.2
108.0	1.2	31.2	1.4	58.1	2.1	143.5	2.1	86.2	2.1	113.1	2.2
102.5	1.2	26.4	1.4	76.1	2.1	118.6	2.1	128.1	2.1	24.7	2.2
8.6	1.2	112.6	1.4	107.6	2.1	138.1	2.1	80.7	2.1	49.7	2.2
31.2	1.2	60.6	1.4	155.5	2.1	104.6	2.1	68.4	2.1	69.8	2.2
26.4	1.2	127.1	1.4	46.6	2.1	101.0	2.1	134.5	2.1	3.1	2.2
112.6	1.2	79.8	1.4	154.2	2.1	139.0	2.1	46.8	2.1	69.1	2.2
60.6	1.2	117.9	1.4	160.5	2.1	8.3	2.1	127.9	2.1	147.7	2.2
127.1	1.2	166.6	1.4	120.6	2.1	92.9	2.1	31.1	2.1	62.1	2.2
79.8	1.2	59.9	1.4	115.9	2.1	130.9	2.1	54.6	2.1	85.6	2.2
117.9	1.2	128.5	1.4	18.1	2.1	125.2	2.1	35.5	2.1	104.7	2.2
40.9	1.3	19.1	1.5	47.4	2.1	93.0	2.1	53.5	2.1	87.5	2.2
58.6	1.3	33.2	1.5	41.9	2.1	149.1	2.1	17.4	2.1	35.6	2.2
92.8	1.3	67.4	1.5	124.4	2.1	86.8	2.1	117.9	2.1	141.0	2.2
151.6	1.3	143.5	1.5	78.1	2.1	54.5	2.1	151.1	2.1	115.8	2.2
30.5	1.3	12.2	1.5	136.3	2.1	20.1	2.1	137.5	2.1	135.8	2.2
149.9	1.3	140.9	1.5	96.1	2.1	26.5	2.1	140.1	2.1	21.5	2.2
158.1	1.3	153.0	1.5	128.8	2.1	66.9	2.1	89.5	2.1	124.6	2.2
108.0	1.3	85.2	1.5	166.9	2.1	99.3	2.1	82.3	2.1	74.0	2.2
102.5	1.3	78.6	1.5	77.4	2.1	120.8	2.1	49.4	2.2	140.0	2.2
8.6	1.3	36.0	1.6	137.3	2.1	81.8	2.1	67.4	2.2	112.4	2.2
31.2	1.3	53.2	1.6	159.2	2.1	156.2	2.1	100.4	2.2	134.0	2.2
26.4	1.3	87.9	1.6	48.6	2.1	63.0	2.1	153.7	2.2	97.2	2.2

Duane plots applied to the data for each failure mode show evidence that a PLP may be able to adequately describe the behavior of system-associated failure times, since the scattered points show an approximately linear trend (see Figure 21).

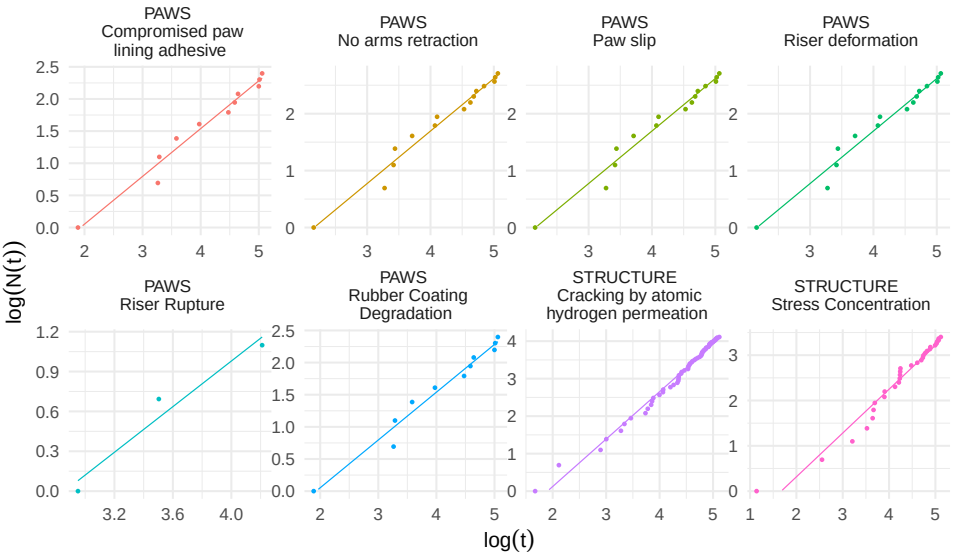


Figure 21 – Duane plots for the failure modes related to the in-pipe robot traction system.

The CML estimates for the parameters associated with each failure mode, as well as their respective CI 95%, are presented in Table 11. From these results, there is evidence that the intensity associated with some failure modes (e.g., *Cracking by atomic hydrogen permeation*) increases. On the other hand, there seems to be a decrease in risks associated with other failure modes (e.g., *Compromised paw lining adhesive* and *Riser rupture*), although their CI 95% contains the one.

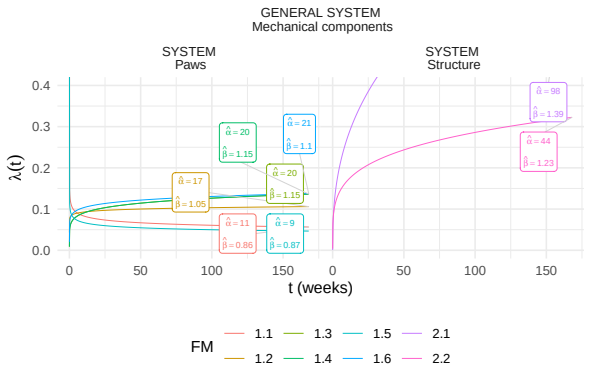
Table 11 – CML estimates and CI 95% of model parameters, considering the in-pipe robot traction system.

Failure Mode	Parameter	Estimate	95% Confidence Interval	
			Lower	Upper
Compromised paw lining adhesive	β_{11}	0.862	0.474	1.586
	α_{11}	11.000	6.201	19.682
No arms retraction	β_{12}	1.047	0.648	1.700
	α_{12}	17.000	10.668	27.219
Paw slip	β_{13}	1.145	0.736	1.789
	α_{13}	20.000	12.999	30.888
Riser deformation	β_{14}	1.145	0.736	1.789
	α_{14}	20.000	12.999	30.888
Riser rupture	β_{15}	0.870	0.448	1.714
	α_{15}	9.000	4.795	17.085
Rubber coating degradation	β_{16}	1.101	0.716	1.700
	α_{16}	21.000	13.787	32.101
Cracking by atomic hydrogen permeation	β_{21}	1.390	1.140	1.696
	α_{21}	98.000	80.462	119.431
Stress concentration	β_{22}	1.232	0.916	1.659
	α_{22}	44.000	32.823	59.068

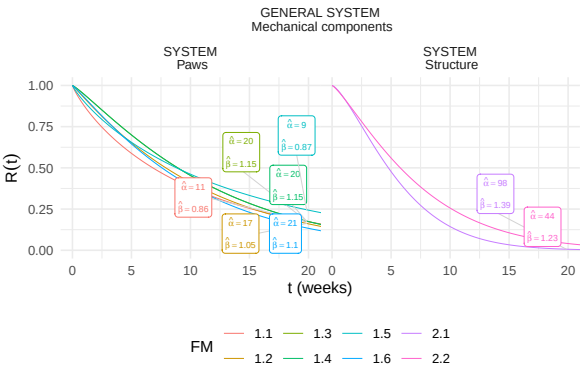
The behavior over time of the estimated intensities and reliabilities, for each specific failure mode, is shown in Figures 22a and 22b, from which we can see that the intensities associated with the failure modes 2.1 and 2.2 (*Structure* system) grow significantly more than others. In addition, the median lifespan of these sub-systems is around five weeks, while the rest is around seven weeks.

The combination of individual intensities and reliabilities, within their respective hierarchies, results in specific functions corresponding to each system. In this context, it is possible to study such measures by considering a level above in the hierarchy. The results are shown in Figure 22c, where it can be seen that the intensity of the *Paws* system grows significantly much less than the *Structure* system; however, in the (approximately) twenty initial weeks, its intensity is lower than the *Paws* system. This can also be observed from the reliability curve, where the median lifetime of the *Paws* system is around one week, while that of the *Structure* system is close to three. In addition, the curves become very close from week fifteen.

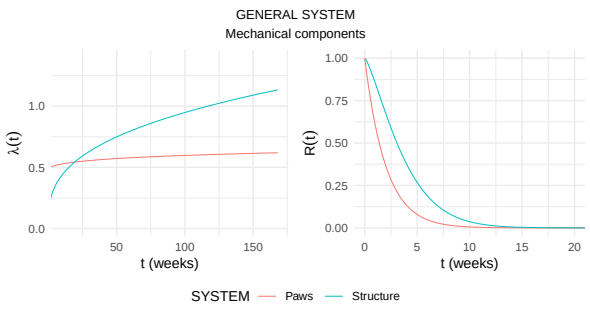
Finally, the combination of the intensities and reliabilities of all failure modes results in their respective functions for the general system, as a whole. Thus, there is a growing intensity associated with it, and a median time of operation close to one week.



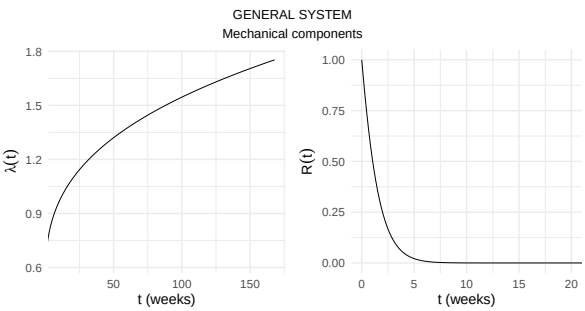
(a) Estimated intensity curves of failure modes.



(b) Estimated reliability curves of failure modes.



(c) Estimated intensity and reliability of systems.



(d) Estimated intensity and reliability of general system.

Figure 22 – General graphical results.

3.5.3 Blowout Preventer System: A Real Example

According Drægebø (2014), BlowOut Preventer (BOP) is a large, specially designed valve that is used to seal, control and monitor oil and gas wells. This valve mounts on top of the well during the drilling and completion stages of operation and serves as an essential barrier against blowouts, that is, the uncontrolled release of crude oil and/or natural gas from a well.

FTA for the BOP system is shown in Figure 23, which was done based on the real data set downloaded from the RAPID-S53 website (<https://www.rapid4s53.com>). This data set is available in Table 12. Analogously to Section 4.5, the failure modes are represented by two numbers, in the order that they appear from left to right in the graphical representation (again, the first number refers to the system, and the second one stands for the sub-system). It is worth mentioning that these failures occur due to a competing risks mechanism (in which we assume that each of them acts independently), and the safety equipment in question (BOP) is considered to be a repairable system.

First, we can evaluate the proportion of the PLP for each failure cause by using the Duane plot. As it can be observed from Figure 24, the values of the sub-systems are, in general, close to the line, which means that this data set comes from a PLP and our methods can be suitably used.

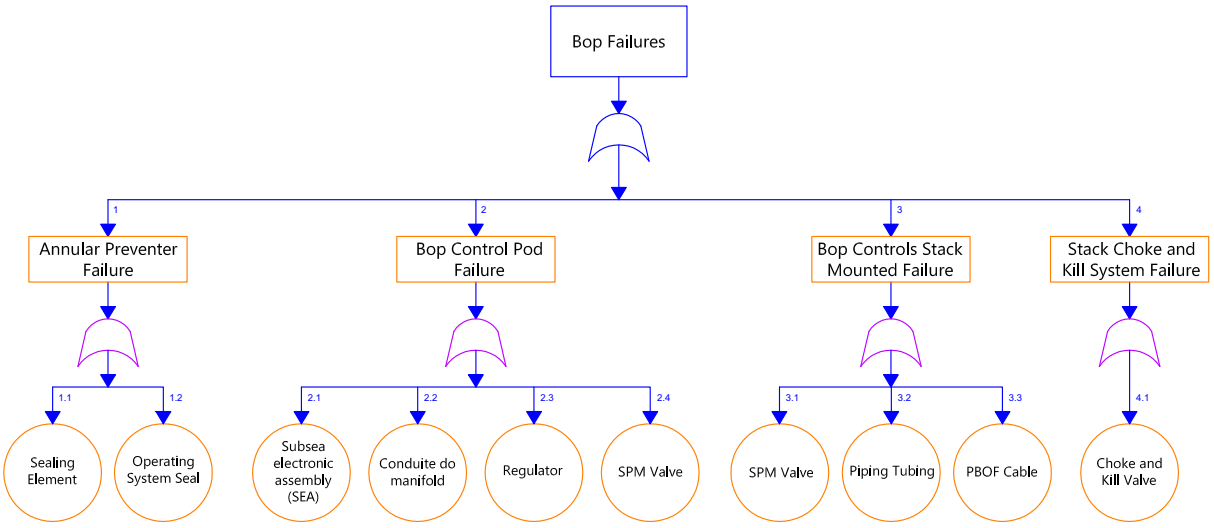


Figure 23 – FTA of BOP failure.

Table 12 – Failure data for a BOP system.

Failure Time	Failure Mode	Failure Time	Failure Mode	Failure Time	Failure Mode	Failure Time	Failure Mode
52	1.1	12	2.1	6	3.1	5,300	4.1
92	1.1	100	2.1	720	3.1	6,192	4.1
5,000	1.1	50	2.2	504	3.1		
4,320	1.2	4,320	2.2	1,800	3.1		
6,566	1.2	720	2.3	48	3.2		
		768	2.3	1,450	3.2		
		3,000	2.3	2,160	3.2		
		8,200	2.3	14,780	3.2		
		7,776	2.4	200,000	3.2		
		8,200	2.4	240	3.3		
		8,500	2.4	5,640	3.3		

Table 13 shows the CML estimates and CI 95% for the model parameters. The results presented in this table suggest that the reliability of all components improves over time since the $\widehat{\beta}_{jk}$s are less than one. Therefore, we can say that the cause-specific intensity functions of all components decrease over time.

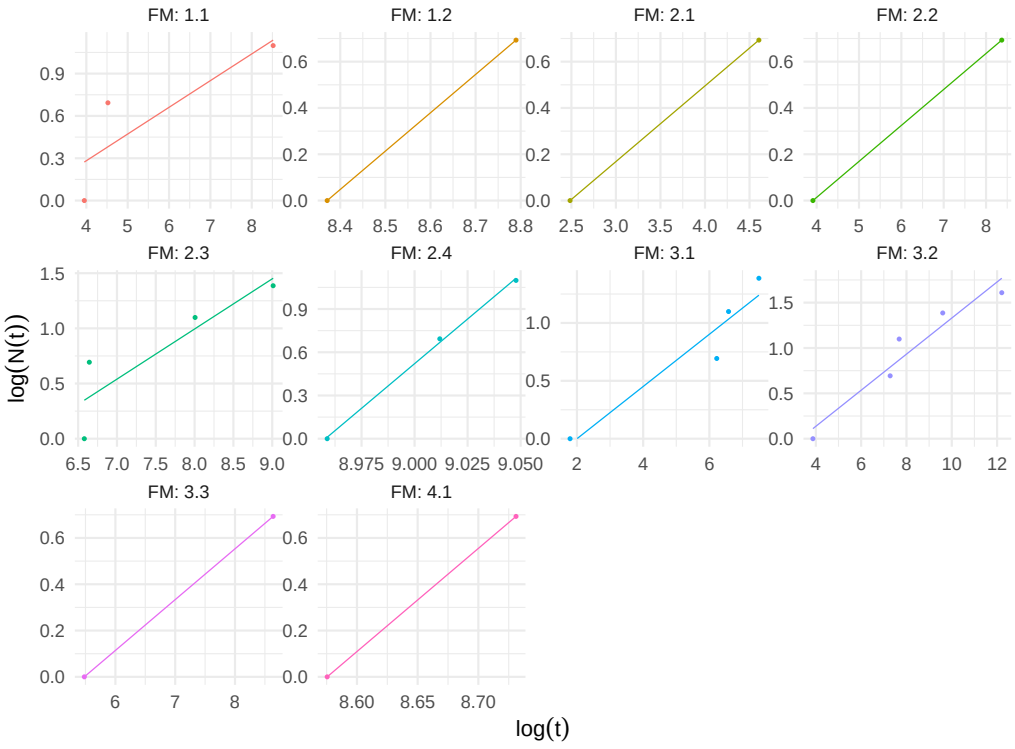


Figure 24 – Duane plots for the failure modes of a BOP system.

Table 13 – CML estimates and CI 95% of parameters β_{jk} and α_{jk}, considering the BOP failure data.

Parameter	Estimate	95% Confidence Interval	
		Lower	Upper
β_{11}	0.102	0.032	0.368
α_{11}	3.000	1.090	8.767
β_{12}	0.138	0.033	0.768
α_{12}	2.000	0.619	7.225
β_{21}	0.058	0.014	0.322
α_{21}	2.000	0.619	7.225
β_{22}	0.082	0.020	0.459
α_{22}	2.000	0.619	7.225
β_{23}	0.161	0.059	0.472
α_{23}	4.000	1.623	10.242
β_{24}	0.208	0.064	0.753
α_{24}	3.000	1.090	8.767
β_{31}	0.112	0.041	0.328
α_{31}	4.000	1.623	10.242
β_{32}	0.196	0.080	0.502
α_{32}	5.000	2.202	11.668
β_{33}	0.097	0.024	0.541
α_{33}	2.000	0.619	7.225
β_{41}	0.141	0.034	0.784
α_{41}	2.000	0.619	7.225

Table 14 – Estimates of the subsystem-specific and overall intensity functions at different times, considering the BOP failure data.

Intensity	Time (hours)				
	4	24	72	1000	9000
λ_{11}	0.0254	0.0051	0.0019	0.0002	0.0000
λ_{12}	0.0155	0.0033	0.0013	0.0001	0.0000
λ_{21}	0.0155	0.0029	0.0010	0.0001	0.0000
λ_{22}	0.0169	0.0033	0.0012	0.0001	0.0000
λ_{23}	0.0281	0.0063	0.0025	0.0003	0.0000
λ_{24}	0.0164	0.0040	0.0017	0.0002	0.0000
λ_{31}	0.0333	0.0068	0.0026	0.0002	0.0000
λ_{32}	0.0294	0.0070	0.0029	0.0003	0.0001
λ_{33}	0.0170	0.0034	0.0012	0.0001	0.0000
λ_{41}	0.0153	0.0033	0.0013	0.0001	0.0000
λ	0.2128	0.0451	0.0175	0.0018	0.0003

Table 14 displays the estimated intensity functions for each sub-system and the estimated hazard function over some fixed failure times. Observe that the results shown in this table are in agreement with the ones presented in Table 13, that is, the intensity functions decrease over time. We also see that the overall hazard function decreases over time, which may be due to the repair and maintenance effects.

The results presented in Table 13 would also allow us to estimate the reliability function of each system or sub-system, and the reliability function of the overall system. The estimated reliability functions can be obtained from (3.7), while the overall reliability function can be obtained from (3.6), with the parameters substituted by their estimates. Table 15 shows the estimated reliability functions for each sub-system and the estimated overall reliability function over some fixed failure times (which are the same times considered in Table 14). Observe that the overall reliability function, as well as the sub-systems' reliability functions, decrease over time, although at different rates.

Table 15 – Estimates of the subsystem-specific and overall reliability functions at different times, considering the BOP failure data.

Reliability	Time (hours)				
	4	24	72	1000	9000
R_{11}	0.3693	0.3025	0.2626	0.1740	0.1122
R_{12}	0.6378	0.5622	0.5117	0.3817	0.2714
R_{21}	0.3427	0.3049	0.2821	0.2292	0.1878
R_{22}	0.4406	0.3867	0.3534	0.2747	0.2125
R_{23}	0.4979	0.3940	0.3289	0.1826	0.0885
R_{24}	0.7299	0.6330	0.5627	0.3698	0.2076
R_{31}	0.3049	0.2340	0.1934	0.1100	0.0593
R_{32}	0.5494	0.4269	0.3479	0.1705	0.0658
R_{33}	0.4970	0.4352	0.3962	0.3026	0.2277
R_{41}	0.6465	0.5704	0.5193	0.3872	0.2745
R	0.0007	0.0001	0.0000	0.0000	0.0000

3.6 Concluding Remarks

In this chapter, we introduced a new statistical model for repairable systems subject to hierarchical competing risks under the assumption that the failure modes act independently. The competing risks approach may be useful in the engineering area, since it may lead to a better comprehension of the several failure modes of a system. Therefore, design strategies improve overall system reliability. The hierarchical structure may also be advantageous because sometimes it may record information about which sub-system of a specific system has resulted in the total system failure.

We assumed that the repairs are minimal, and the failure intensity follows a PLP model after a convenient reparametrization. Under a classical framework, we proposed estimators and confidence intervals for the model parameters, whose performances were investigated using a simulation study. In short, the simulation results revealed that the CMLE provide better estimates, mainly for the β_{jk} parameters, than the MLE. Besides, the exact confidence intervals for the α_{jk} parameters give CPs closer to the nominal value (0.95) than the asymptotic confidence intervals.

Finally, the proposed methodology is illustrated through a toy example on a butterfly valve system, an example of a real early-stage project related to an in-pipe robot traction system, and also a real example on a BOP system.

CHAPTER

4

IMPROVED OBJECTIVE BAYESIAN ESTIMATOR FOR A PLP MODEL HIERARCHICALLY REPRESENTED SUBJECT TO COMPETING RISKS UNDER MR REGIME

In this chapter, we propose a hierarchical statistical model for a single repairable system subject to several failure modes (competing risks). The chapter describes how complex engineered systems may be modelled hierarchically by use of Bayesian methods. It is also assumed that repairs are minimal and each failure mode has a power-law intensity. Our proposed model generalizes the one presented in Somboonsavatdee and Sen (2015) and continues the study begun in Louzada *et al.* (2019). Some properties of the new model are discussed. We conduct statistical inference under an objective Bayesian framework. A simulation study is carried out to investigate the efficiency of the proposed methods. Finally, our methodology is illustrated by two practical situations currently addressed in a project under development arising from a partnership between Petrobras and six research institutes.

Contents of Chapter

4.1 Introduction

The repeated occurrence of an event of interest in the same subject is frequent in many areas, such as manufacturing, software development, medicine, social sciences, risk analysis, among others. In reliability engineering, when studying a complex system, such as supercomputers, cars and airplanes, multiple defects or vulnerabilities in the product design, manufacture, operation, maintenance and handling can cause a number of unexpected failures (JIANG; LIU, 2017). Failure process models, in the context of repairable systems, are often described in terms of competing risks, or equivalently, a system with many components connected in series, such that the failure of a single component will result in a whole system failure. However, recently in engineering, the evaluation of repairable systems with multiple failure modes has drawn attention due to their potential applicability (ALMEIDA, 2019; MULLOR; MULERO; TROTTINI, 2019; SYAMSUNDAR; NAIKAN; WU, 2020; YAN, 2019).

In a competing risks framework, a system fails due to the first occurrence of possible failure modes. In this context, we use a model for any individual component, whose failures occur due to one of the causal mechanisms and which each one acts independently on the system.

A system can be thought as the joint union of different subsystems which, in turn, can also be thought as unions of other more particular subsystems, to an arbitrary level of undividing. In a well-defined hierarchical structure, the functions and relationships between components of a system can be better understood, highlighting their importance to the system as a whole. This makes it possible to clearly define acceptable levels of damage for each part of the structure, and to delimit its impacts on the system when exposed to different sources of external variation (LIEPING; ZHE, 2009).

In an industrial context, Langseth and Lindqvist (2006) recorded the service times of a component spanning over 1,600 units of time. Each failure had its respective mode also recorded. In this case, the causes of failure were categorized into two main groups, each with its respective subcauses. In health care, for example, Tuli *et al.* (2000) analyzed repeated shunt failures in children diagnosed with hydrocephalus; failures in this context are known to result from a variety of causes.

In complex systems with several hierarchical levels, redundancy can be implemented in any of the hierarchical levels. Finding the specific optimal configuration of a specific system is addressed by the reliability allocation problem. At the lowest level of the hierarchy, a unit can have different failure modes. Considering the modes separately might be of importance as either the consequences of the failures might be different or the maintenance actions that each

failure mode triggers might be different. In general, the failure of any single component can be considered in a competing risks framework where every failure mode is competing against the others to make the component fail in that mode.

A repairable system is defined as a system which, after failing to perform one or more of its functions satisfactorily, can be restored to fully satisfactory performance by a method other than replacement of the entire system. Traditional studies on repairable systems focus on modeling failure times, using point process theory as the main tool. In the literature, it is commonly assumed that failures in a repairable system occur due to a NHPP with the intensity described by a power-law. The resulting method is generally referred to as the PLP. The PLP is convenient in many ways, especially for its flexibility, easy implementation, and the interpretability of its parameters (ASCHER; FEINGOLD, 1984; RIGDON; BASU, 2000).

Considering the fault-causing mechanisms known, it is also important to observe how to repair such failures, including preventive maintenance. In this context, the books of Crowder (2001), Pintilie (2006), among others, illustrate with some examples the need for considering the setting of competing risks in the application of reliability techniques (in industrial statistics) or survival analysis tools (in health sciences).

Under a Bayesian perspective, the inference of a problem is on the basis of the posterior distribution of the quantity of interest, which combines the information provided by the data with the available prior information. The elicitation of an appropriate prior distribution becomes the main task for Bayesian statisticians in practice. Subjective priors, which always depend on the experts' belief, are not easy to derive in a limited time period. Therefore, given little prior information, we prefer to use objective (non-informative) priors to make inference.

An important objective prior distribution is the reference prior, introduced by Bernardo (1979) and later refined by other authors (BERGER; BERNARDO, 1989; BERGER; BERNARDO, 1992a; BERGER; BERNARDO, 1992b; BERGER, 1992). The reference prior is minimally informative in a precise theoretical sense about information. The intent is to make information from data dominate a prior information, reflecting the vague nature of a prior knowledge. To obtain such prior, the expected Kullback-Leibler divergence between a prior distribution and a posterior distribution was maximized. The posterior distribution obtained using this prior has interesting properties, such as invariance and consistency in marginalization and sample properties (BERNARDO, 2005). Some recent reference priors were obtained for the Pareto (FU; XU; TANG, 2012), Poisson-exponential (TOMAZELLA; CANCHO; LOUZADA, 2013), extended exponential-geometric (RAMOS; MOALA; ACHCAR, 2014), inverse Weibull (KIM; LEE; KANG, 2014), generalized half-normal (KANG; LEE; LEE, 2014) and Lomax (FERREIRA *et al.*, 2020) distributions.

This chapter aims to continue the study begun in Louzada *et al.* (2019), which generalized model of Somboonsavatdee and Sen (2015). For this, we describe a statistical model for a repairable system hierarchically represented subject to competing risks under

MR regime with PLP intensity function. Working under an objective Bayesian framework, we consider reference and matching priors for the model parameters. The proposed methodology is applied to two real problems arising from the development of the robotic unit Annelida, as described previously.

The next sections of the chapter are organized as follows. In Section 4.2, we introduce the proposed statistical model, which considers, in its basic assumptions, a hierarchically represented repairable system (with an arbitrary number of hierarchical levels) subject to competing risks in a MR regime governed by a PLP intensity function. In Section 4.3, under the perspective of an objective Bayesian inference framework, we derive the Maximum *a posteriori* Probability (MAP) estimators for the parameters of the proposed model, correct the biases of such estimators and expose credibility intervals with closed forms. In Section 4.4, we evaluate the properties of the estimators through a simulation study. In Section 4.5, we use the proposed model (and methods) to assess the reliability of two subsystems of the robotic unit (pressure vessels and traction system) that motivated this research, in the light of currently available information. Finally, in Section 4.6 we draw some final remarks and suggestions for future research.

4.2 Model Formulation

In this section, we introduce the proposed statistical modeling for reliability data arising from a single repairable system subject to both MR and hierarchical competing risks, whose successive failures are assumed to be governed by a PLP. Our model can be regarded as a generalization of the Somboonsavatdee and Sen (SOMBOONSAVATDEE; SEN, 2015)'s model for the cases where there are two or more levels of hierarchy, that is, secondary, tertiary, quaternary and so on failure causes (or subsystems). The model proposed here is also an extension of the work by Louzada *et al.* (LOUZADA *et al.*, 2019). This situation is illustrated in Figure 25, which depicts a fault tree. The general feature illustrated in this figure includes the composition of a system by multiple subsystems, and the composition of these subsystems by further subsystems and components.

In order to model these kinds of systems, we first assume that the failure probabilities of components in distinct branches of the fault tree are conditionally independent and that success of the systems requires successful functioning of all components.

Let us suppose a repairable system with p levels of hierarchy. Then, the hierarchical competing risks model's data consist of the $(p+1)$-tuples $(t, \delta_1, \delta_2, \ldots, \delta_p)$, where $t > 0$ denotes the failure time, δ_1 is the indicator of the leading failure cause (system), and δ_j is the indicator of the subcause (subsystem) at hierarchical level j, for $j = 2, \ldots, p$. Let $N_p(\delta_p)$ be the counting process associated with the system failure of type p. To consider the natural hierarchy of this model, the δ_2 indicator, for example, refers to the cause in the second hierarchical level, which

is nested with a specific cause of the first hierarchical level, represented by δ_1. In this sense, we say that $\delta_1 = 1,\ldots,N_1$, that is, there are N_1 primary causes of system failure. On the other hand, $\delta_2 = 1,\ldots,N_2(\delta_1)$, that is, the number of causes $N_2(\delta_1)$ (denoted later simply by N_2, for simplicity) closely depends on the primary cause δ_1. This logic extends to the p-th cause, indicated by $\delta_p = 1,\ldots,N_p(\delta_1,\ldots,\delta_{p-1})$.

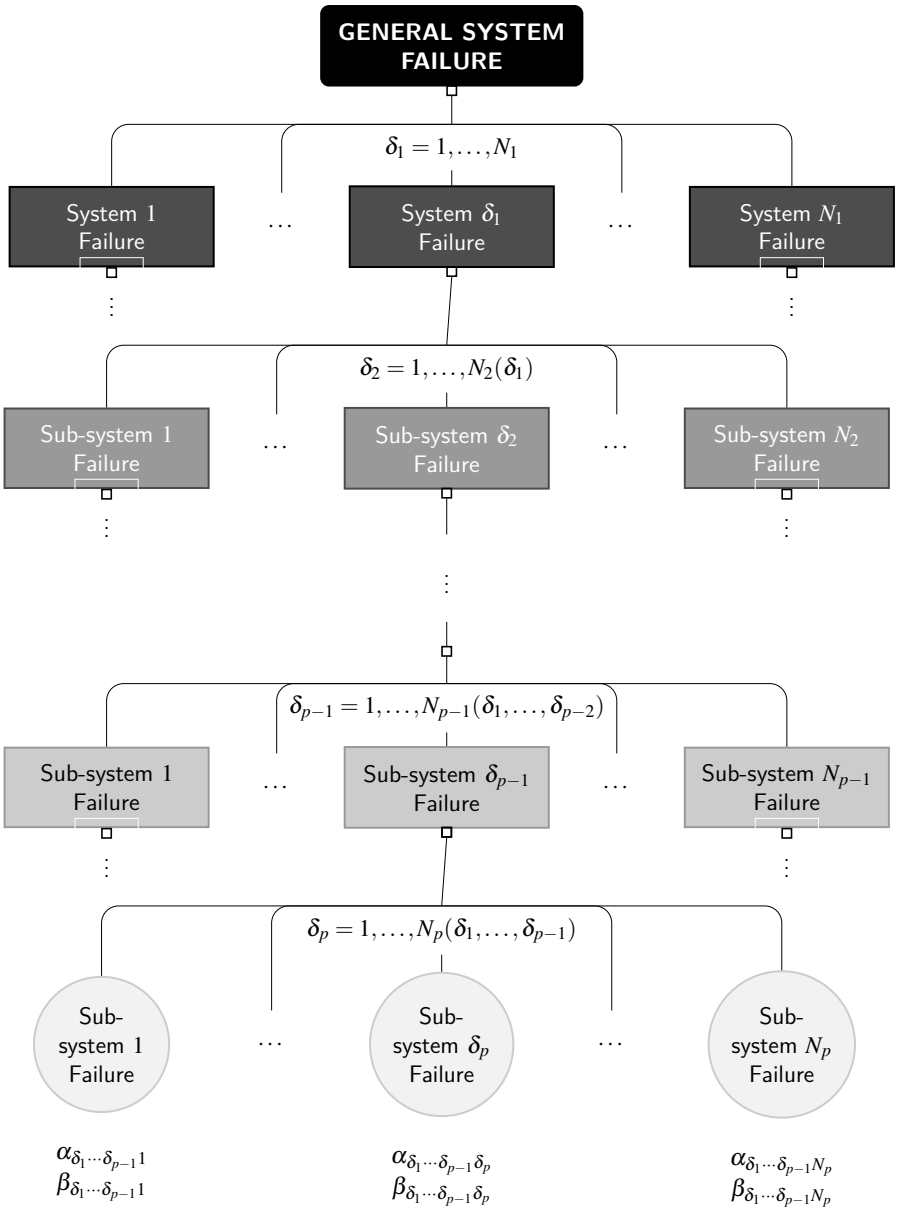


Figure 25 – FTA of the general system with hierarchical failure modes (p levels of nesting).

Our proposed model for failure data analysis can be formulated as follows. Let us assume that the failures from a hierarchical (at level p) subsystem follow a NHPP with the PLP intensity function given by

$$\lambda_{\delta_1\cdots\delta_p}(t) = \left(\frac{\beta_{\delta_1\cdots\delta_p}}{\mu_{\delta_1\cdots\delta_p}}\right)\left(\frac{t}{\mu_{\delta_1\cdots\delta_p}}\right)^{\beta_{\delta_1\cdots\delta_p}-1},$$

where $\mu_{\delta_1\cdots\delta_p} > 0$ and $\beta_{\delta_1\cdots\delta_p} > 0$ are, respectively, the scale and shape parameters. Or equivalently, $\mu_{\delta_1\cdots\delta_p}$ represents the time for which we expect to observe a single event, and $\beta_{\delta_1\cdots\delta_p}$ is the elasticity of the mean number of events with regard to time (OLIVEIRA; COLOSIMO; GILARDONI, 2012).

Thus, it follows that the overall intensity function at time t is given by

$$\begin{aligned}\lambda(t) &= \sum_{\delta_1=1}^{N_1}\cdots\sum_{\delta_p=1}^{N_p}\lambda_{\delta_1\cdots\delta_p}(t)\\ &= \sum_{\delta_1=1}^{N_1}\cdots\sum_{\delta_p=1}^{N_p}\left(\frac{\beta_{\delta_1\cdots\delta_p}}{\mu_{\delta_1\cdots\delta_p}}\right)\left(\frac{t}{\mu_{\delta_1\cdots\delta_p}}\right)^{\beta_{\delta_1\cdots\delta_p}-1},\end{aligned} \tag{4.1}$$

where N_j denotes the number of components in the $(\delta_1,\ldots,\delta_j)$-th hierarchical subsystem, for $j = 1,\ldots,p$.

Let us assume that $n \geq 1$ failures have occurred in the time interval $(0,T]$. Then, the hierarchical (at level p) subsystem-specific cumulative intensity up to time T becomes

$$\Lambda_{\delta_1\cdots\delta_p}(T) = \left(\frac{T}{\mu_{\delta_1\cdots\delta_p}}\right)^{\beta_{\delta_1\cdots\delta_p}}.$$

From (4.1), it follows that

$$\Lambda(T) = \sum_{\delta_1=1}^{N_1}\cdots\sum_{\delta_p=1}^{N_p}\Lambda_{\delta_1\cdots\delta_p}(T)$$

is the overall cumulative intensity up to time T. Hence, we have that the overall reliability up to time T is

$$\begin{aligned}R(T) &= \exp\{-\Lambda(T)\}\\ &= \exp\left\{-\sum_{\delta_1=1}^{N_1}\cdots\sum_{\delta_p=1}^{N_p}\Lambda_{\delta_1\cdots\delta_p}(T)\right\}\\ &= \exp\left\{-\sum_{\delta_1=1}^{N_1}\cdots\sum_{\delta_p=1}^{N_p}\left(\frac{T}{\mu_{\delta_1\cdots\delta_p}}\right)^{\beta_{\delta_1\cdots\delta_p}}\right\},\end{aligned}$$

while the hierarchical (at level p) subsystem-specific reliability up to time T is given by

$$\begin{aligned}R_{\delta_1\cdots\delta_p}(T) &= \exp\left\{-\Lambda_{\delta_1\cdots\delta_p}(T)\right\}\\ &= \exp\left\{-\left(\frac{T}{\mu_{\delta_1\cdots\delta_p}}\right)^{\beta_{\delta_1\cdots\delta_p}}\right\}.\end{aligned}$$

As suggested by Oliveira, Colosimo and Gilardoni (2012), we will reparametrize our proposed model in terms of $\beta_{\delta_1 \cdots \delta_p}$ and

$$\alpha_{\delta_1 \cdots \delta_p} = \mathbb{E}[N_{\delta_1 \cdots \delta_p}(T)] = \left(\frac{T}{\mu_{\delta_1 \cdots \delta_p}} \right)^{\beta_{\delta_1 \cdots \delta_p}},$$

where $N_{\delta_1 \cdots \delta_p}(T)$ is the hierarchical (at level p) subsystem-specific counting process, which denotes the number of failures before time T. It follows that $\beta_{\delta_1 \cdots \delta_p}$ and $\alpha_{\delta_1 \cdots \delta_p}$ are orthogonal parameters.

The orthogonal reparametrization of the PLP model enables us to obtain a likelihood function whose parameters $\beta_{\delta_1 \cdots \delta_p}$ and $\alpha_{\delta_1 \cdots \delta_p}$ are independent with desirable properties. In this case, the hierarchical (at level p) subsystem-specific likelihood function is given by

$$\begin{aligned} L(\beta_{\delta_1 \cdots \delta_p}, \alpha_{\delta_1 \cdots \delta_p} \mid n, \boldsymbol{t}) &= c \beta_{\delta_1 \cdots \delta_p}^{n} e^{-n\beta_{\delta_1 \cdots \delta_p}/\hat{\beta}_{\delta_1 \cdots \delta_p}} \\ &\quad \times \alpha_{\delta_1 \cdots \delta_p}^{n} e^{-\alpha_{\delta_1 \cdots \delta_p}} \\ &\propto \gamma\left(\beta_{\delta_1 \cdots \delta_p} \mid n+1, n/\hat{\beta}_{\delta_1 \cdots \delta_p}\right) \\ &\quad \times \gamma\left(\alpha_{\delta_1 \cdots \delta_p} \mid n+1, 1\right), \end{aligned}$$

where $c = \prod_{i=1}^{n} t_i^{-1}$ and $\gamma(x \mid a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}$, for $x, a, b > 0$, is the probability density function of a gamma distribution with shape parameter a and scale parameter b. Moreover, $\hat{\beta}_{\delta_1 \cdots \delta_p}$ is the (biased) MLE of $\beta_{\delta_1 \cdots \delta_p}$, which is given by [1]

$$\hat{\beta}_{\delta_1 \cdots \delta_p} = \frac{n}{\sum_{i=1}^{n} \log\left(\frac{T}{t_i}\right)}.$$

For a further discussion on the advantages of having orthogonal parameters, see Cox and Reid (COX; REID, 1987).

4.3 Bayesian Inference

In this chapter, we investigate the repairable system in the presence of hierarchical competing risks via objective Bayesian approach. A non-informative prior is used to depict lack of prior knowledge about the quantity of interest. There are different ways to obtain objective priors for the parameters of our model. Although the Jeffreys prior is the most commonly used, this prior may not be adequate in multivariate case (BERNARDO, 2005). Tibshirani (1989) proposed an alternative method to derive a class of objective priors $\pi(\theta_1, \theta_2)$, where θ_1 is the parameter of interest so that the credible interval for θ_1 has a coverage error $O\left(n^{-1}\right)$ in the

[1] It is worth noting that n, t_i and $\boldsymbol{t}$ should also carry $\delta_1 \ldots \delta_p$ as a subscript (i.e., $n_{\delta_1 \ldots \delta_p}$, $t_{i;\delta_1 \ldots \delta_p}$ and $\boldsymbol{t}_{\delta_1 \ldots \delta_p}$), but we omit it so as not to clutter the notation.

frequentist sense, i.e.,

$$P\left[\theta_1 \leq \theta_1^{1-\xi}(\pi;\boldsymbol{t}) \mid (\theta_1,\theta_2)\right] = 1-\xi-O\left(n^{-1}\right), \tag{4.2}$$

where $\theta_1^{1-\xi}(\pi;\boldsymbol{t}) \mid (\theta_1,\theta_2)$ denotes the $(1-\xi)$-th quantile of the posterior distribution of θ_1. The priors that satisfy (4.2) are known as matching priors.

To obtain such priors, Tibshirani (1989) showed that if θ_1 and θ_2 are orthogonal parameters in the sense discussed by Cox and Reid (1987), i.e., $I_{\theta_1,\theta_2}(\theta_1,\theta_2)=0$, where θ_1 is the parameter of interest and θ_2 is the orthogonal nuisance parameter, then the matching priors are all priors of the form

$$\pi(\theta_1,\theta_2) = g(\theta_2)\sqrt{I_{\theta_1,\theta_1}(\theta_1,\theta_2)}, \tag{4.3}$$

where $g(\theta_2)>0$ is an arbitrary function and $I_{\theta_1,\theta_1}(\theta_1,\theta_2)$ is the θ_1 entry of the Fisher information matrix. Tibshirani (1989) showed that (4.3) is also a matching prior when θ_2 is a vector of nuisance parameters.

Considering the proposed model, and assuming that

$$\delta_1 = \{1,\ldots,N_1\},\ldots,\delta_p = \left\{1,\ldots,N_p\left(\delta_1,\ldots,\delta_{p-1}\right)\right\},$$

the elements of the Fisher information matrix can be expressed as

$$\begin{aligned}
I_{\beta_{\delta_1\cdots\delta_p},\beta_{\delta_1\cdots\delta_p}}\left(\beta_{\delta_1\cdots\delta_p},\alpha_{\delta_1\cdots\delta_p}\right) &= \alpha_{\delta_1\cdots\delta_p}\beta^{-2}_{\delta_1\cdots\delta_p},\\
I_{\alpha_{\delta_1\cdots\delta_p},\alpha_{\delta_1\cdots\delta_p}}\left(\beta_{\delta_1\cdots\delta_p},\alpha_{\delta_1\cdots\delta_p}\right) &= \alpha^{-1}_{\delta_1\cdots\delta_p},\\
I_{\beta_{\delta_1\cdots\delta_p},\alpha_{\delta_1\cdots\delta_p}}\left(\beta_{\delta_1\cdots\delta_p},\alpha_{\delta_1\cdots\delta_p}\right) &= 0,\\
I_{\alpha_{\delta_1\cdots\delta_p},\beta_{\delta_1\cdots\delta_p}}\left(\beta_{\delta_1\cdots\delta_p},\alpha_{\delta_1\cdots\delta_p}\right) &= 0.
\end{aligned}$$

From (4.3), one of the possible solutions is given by

$$\pi\left(\beta_{\delta_1\cdots\delta_p},\alpha_{\delta_1\cdots\delta_p}\right) = \prod_{\delta_1=1}^{N_1}\cdots\prod_{\delta_p=1}^{N_p}\frac{1}{\beta_{\delta_1\cdots\delta_p}\sqrt{\alpha_{\delta_1\cdots\delta_p}}}. \tag{4.4}$$

The prior given above satisfies (4.3) for all $\beta_{\delta_1\cdots\delta_p}$ and $\alpha_{\delta_1\cdots\delta_p}$ selected as interested parameters. Hence, the obtained prior is a matching prior for all the parameters, which implies that the credibility interval for any parameter has a coverage error $O\left(n^{-1}\right)$ in the frequentist sense.

Another important objective prior is the reference prior introduced by Bernardo (1979) with further developments by Berger and Bernardo (1992b), Berger (1992). This prior is defined as the prior that maximizes the expected Kullback-Leibler distance between the posterior distribution and the prior distribution based on the experimental data. Bernardo

(2005) proved that the reference prior has desirable properties, such as invariance, consistency under marginalization and consistent sampling properties. If the parameters of the model are orthogonal, the following lemma (see Berger *et al.* (2015)) can be used to easily obtain a one-at-a-time reference prior to any chosen parameter of interest and any ordering of the nuisance parameters (hereafter referred to as overall reference prior).

Consider the unknown parameters $\boldsymbol{\theta} = (\theta_1, \theta_2)$ with associated Fisher information matrix $I(\theta_1, \theta_2)$. If $I(\theta_1, \theta_2)$ is of the form

$$I(\theta_1, \theta_2) = \text{Diag}\left(f(\theta_1) g(\theta_2), h(\theta_2) w(\theta_1)\right),$$

where f, g, h and w are positive functions of $\boldsymbol{\theta}$, then the overall reference prior is given by

$$\pi_R(\theta_1, \theta_2) = \sqrt{f(\theta_1) h(\theta_2)}. \tag{4.5}$$

Assuming that $\boldsymbol{\theta} \in \mathbb{R}^k$, where k is the number of parameters, the same approach can be applied to obtain the overall reference prior related to the vector of parameters. Here, we have that $f_{\delta_1 \cdots \delta_p}\left(\beta_{\delta_1 \cdots \delta_p}\right) = \beta^{-2}_{\delta_1 \cdots \delta_p}$ and $h_{\delta_1 \cdots \delta_p}\left(\beta_{\delta_1 \cdots \delta_p}\right) = \alpha^{-1}_{\delta_1 \cdots \delta_p}$. Hence, from (4.5), the overall reference prior is given by

$$\pi_R\left(\beta_{\delta_1 \cdots \delta_p}, \alpha_{\delta_1 \cdots \delta_p}\right) = \prod_{\delta_1=1}^{N_1} \cdots \prod_{\delta_p=1}^{N_p} \frac{1}{\beta_{\delta_1 \cdots \delta_p} \sqrt{\alpha_{\delta_1 \cdots \delta_p}}}.$$

Therefore, the prior (4.4) is an overall reference prior and also a matching prior for all the parameters. The obtained posterior distribution is given by

$$\pi_R\left(\beta_{\delta_1 \cdots \delta_p}, \alpha_{\delta_1 \cdots \delta_p} \mid n_{\delta_1 \cdots \delta_p}, \boldsymbol{t}\right) \propto \prod_{\delta_1=1}^{N_1} \cdots \prod_{\delta_p=1}^{N_p} \gamma\left(\beta_{\delta_1 \cdots \delta_p} \mid n_{\delta_1 \cdots \delta_p}, n_{\delta_1 \cdots \delta_p} / \hat{\beta}_{\delta_1 \cdots \delta_p}\right) \gamma\left(\alpha_{\delta_1 \cdots \delta_p} \mid n_{\delta_1 \cdots \delta_p} + \frac{1}{2}, 1\right).$$

Due to the consistent marginalization property of the overall reference prior, the marginal reference posteriors are given by

$$\pi_R(\beta_{\delta_1 \cdots \delta_p} \mid n_{\delta_1 \cdots \delta_p}, \boldsymbol{t}) \propto \gamma\left(\beta_{\delta_1 \cdots \delta_p} \mid n_{\delta_1 \cdots \delta_p}, n_{\delta_1 \cdots \delta_p} / \hat{\beta}_{\delta_1 \cdots \delta_p}\right),$$

and

$$\pi_R(\alpha_{\delta_1 \cdots \delta_p} \mid n_{\delta_1 \cdots \delta_p}, \boldsymbol{t}) \propto \gamma\left(\alpha_{\delta_1 \cdots \delta_p} \mid n_{\delta_1 \cdots \delta_p} + \frac{1}{2}, 1\right).$$

From the marginal posterior distribution, we can obtain the Bayes estimator assuming some rule, such as the posterior mean, median or mode. Here, we assume the posterior mode, also known as MAP estimator, since this approach leads to an unbiased estimator for $\beta_{\delta_1 \cdots \delta_p}$ in the frequentist sense. The MAP estimator for $\beta_{\delta_1 \cdots \delta_p}$ is given by

$$\hat{\beta}^B_{\delta_1 \cdots \delta_p} = \frac{n_{\delta_1 \cdots \delta_p} - 1}{\sum_{i=1}^n \log\left(\frac{T}{t_i}\right) \mathbb{I}(\delta_1 \cdots \delta_p)},$$

where $\mathbb{I}(\delta_1 \cdots \delta_p)$ is the indicator function that equals one if the observation t_i belongs to the subsystem $\delta_1 \cdots \delta_p$. From the estimator above, we have that $\mathbb{E}\left[\hat{\beta}^B_{\delta_1 \cdots \delta_p}\right] = \beta_{\delta_1 \cdots \delta_p}$, i.e., such a Bayes estimator is unbiased in the frequentist sense. On the other hand, for $\alpha_{\delta_1 \cdots \delta_p}$, the MAP estimator is given by

$$\hat{\alpha}^B_{\delta_1 \cdots \delta_p} = n_{\delta_1 \cdots \delta_p} - \frac{1}{2}.$$

In this case, we have that

$$\mathbb{E}\left[\hat{\alpha}^B_{\delta_1 \cdots \delta_p}\right] = \alpha_{\delta_1 \cdots \delta_p} - \frac{1}{2}.$$

Hence, such a Bayes estimator for $\alpha_{\delta_1 \cdots \delta_p}$ has a systematic bias of -0.5. Once we have identified the bias, we can remove it. In this case, we will not have that MAP estimator, but a Bias-Corrected MAP (BMAP) estimator. Hereafter, we will consider the BMAP estimator for the model parameters, which will be computed by

$$\hat{\beta}^{BC}_{\delta_1 \cdots \delta_p} = \frac{n_{\delta_1 \cdots \delta_p} - 1}{\sum_{i=1}^{n} \log\left(\frac{T}{t_i}\right) \mathbb{I}(\delta_1 \cdots \delta_p)},$$

and

$$\hat{\alpha}^{BC}_{\delta_1 \cdots \delta_p} = n_{\delta_1 \cdots \delta_p}.$$

Now, since the marginal posterior distributions have closed-form expressions, we have that the $\upsilon = 100(1-\xi)\%$ credibility intervals for $\beta_{\delta_1 \cdots \delta_p}$ and $\alpha_{\delta_1 \cdots \delta_p}$ can be obtained directly from the quantile function of the gamma distribution, that is,

$$CI(\beta_\circ; \upsilon) = \left[\gamma_Q\left(n_\circ, \frac{n_\circ}{\hat{\beta}_\circ}; \frac{\xi}{2}\right); \gamma_Q\left(n_\circ, \frac{n_\circ}{\hat{\beta}_\circ}; 1 - \frac{\xi}{2}\right)\right]$$

and

$$CI(\alpha_\circ; \upsilon) = \left[\gamma_Q\left(n_\circ + \frac{1}{2}, 1; \frac{\xi}{2}\right); \gamma_Q\left(n_\circ + \frac{1}{2}, 1; 1 - \frac{\xi}{2}\right)\right],$$

where $\circ$ denotes the index $\delta_1 \cdots \delta_p$ and $\gamma_Q(a, b; \upsilon)$ is the quantile function of the gamma distribution with shape parameter a and scale parameter b, and $0 \leq \upsilon \leq 1$. This quantile function is available in most of the standard statistical softwares. For example, in R it can be computed by using the `qgamma` function. Therefore, the exact confidence intervals for the model parameters can be obtained without the use of intensive computation.

4.4 Simulation

In this section, we carry out a simulation study to investigate and compare the performance of the proposed Bayes estimators. To evaluate the estimators' behavior, two metrics are used: the Mean Relative Estimate (MRE) and the Root Mean Squared Error (RMSE), which are calculated, respectively, by

$$MRE(\hat{\theta}_w) = \frac{1}{M}\sum_{m=1}^{M} \frac{\hat{\theta}_w^{(m)}}{\theta_w} \qquad \text{and} \qquad RMSE(\hat{\theta}_w) = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(\hat{\theta}_w^{(m)} - \theta_w\right)^2},$$

for $w = 1, \ldots, \kappa$, where $M = 100{,}000$ is the number of Monte Carlo simulations and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_\kappa) = \left(\beta_{1\cdots 1}, \ldots, \beta_{N_1 \cdots N_p}, \alpha_{1\cdots 1}, \ldots, \alpha_{N_1 \cdots N_p}\right)$ denotes the parameter vector. Besides, $\hat{\theta}_w^{(m)}$ represents the estimate of θ_w obtained from sample m, for $m = 1, \ldots, M$.

Through this approach, it is expected that good estimators return MREs close to one and RMSEs close to zero. On the other hand, the 90% credibility intervals, which are obtained directly from the 5% and 95% quantiles of the gamma posterior distributions, are expected to have CP near the nominal value of 90%.

By considering the well-known results regarding NHPPs (RIGDON; BASU, 2000), and also from the assumption that the failure modes are independent, we can generate the failure times, for each Monte Carlo replication, according to the steps described in Algorithm 4. All numerical computations and simulations were done using the R programming language (R Core Team, 2018). Due to space constraints, the results are reported only for six scenarios. However, similar findings are obtained for other parameter choices.

Input:

$p, T, N_1, N_2(\delta_1), \ldots, N_p(\delta_1, \ldots, \delta_{p-1}), \boldsymbol{\beta} = (\beta_{1\cdots 1}, \ldots, \beta_{N_1 \cdots N_p}), \boldsymbol{\alpha} = (\alpha_{1\cdots 1}, \ldots, \alpha_{N_1 \cdots N_p})$

Output:

$\{(\boldsymbol{t}, \boldsymbol{\psi})\} = \{(\boldsymbol{t}_{1\ldots 1}, \boldsymbol{\psi}_{1\ldots 1}), \ldots, (\boldsymbol{t}_{N_1 \ldots N_p}, \boldsymbol{\psi}_{N_1 \ldots N_p})\}$

Procedure:

for $\delta_1 := 1$ **to** N_1 **do**
  for $\delta_2 := 1$ **to** $N_2(\delta_1)$ **do**
    $\vdots$
    for $\delta_p := 1$ **to** $N_p(\delta_1, \ldots, \delta_{p-1})$ **do**
      $n_{\delta_1,\ldots,\delta_p} \sim \text{Poisson}(\alpha_{\delta_1,\ldots,\delta_p})$
      for $i := 1$ **to** $n_{\delta_1,\ldots,\delta_p}$ **do**
        $U_{i,\delta_1,\ldots,\delta_p} \sim \text{Uniform}(0,1)$
        $t_{i,\delta_1,\ldots,\delta_p} = T\, U_{i,\delta_1,\ldots,\delta_p}^{1/\beta_{\delta_1,\ldots,\delta_p}}$
        $\psi_{i,\delta_1,\ldots,\delta_p} = \delta_1, \ldots, \delta_p$
      end
    end
    $\vdots$
  end
end

Algorithm 2: Generator of random numbers from the proposed model.

In what follows, we present the results for two distinct structures of a single system,

both under the assumption that the components system is observed in the fixed time interval $(0, T]$, where $T = 60$. The first is a system subject to 3 failure causes each with 2 subcauses; and the second is a system subject to 2 main causes each one subject to 2 subcauses which, in turn, are subject to 2 other causes which, in the end, are also subject to 2 causes, in a 4-level structure. These structures can be seen in Figures 57 and 58, in which the parameters set for each underlying cause is presented in Table 16.

Table 16 – Scenarios for a single system subject to 3 failure causes each one with 2 subcauses (Scenarios 1, 2 and 3); and a single system subject to 2 failure causes each one with 2 subcauses which, in turn, have 2 causes of failure and, finally, also 2 other causes of failure (Scenarios 4, 5 and 6).

System	β	α	System	β	α	System	β	α
Scenario 1			Scenario 2			Scenario 3		
1.1	1.5	22	1.1	2.2	30	1.1	0.20	35
1.2	2.3	18	1.2	0.7	25	1.2	0.50	24
2.1	1.2	20	2.1	1.5	27	2.1	0.90	18
2.2	3.5	25	2.2	0.2	32	2.2	1.00	20
3.1	1.0	21	3.1	1.0	23	3.1	0.80	32
3.2	1.1	23	3.2	1.8	20	3.2	0.05	25
Scenario 4			Scenario 5			Scenario 6		
1.1.1.1	1.3	20	1.1.1.1	1.0	20	1.1.1.1	0.10	20
1.1.1.2	1.5	22	1.1.1.2	1.2	25	1.1.1.2	0.10	22
1.1.2.1	2.0	23	1.1.2.1	0.8	26	1.1.2.1	0.50	30
1.1.2.2	1.8	25	1.1.2.2	1.5	23	1.1.2.2	0.30	28
1.2.1.1	1.1	22	1.2.1.1	0.3	18	1.2.1.1	0.90	18
1.2.1.2	2.5	21	1.2.1.2	1.8	17	1.2.1.2	0.80	25
1.2.2.1	1.2	30	1.2.2.1	0.7	23	1.2.2.1	0.05	19
1.2.2.2	1.5	32	1.2.2.2	1.5	22	1.2.2.2	0.40	17
2.1.1.1	1.0	17	2.1.1.1	0.5	20	2.1.1.1	0.20	20
2.1.1.2	1.1	18	2.1.1.2	0.9	17	2.1.1.2	0.01	26
2.1.2.1	1.0	20	2.1.2.1	2.0	28	2.1.2.1	0.30	32
2.1.2.2	2.1	18	2.1.2.2	1.0	22	2.1.2.2	0.70	19
2.2.1.1	1.6	29	2.2.1.1	0.4	19	2.2.1.1	0.40	23
2.2.1.2	1.5	25	2.2.1.2	0.1	23	2.2.1.2	0.90	25
2.2.2.1	2.5	24	2.2.2.1	1.4	20	2.2.2.1	0.10	26
2.2.2.2	3.0	23	2.2.2.2	1.9	18	2.2.2.2	0.20	20

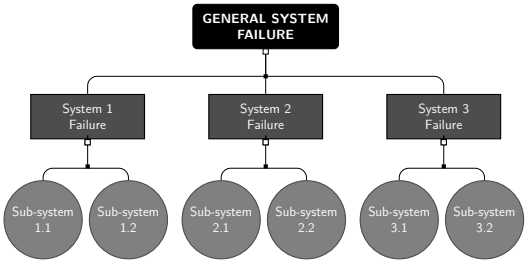


Figure 26 – Simulated failure structure for Scenarios 1, 2 and 3.

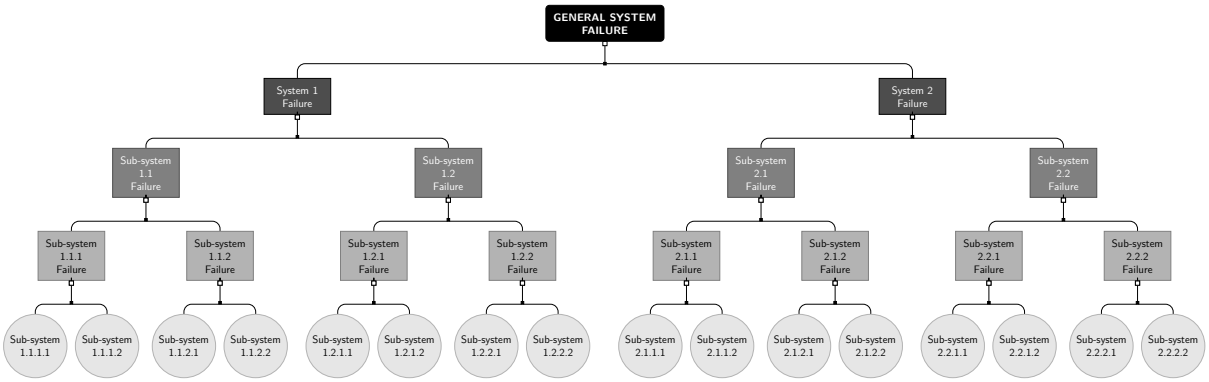


Figure 27 – Simulated failure structure for Scenarios 4, 5 and 6.

As can be seen in Tables 17 and 18, the MREs are very close to one, with no exception, especially for the BMAP estimator. On the other hand, the observed values of RMSE are, in general, less than 0.5 for the MAP/BMAP estimator of $\boldsymbol{\beta}$ and less than 5 in the case of $\boldsymbol{\alpha}$.

The CPs are close to the nominal value of 90%, especially in the case of the exact intervals (CB) when compared to the asymptotic ones (B).

Table 17 – MREs, RMSEs and CPs from the MAP (B) and BMAP (CB) estimators, considering different parameter values (Scenarios 1, 2 and 3).

Parameter	Scenario 1							Scenario 2							Scenario 3						
	Value	MRE		RMSE		CP		Value	MRE		RMSE		CP		Value	MRE		RMSE		CP	
		B	CB	B	CB	B	CB		B	CB	B	CB	B	CB		B	CB	B	CB	B	CB
$\alpha_{1.1}$	22.0	0.977	1.000	4.719	4.693	0.856	0.892	30.0	0.983	1.000	5.505	5.483	0.887	0.899	35.00	0.986	1.000	5.942	5.921	0.871	0.892
$\beta_{1.1}$	1.5	1.000	1.000	0.348	0.348	0.885	0.899	2.2	1.000	1.000	0.427	0.427	0.888	0.898	0.20	0.999	0.999	0.035	0.035	0.890	0.899
$\alpha_{1.2}$	18.0	0.972	1.000	4.276	4.247	0.880	0.900	25.0	0.980	1.000	5.027	5.002	0.886	0.889	24.00	0.979	1.000	4.927	4.902	0.892	0.897
$\beta_{1.2}$	2.3	1.000	1.000	0.599	0.599	0.884	0.900	0.7	1.001	1.001	0.151	0.151	0.887	0.900	0.50	1.000	1.000	0.110	0.110	0.886	0.899
$\alpha_{2.1}$	20.0	0.975	1.000	4.504	4.476	0.873	0.908	27.0	0.981	1.000	5.227	5.203	0.872	0.899	18.00	0.972	1.000	4.276	4.247	0.880	0.900
$\beta_{2.1}$	1.2	1.000	1.000	0.294	0.294	0.885	0.898	1.5	1.001	1.001	0.309	0.309	0.888	0.899	0.90	1.000	1.000	0.234	0.234	0.884	0.900
$\alpha_{2.2}$	25.0	0.980	1.000	5.027	5.002	0.886	0.889	32.0	0.984	1.000	5.683	5.661	0.886	0.908	20.00	0.975	1.000	4.504	4.476	0.873	0.908
$\beta_{2.2}$	3.5	1.001	1.001	0.753	0.753	0.887	0.900	0.2	1.000	1.000	0.037	0.037	0.889	0.898	1.00	1.000	1.000	0.245	0.245	0.885	0.898
$\alpha_{3.1}$	21.0	0.976	1.000	4.616	4.589	0.864	0.900	23.0	0.978	1.000	4.824	4.798	0.889	0.904	32.00	0.984	1.000	5.683	5.661	0.886	0.908
$\beta_{3.1}$	1.0	1.000	1.000	0.238	0.238	0.885	0.898	1.0	1.000	1.000	0.226	0.226	0.886	0.899	0.80	1.000	1.000	0.150	0.150	0.889	0.898
$\alpha_{3.2}$	23.0	0.978	1.000	4.824	4.798	0.889	0.904	20.0	0.975	1.000	4.504	4.476	0.873	0.908	25.00	0.980	1.000	5.027	5.002	0.886	0.889
$\beta_{3.2}$	1.1	1.000	1.000	0.248	0.248	0.886	0.899	1.8	1.000	1.000	0.440	0.440	0.885	0.898	0.05	1.001	1.001	0.011	0.011	0.887	0.900

Table 18 – MREs, RMSEs and CPs from the MAP (B) and BMAP (CB) estimators, considering different parameter values (Scenarios 4, 5 and 6).

Parameter	Scenario 4							Scenario 5							Scenario 6						
	Value	MRE		RMSE		CP		Value	MRE		RMSE		CP		Value	MRE		RMSE		CP	
		B	CB	B	CB	B	CB		B	CB	B	CB	B	CB		B	CB	B	CB	B	CB
$\alpha_{1.1.1.1}$	20.0	0.975	1.000	4.504	4.476	0.873	0.908	20.0	0.975	1.000	4.504	4.476	0.873	0.908	20.00	0.975	1.000	4.504	4.476	0.873	0.908
$\beta_{1.1.1.1}$	1.3	1.000	1.000	0.318	0.318	0.885	0.898	1.0	1.000	1.000	0.245	0.245	0.885	0.898	0.10	1.000	1.000	0.024	0.024	0.885	0.898
$\alpha_{1.1.1.2}$	22.0	0.977	1.000	4.719	4.693	0.856	0.892	25.0	0.980	1.000	5.027	5.002	0.886	0.889	22.00	0.977	1.000	4.719	4.693	0.856	0.892
$\beta_{1.1.1.2}$	1.5	1.000	1.000	0.348	0.348	0.885	0.899	1.2	1.001	1.001	0.258	0.258	0.887	0.900	0.10	1.000	1.000	0.023	0.023	0.885	0.899
$\alpha_{1.1.2.1}$	23.0	0.978	1.000	4.824	4.798	0.889	0.904	26.0	0.981	1.000	5.124	5.100	0.879	0.906	30.00	0.983	1.000	5.505	5.483	0.887	0.899
$\beta_{1.1.2.1}$	2.0	1.000	1.000	0.451	0.451	0.886	0.899	0.8	1.001	1.001	0.169	0.169	0.887	0.899	0.50	1.000	1.000	0.097	0.097	0.888	0.898
$\alpha_{1.1.2.2}$	25.0	0.980	1.000	5.027	5.002	0.886	0.889	23.0	0.978	1.000	4.824	4.798	0.889	0.904	28.00	0.982	1.000	5.321	5.297	0.865	0.893
$\beta_{1.1.2.2}$	1.8	1.001	1.001	0.387	0.387	0.887	0.900	1.5	1.000	1.000	0.338	0.338	0.886	0.899	0.30	1.000	1.000	0.061	0.061	0.888	0.899
$\alpha_{1.2.1.1}$	22.0	0.977	1.000	4.719	4.693	0.856	0.892	18.0	0.972	1.000	4.276	4.247	0.880	0.900	18.00	0.972	1.000	4.276	4.247	0.880	0.900
$\beta_{1.2.1.1}$	1.1	1.000	1.000	0.255	0.255	0.885	0.899	0.3	1.000	1.000	0.078	0.078	0.884	0.900	0.90	1.000	1.000	0.234	0.234	0.884	0.900
$\alpha_{1.2.1.2}$	21.0	0.976	1.000	4.616	4.589	0.864	0.900	17.0	0.971	1.000	4.155	4.125	0.890	0.910	25.00	0.980	1.000	5.027	5.002	0.886	0.889
$\beta_{1.2.1.2}$	2.5	1.000	1.000	0.596	0.596	0.885	0.898	1.8	1.000	1.000	0.486	0.486	0.882	0.899	0.80	1.001	1.001	0.172	0.172	0.887	0.900
$\alpha_{1.2.2.1}$	30.0	0.983	1.000	5.505	5.483	0.887	0.899	23.0	0.978	1.000	4.824	4.798	0.889	0.904	19.00	0.974	1.000	4.395	4.367	0.882	0.890

$\beta_{1.2.2.1}$	1.2	1.000	1.000	0.233	0.233	0.888	0.898	0.7	1.000	1.000	0.158	0.158	0.886	0.899	0.05	1.000	1.000	0.013	0.013	0.884	0.899
$\alpha_{1.2.2.2}$	32.0	0.984	1.000	5.683	5.661	0.886	0.908	22.0	0.977	1.000	4.719	4.693	0.856	0.892	17.00	0.971	1.000	4.155	4.125	0.890	0.910
$\beta_{1.2.2.2}$	1.5	1.000	1.000	0.281	0.281	0.889	0.898	1.5	1.000	1.000	0.348	0.348	0.885	0.899	0.40	1.000	1.000	0.108	0.108	0.882	0.899
$\alpha_{2.1.1.1}$	17.0	0.971	1.000	4.155	4.125	0.890	0.910	20.0	0.975	1.000	4.504	4.476	0.873	0.908	20.00	0.975	1.000	4.504	4.476	0.873	0.908
$\beta_{2.1.1.1}$	1.0	1.000	1.000	0.270	0.270	0.882	0.899	0.5	1.000	1.000	0.122	0.122	0.885	0.898	0.20	1.000	1.000	0.049	0.049	0.885	0.898
$\alpha_{2.1.1.2}$	18.0	0.972	1.000	4.276	4.247	0.880	0.900	17.0	0.971	1.000	4.155	4.125	0.890	0.910	26.00	0.981	1.000	5.124	5.100	0.879	0.906
$\beta_{2.1.1.2}$	1.1	1.000	1.000	0.286	0.286	0.884	0.900	0.9	1.000	1.000	0.243	0.243	0.882	0.899	0.01	0.989	0.989	0.002	0.002	0.880	0.903
$\alpha_{2.1.2.1}$	20.0	0.975	1.000	4.504	4.476	0.873	0.908	28.0	0.982	1.000	5.321	5.297	0.865	0.893	32.00	0.984	1.000	5.683	5.661	0.886	0.908
$\beta_{2.1.2.1}$	1.0	1.000	1.000	0.245	0.245	0.885	0.898	2.0	1.000	1.000	0.404	0.404	0.888	0.899	0.30	1.000	1.000	0.056	0.056	0.889	0.898
$\alpha_{2.1.2.2}$	18.0	0.972	1.000	4.276	4.247	0.880	0.900	22.0	0.977	1.000	4.719	4.693	0.856	0.892	19.00	0.974	1.000	4.395	4.367	0.882	0.890
$\beta_{2.1.2.2}$	2.1	1.000	1.000	0.547	0.547	0.884	0.900	1.0	1.000	1.000	0.232	0.232	0.885	0.899	0.70	1.000	1.000	0.177	0.177	0.884	0.899
$\alpha_{2.2.1.1}$	29.0	0.983	1.000	5.413	5.390	0.893	0.905	19.0	0.974	1.000	4.395	4.367	0.882	0.890	23.00	0.978	1.000	4.824	4.798	0.889	0.904
$\beta_{2.2.1.1}$	1.6	1.000	1.000	0.316	0.316	0.888	0.899	0.4	1.000	1.000	0.101	0.101	0.884	0.899	0.40	1.000	1.000	0.090	0.090	0.886	0.899
$\alpha_{2.2.1.2}$	25.0	0.980	1.000	5.027	5.002	0.886	0.889	23.0	0.978	1.000	4.824	4.798	0.889	0.904	25.00	0.980	1.000	5.027	5.002	0.886	0.889
$\beta_{2.2.1.2}$	1.5	1.001	1.001	0.323	0.323	0.887	0.900	0.1	1.000	1.000	0.023	0.023	0.886	0.899	0.90	1.001	1.001	0.194	0.194	0.887	0.900
$\alpha_{2.2.2.1}$	24.0	0.979	1.000	4.927	4.902	0.892	0.897	20.0	0.975	1.000	4.504	4.476	0.873	0.908	26.00	0.981	1.000	5.124	5.100	0.879	0.906
$\beta_{2.2.2.1}$	2.5	1.000	1.000	0.551	0.551	0.886	0.899	1.4	1.000	1.000	0.343	0.343	0.885	0.898	0.10	1.001	1.001	0.021	0.021	0.887	0.899
$\alpha_{2.2.2.2}$	23.0	0.978	1.000	4.824	4.798	0.889	0.904	18.0	0.972	1.000	4.276	4.247	0.880	0.900	20.00	0.975	1.000	4.504	4.476	0.873	0.908
$\beta_{2.2.2.2}$	3.0	1.000	1.000	0.677	0.677	0.886	0.899	1.9	1.000	1.000	0.494	0.494	0.884	0.900	0.20	1.000	1.000	0.049	0.049	0.885	0.898

4.5 Applications

In Louzada *et al.* (2019), the Annelida's traction subsystem was used to illustrate the methodology discussed at the time. With the advance in the development of the robotic unit, the applications exposed here refine the previous one with the inclusion of more details about this important subsystem (Section 4.5.2), and also adding the new results obtained in this chapter to another subsystem as well (pressure vessel subsystem in Section 4.5.1). The information available so far comes from their respective FMEA (Failure Mode and Effects Analysis) tables, which were reviewed by the technical team to further deepen the knowledge they have about their idealizations and tests already carried out.

In both subsystems, the parameters of the model proposed here were associated with the Severity (S), Occurrence (O) and Detectability (D) indices. Thus, it was possible to take random realizations (based on Algorithm 4) of the failure times that represent the perspective provided by the technical team at FMEA, from the perspective of reliability. This approach ensures the addition of information at this early stage of the project.

The current reliability requirements for both subsystems include, among numerous other factors, that: (i) with high probability, the system must remain in operation (without any failure) for at least a minimum number of days; (ii) on the other hand, the median lifetime of the first failure is expected to be around another number of days. These requirements were determined by considering the expected number of annual missions, the size of the step taken by the robotic unit inside the oil pipelines, its respective speed and the estimated time for the hydrate block to melt. Such estimates also allowed to assess the time needed per mission and, therefore, the minimum desired lifetime.

4.5.1 In-Pipe Robot – Pressure Vessel Set

Based on the reliability requirements for the pressure vessel system, we consider that with high probability ($\approx 95\%$) the first failure should not occur before 68 days. Similarly, the median of first-failure time should be around 136 days. In this way, we can plot the reliability curve associated with the proposed model (Figure 28). This curve is a partial reference for the reliability that we want to achieve at the end of the development of this system, in the light of the currently proposed modeling. We will see that all the uncertainty involved in this initial stage of the project exposes us to how far we are from this objective. However, it does allow for more targeted research routes.

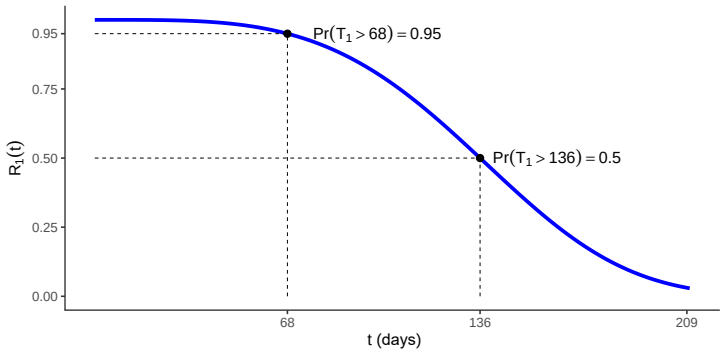


Figure 28 – Pressure vessel set's reliability requirements.

The project designed 11 pressure vessels connected by connecting cables, as shown in Figure 19. Although the whole set of vessels will be exposed to the same environmental characteristics, their functions have different purposes and importances. However, the individual evaluation of the vessels will not be considered in this work, so that the same FMEA will represent, here, all 11 subsystems and the indices used, as well as the FTA, are shown in Table 19.

Table 19 – FTA (with FMEA indices) for the in-pipe robot's pressure vessel system.

			S	O	D
All Pressure Vessel / Pressure Vessel Set (*there are 11 of these subsystems*)	Cable Conductor	—	7	7	5
	Pressure Vessel	Rear Vessel Cap	8	5	4
		Heatpipe	6	5	3
		Carbon Fiber Protection	8	4	7
		Vessel Base	8	4	7
		Front Vessel Cap	7	4	4

An illustration of the data set generated can be found in Table 27 of Appendix B.1, and obtained in full via an individual request for the authors. Due to the approximately linear

behavior in the Duane plots, for each failure mode (see Figure 52 of Appendix B.2), we have indications that the theoretical PLP model may be appropriate to model a problem like this.

The parameter estimates (see Table 20) express that, in a practical situation, the Cable Conductor subsystem would be undergoing a degradation process; there is statistical evidence of this in some cases (6/11), and in others only indications (5/11). On the other hand, the same occurs less frequently in the other subsystems.

Table 20 – BMAP estimates and 95% credibility intervals (95% CIs) for the model parameters, considering the in-pipe robot's pressure vessel system's failure data.

		95% CI			95% CI				95% CI			95% CI	
System	$\widehat{\alpha}$	Lower	Upper	$\widehat{\beta}$	Lower	Upper	System	$\widehat{\alpha}$	Lower	Upper	$\widehat{\beta}$	Lower	Upper
1.1.1	47	34.9624	61.9290	**1.4684**	**1.1024**	**1.9587**	1.7.1	47	34.9624	61.9290	**1.0539**	0.7912	1.4057
1.1.2.1	28	19.0134	39.8761	0.6813	0.4695	0.9913	1.7.2.1	33	23.1305	45.7597	0.5341	0.3792	0.7541
1.1.2.2	29	19.8309	41.0587	0.8647	0.5998	1.2498	1.7.2.2	29	19.8309	41.0587	0.8975	0.6225	1.2971
1.1.2.3	26	17.3882	37.5009	**1.0627**	0.7219	1.5687	1.7.2.3	35	24.7961	48.0944	0.8997	0.6451	1.2572
1.1.2.4	28	19.0134	39.8761	0.9132	0.6293	1.3286	1.7.2.4	33	23.1305	45.7597	**1.0314**	0.7322	1.4560
1.1.2.5	39	28.1545	52.7364	**1.2672**	0.9248	1.7393	1.7.2.5	40	28.9992	53.8917	**1.2499**	0.9158	1.7086
1.2.1	52	39.2682	67.6235	**1.4171**	**1.0791**	**1.8633**	1.8.1	52	39.2682	67.6235	**1.1184**	0.8516	1.4705
1.2.2.1	24	15.7775	35.1112	0.5534	0.3700	0.8303	1.8.2.1	29	19.8309	41.0587	0.4055	0.2812	0.5860
1.2.2.2	37	26.4710	50.4197	0.7335	0.5308	1.0154	1.8.2.2	37	26.4710	50.4197	0.8517	0.6164	1.1791
1.2.2.3	36	25.6324	49.2581	0.9266	0.6675	1.2887	1.8.2.3	38	27.3117	51.5791	0.8932	0.6492	1.2312
1.2.2.4	23	14.9781	33.9103	0.5451	0.3613	0.8253	1.8.2.4	30	20.6516	42.2382	0.7565	0.5280	1.0865
1.2.2.5	52	39.2682	67.6235	0.9670	0.7364	1.2715	1.8.2.5	29	19.8309	41.0587	**1.0210**	0.7082	1.4756
1.3.1	39	28.1545	52.7364	**1.2904**	0.9418	1.7712	1.9.1	49	36.6805	64.2110	**1.3643**	**1.0304**	**1.8089**
1.3.2.1	15	8.7694	24.1159	**1.0601**	0.6357	1.7786	1.9.2.1	23	14.9781	33.9103	0.5200	0.3446	0.7873
1.3.2.2	33	23.1305	45.7597	**1.1131**	0.7901	1.5713	1.9.2.2	37	26.4710	50.4197	0.7737	0.5599	1.0712
1.3.2.3	32	22.3015	44.5886	**1.0118**	0.7144	1.4362	1.9.2.3	33	23.1305	45.7597	0.7979	0.5664	1.1264
1.3.2.4	33	23.1305	45.7597	0.8795	0.6243	1.2416	1.9.2.4	29	19.8309	41.0587	0.8174	0.5670	1.1813
1.3.2.5	37	26.4710	50.4197	0.9408	0.6808	1.3025	1.9.2.5	47	34.9624	61.9290	0.9518	0.7145	1.2695
1.4.1	33	23.1305	45.7597	**1.8210**	**1.2927**	**2.5707**	1.10.1	55	41.8675	71.0243	**1.3868**	**1.0641**	**1.8095**
1.4.2.1	26	17.3882	37.5009	0.4658	0.3165	0.6876	1.10.2.1	31	21.4751	43.4148	0.6620	0.4648	0.9450
1.4.2.2	26	17.3882	37.5009	0.8439	0.5733	1.2458	1.10.2.2	34	23.9621	46.9282	0.7932	0.5660	1.1140
1.4.2.3	32	22.3015	44.5886	**1.1062**	0.7811	1.5702	1.10.2.3	38	27.3117	51.5791	0.9273	0.6740	1.2782
1.4.2.4	30	20.6516	42.2382	**1.2512**	0.8733	1.7969	1.10.2.4	30	20.6516	42.2382	**1.0195**	0.7116	1.4642
1.4.2.5	39	28.1545	52.7364	**1.0461**	0.7635	1.4359	1.10.2.5	37	26.4710	50.4197	**1.0857**	0.7857	1.5031
1.5.1	38	27.3117	51.5791	**1.2270**	0.8917	1.6912	1.11.1	46	34.1056	60.7857	**1.0990**	0.8225	1.4705
1.5.2.1	35	24.7961	48.0944	0.7052	0.5057	0.9855	1.11.2.1	27	18.1991	38.6902	0.5680	0.3887	0.8323
1.5.2.2	26	17.3882	37.5009	0.6888	0.4680	1.0169	1.11.2.2	37	26.4710	50.4197	0.8204	0.5937	1.1358
1.5.2.3	30	20.6516	42.2382	0.7637	0.5331	1.0969	1.11.2.3	34	23.9621	46.9282	0.8185	0.5840	1.1495
1.5.2.4	30	20.6516	42.2382	0.9083	0.6340	1.3045	1.11.2.4	38	27.3117	51.5791	0.6382	0.4638	0.8797
1.5.2.5	39	28.1545	52.7364	0.9437	0.6888	1.2954	1.11.2.5	34	23.9621	46.9282	**1.0132**	0.7229	1.4228
1.6.1	68	53.2453	85.6470	**1.2945**	**1.0202**	**1.6440**							
1.6.2.1	35	24.7961	48.0944	**1.1473**	0.8226	1.6032							
1.6.2.2	28	19.0134	39.8761	0.9043	0.6231	1.3157							
1.6.2.3	24	15.7775	35.1112	0.9041	0.60045	1.3566							
1.6.2.4	30	20.6516	42.2382	0.9267	0.6468	1.3310							
1.6.2.5	33	23.1305	45.7597	**1.0657**	0.7565	1.5045							

Some subsystems did not have an increasing intensity function (which would indicate a degradation process), however, the failure intensity in the initial times was higher, which results in the high occurrence of failures in the first moments of activity and, therefore, significantly reduces the component's survival time. This occurred more frequently in the Heatpipe and Carbon Fiber Protection subsystems. This information cannot be perceived directly on the parameter estimates (Table 20), however, the graphs associated with the first-failure time reliability and intensity curves, obtained by the adjustment, are shown in Figure 29.

From Figure 29, it is possible to notice that the desired reliability requirements are still far away. The median first-failure time of a pressure vessel is close to one day (depending on the randomization of the simulated data), which is still far from the desired 68 days. The result obtained by the model does not reflect any time observed in practice, however, it describes, from the perspective of the reliability analysis, all the uncertainty that still surrounds the development of this system component.

The graph of the observed number of failures *versus* the number of failures estimated over time can be used to assess the quality of the model's fit. In Figure 53 of Appendix B.3, the plots for each component are presented and, in general, in a practical situation we would understand that the model was able to describe the observed behavior.

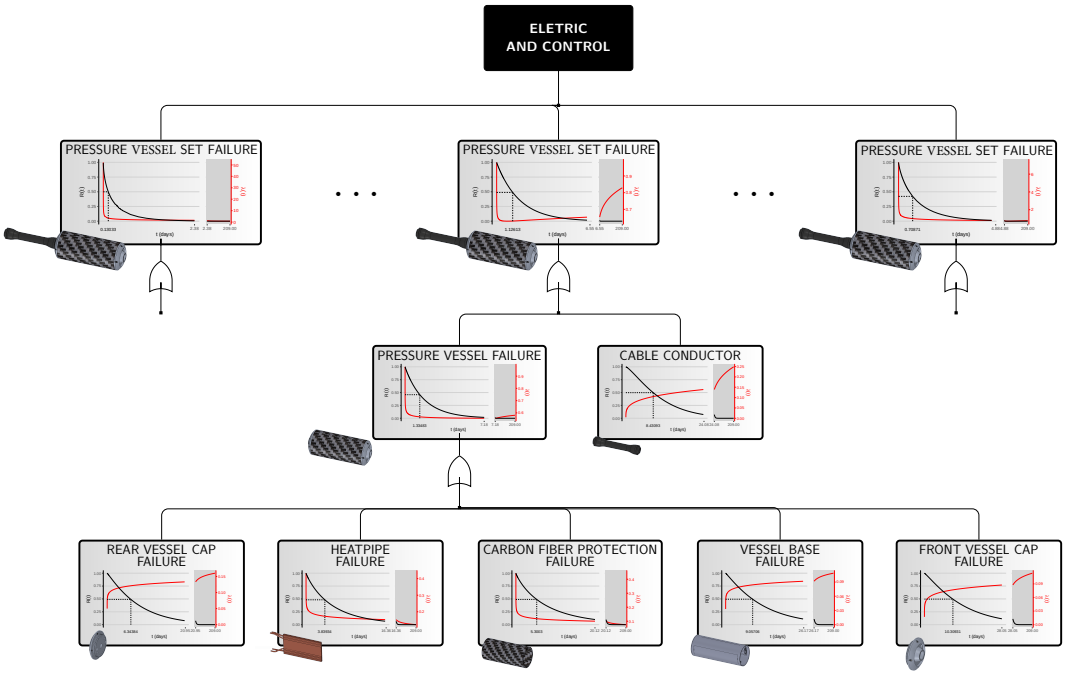


Figure 29 – First-failure time reliability (in black) and intensity (in red) functions by components, subsystems and systems, for the in-pipe robotic unit's pressure vessel set system.

4.5.2 *In-Pipe Robot – Traction System*

In this section, we return to the situation described in Section **??**. We obtained a suitable data set (an example is given in Table 28 of Appendix B.5) for this problem using a

similar approach as proposed in the previous section, i.e., based on the limited but available information provided by the revised FMEA and FTA tools (see Figure 54).

The required reliability for the traction system is graphically exposed in Figure 30 of Appendix B.4, and claims, with high probability, that the system remains in operation for at least 102 days, without any failure. Also, the median failure time of the system has to be approximately 170 days.

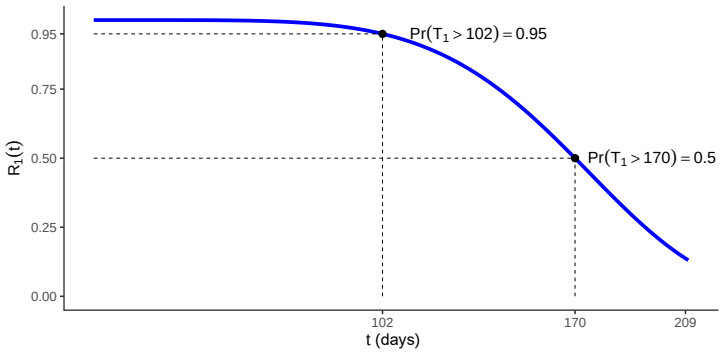


Figure 30 – Traction system's reliability requirements.

The Duane plots built for each failure mode (see Figure 55 of Appendix B.6) have an approximately linear behavior, which demonstrates the theoretical suitability that a PLP model needs for adjustment.

In a practical context, if the failure times came from some missions, the BMAP estimates of the adjusted model for the traction system (see Table 21) would show that there are evidences of some components having an increasing failure intensity function (systems 1.3.5.5, 1.3.6.4-5, 1.3.7.4, 1.3.8.4, 1.3.9.4-5, 1.3.10.4-5); for many others, there are some indications; and for some, evidences of non-deterioration (systems 1.1.1.2-3, 1.1.2.3, 1.1.3.3, 1.2.1, 1.2.5, 1.3.2, 1.3.5.2, 1.3.6.2, 1.3.8.2, 2.1.3.3-5, 2.1.4.3-5, 2.1.5.3-5, 2.3.3.3, 2.3.4.3, 2.3.5.3).

From these results, we identify that the components responsible for the main failure causes, which represent the biggest obstacles to reaching the reliability requirements, are the rubber (1.3.5-10.4) and adhesive (1.3.5-10.5) components. In particular, these units express an evident degradation behavior, which suggests the need for a preventive maintenance regime dedicated especially to these components, or even the renewal of the design idealized for the process performed by them. Indeed, the latter is what is currently being carried out, since the preliminary practical tests of the traction system exposed serious failures associated with the strength of the adhesive and the rubber to withstand the necessary force for the locomotion.

Table 21 – BMAP estimates and 95% CIs for the model parameters, considering the in-pipe robot's traction system's failure data.

		95% CI			95% CI				95% CI			95% CI	
System	$\widehat{\alpha}$	Lower	Upper	$\widehat{\beta}$	Lower	Upper	System	$\widehat{\alpha}$	Lower	Upper	$\widehat{\beta}$	Lower	Upper
1.1.1.1	3	0.8449	8.0064	1.8669	0.5775	6.7439	2.1.1	6	2.5044	12.3678	1.2642	0.5567	2.9503
1.1.1.2	12	6.5599	20.3232	0.8939	0.5039	1.5995	2.1.2	8	3.7821	15.0955	1.4402	0.7106	2.9674
1.1.1.3	29	19.8309	41.0587	0.9935	0.6891	1.4359	2.1.3.1	4	1.3502	9.5114	2.5054	0.9102	7.3218
1.1.1.4	43	31.5447	57.3465	1.1332	0.8396	1.5317	2.1.3.2	16	9.5233	25.3625	1.1185	0.6819	1.8448
1.1.1.5	36	25.6324	49.2581	1.1339	0.8169	1.5770	2.1.3.3	28	19.0134	39.8761	0.9348	0.6442	1.3601
1.1.2.1	5	1.9079	10.9600	1.8970	0.7699	4.8570	2.1.3.4	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.1.2.2	25	16.5809	36.3080	1.1912	0.8030	1.7724	2.1.3.5	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.1.2.3	34	23.9621	46.9282	0.9243	0.6595	1.2981	2.1.4.1	4	1.3502	9.5114	2.5054	0.9102	7.3218
1.1.2.4	37	26.4710	50.4197	1.1349	0.8213	1.5712	2.1.4.2	16	9.5233	25.3625	1.1185	0.6819	1.8448
1.1.2.5	42	30.6944	56.1967	1.1370	0.8394	1.5425	2.1.4.3	28	19.0134	39.8761	0.9348	0.6442	1.3601
1.1.3.1	4	1.3502	9.5114	2.3031	0.8367	6.7307	2.1.4.4	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.1.3.2	10	5.1414	17.7394	1.0611	0.5654	2.0142	2.1.4.5	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.1.3.3	38	27.3117	51.5791	0.9610	0.6985	1.3246	2.1.5.1	4	1.3502	9.5114	2.5054	0.9102	7.3218
1.1.3.4	37	26.4710	50.4197	1.1349	0.8213	1.5712	2.1.5.2	16	9.5233	25.3625	1.1185	0.6819	1.8448
1.1.3.5	43	31.5447	57.3465	1.1332	0.8396	1.5317	2.1.5.3	28	19.0134	39.8761	0.9348	0.6442	1.3601
1.1.4	15	8.7694	24.1159	1.0673	0.6401	1.7908	2.1.5.4	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.1.5	8	3.7821	15.0955	1.3721	0.6770	2.8271	2.1.5.5	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.2.1	28	19.0134	39.8761	0.9679	0.6670	1.4083	2.2.1	21	13.3927	31.4952	1.0466	0.6803	1.6164
1.2.2	3	0.8449	8.0064	1.8054	0.5585	6.5216	2.2.2	5	1.9079	10.9600	1.9708	0.7999	5.0459
1.2.3	25	16.5809	36.3080	1.0826	0.7298	1.6109	2.2.3	23	14.9781	33.9103	1.0231	0.6780	1.5490
1.2.4	19	11.8272	29.0600	1.0910	0.6933	1.7242	2.2.4	12	6.5599	20.3232	1.0294	0.5802	1.8418
1.2.5	13	7.2867	21.5973	0.9651	0.5567	1.6858	2.3.1	10	5.1414	17.7394	1.2872	0.6858	2.4435
1.3.1	14	8.0235	22.8611	1.0205	0.6009	1.7452	2.3.2	12	6.5599	20.3232	1.0294	0.5802	1.8418
1.3.2	11	5.8443	19.0378	0.9382	0.5152	1.7254	2.3.3.1	5	1.9079	10.9600	1.9708	0.7999	5.0459
1.3.3	15	8.7694	24.1159	1.0193	0.6112	1.7102	2.3.3.2	16	9.5233	25.3625	1.1185	0.6819	1.8448
1.3.4	7	3.1311	13.7442	1.0916	0.5120	2.3760	2.3.3.3	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.3.5.1	11	5.8443	19.0378	1.1527	0.6329	2.1198	2.3.3.4	40	28.9992	53.8917	1.0097	0.7398	1.3803
1.3.5.2	9	4.4533	16.4262	0.9203	0.4734	1.8133	2.3.3.5	40	28.9992	53.8917	1.0097	0.7398	1.3803
1.3.5.3	31	21.4751	43.4148	1.1724	0.8232	1.6737	2.3.4.1	5	1.9079	10.9600	1.9708	0.7999	5.0459
1.3.5.4	39	28.1545	52.7364	1.3199	0.9633	1.8116	2.3.4.2	16	9.5233	25.3625	1.1185	0.6819	1.8448
1.3.5.5	48	35.8208	63.0707	**1.4029**	**1.0564**	**1.8656**	2.3.4.3	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.3.6.1	6	2.5044	12.3678	1.3407	0.5904	3.1286	2.3.4.4	40	28.9992	53.8917	1.0097	0.7398	1.3803
1.3.6.2	17	10.2847	26.6017	0.9860	0.6103	1.6012	2.3.4.5	40	28.9992	53.8917	1.0097	0.7398	1.3803
1.3.6.3	39	28.1545	52.7364	1.0998	0.8026	1.5095	2.3.5.1	5	1.9079	10.9600	1.9708	0.7999	5.0459
1.3.6.4	55	41.8675	71.0243	**1.3636**	**1.0463**	**1.7792**	2.3.5.2	16	9.5233	25.3625	1.1185	0.6819	1.8448
1.3.6.5	50	37.5417	65.3499	**1.3782**	**1.0438**	**1.8221**	2.3.5.3	33	23.1305	45.7597	0.8843	0.6278	1.2484
1.3.7.1	6	2.5044	12.3678	1.3407	0.5904	3.1286	2.3.5.4	40	28.9992	53.8917	1.0097	0.7398	1.3803
1.3.7.2	16	9.5233	25.3625	1.0194	0.6215	1.6813	2.3.5.5	40	28.9992	53.8917	1.0097	0.7398	1.3803
1.3.7.3	38	27.3117	51.5791	1.0821	0.7865	1.4916							
1.3.7.4	51	38.4043	66.4873	**1.3946**	**1.0591**	**1.8386**							
1.3.7.5	42	30.6944	56.1967	1.3054	0.9638	1.7710							
1.3.8.1	9	4.4533	16.4262	1.2405	0.6382	2.4443							
1.3.8.2	17	10.2847	26.6017	0.9860	0.6103	1.6012							
1.3.8.3	44	32.3967	58.4945	1.1088	0.8244	1.4935							
1.3.8.4	50	37.5417	65.3499	**1.3782**	**1.0438**	**1.8221**							
1.3.8.5	41	29.8459	55.0451	1.3312	0.9792	1.8127							
1.3.9.1	11	5.8443	19.0378	1.1527	0.6329	2.1198							
1.3.9.2	18	11.0528	27.8340	1.0040	0.6300	1.6075							
1.3.9.3	37	26.4710	50.4197	1.0857	0.7857	1.5031							
1.3.9.4	51	38.4043	66.4873	**1.3946**	**1.0591**	**1.8386**							
1.3.9.5	53	40.1334	68.7584	**1.4198**	**1.0840**	**1.8619**							
1.3.10.1	4	1.3502	9.5114	2.2272	0.8091	6.5089							
1.3.10.2	18	11.0528	27.8340	1.0040	0.6300	1.6075							
1.3.10.3	51	38.4043	66.4873	1.1621	0.8825	1.5320							
1.3.10.4	48	35.8208	63.0707	**1.4029**	**1.0564**	**1.8656**							
1.3.10.5	46	34.1056	60.7857	**1.3724**	**1.0271**	**1.8363**							

The estimated reliability and intensity functions for the first-failure times are shown in Figure 31. From this figure, it can be seen that the median first-failure time for some components, such as Spring (1.1.1-3.1) and Copling Rod (1.2.2), is high, with approximately 90 days. Other components, such as rubber (1.1.1-3.4), adhesive (1.1.1-3.5), paw (1.1.1-3.3) and main hydraulic piston (1.2.1), have a small first-failure median time, around 5 days. However, at the end of the interaction between all components and subsystems, the median time of the first-failure is around 0.12543 days for the return locomotive, and 0.21694 days for the forward locomotive.

The goodness of fit can be seen by the comparison between the observed and estimated number of failures along the time. These graphs, for each component, can be found in Figure 56 of Appendix B.7.

4.6 Concluding Remarks

In this chapter, we have continued the study started in (LOUZADA *et al.*, 2019), presenting the model under consideration of an arbitrary number of hierarchical levels and maintaining the assumptions of competing risks with independent failure modes, in a MR regime with a reparametrized PLP intensity function. In this context, we have obtained Bayesian estimators with corrected biases, as well as we have derived exact credibility intervals for the parameters. The properties of these estimators were evaluated in a simulation study, which returned good results.

The model structured in the proposed way allowed to highlight analytically and graphically the reliability associated with the first-failure times of each one of the subsystems (at any hierarchical level) of two arbitrary systems that illustrate the use of modeling. Namely, the pressure vessels set and the traction system of the developing system that has served as a practical motivation for this theoretical development.

As future works, we intend to evaluate the quality of these estimators in a context with outliers, their behavior (in terms of quality loss) when exposed to data from an IR regime. In addition, we wish to evaluate the change in reliability based on the increase in redundancy of some subsystems. We also intend to assume that repairs are either perfect or imperfect, and model the dependence among the failure modes via shared frailty models.

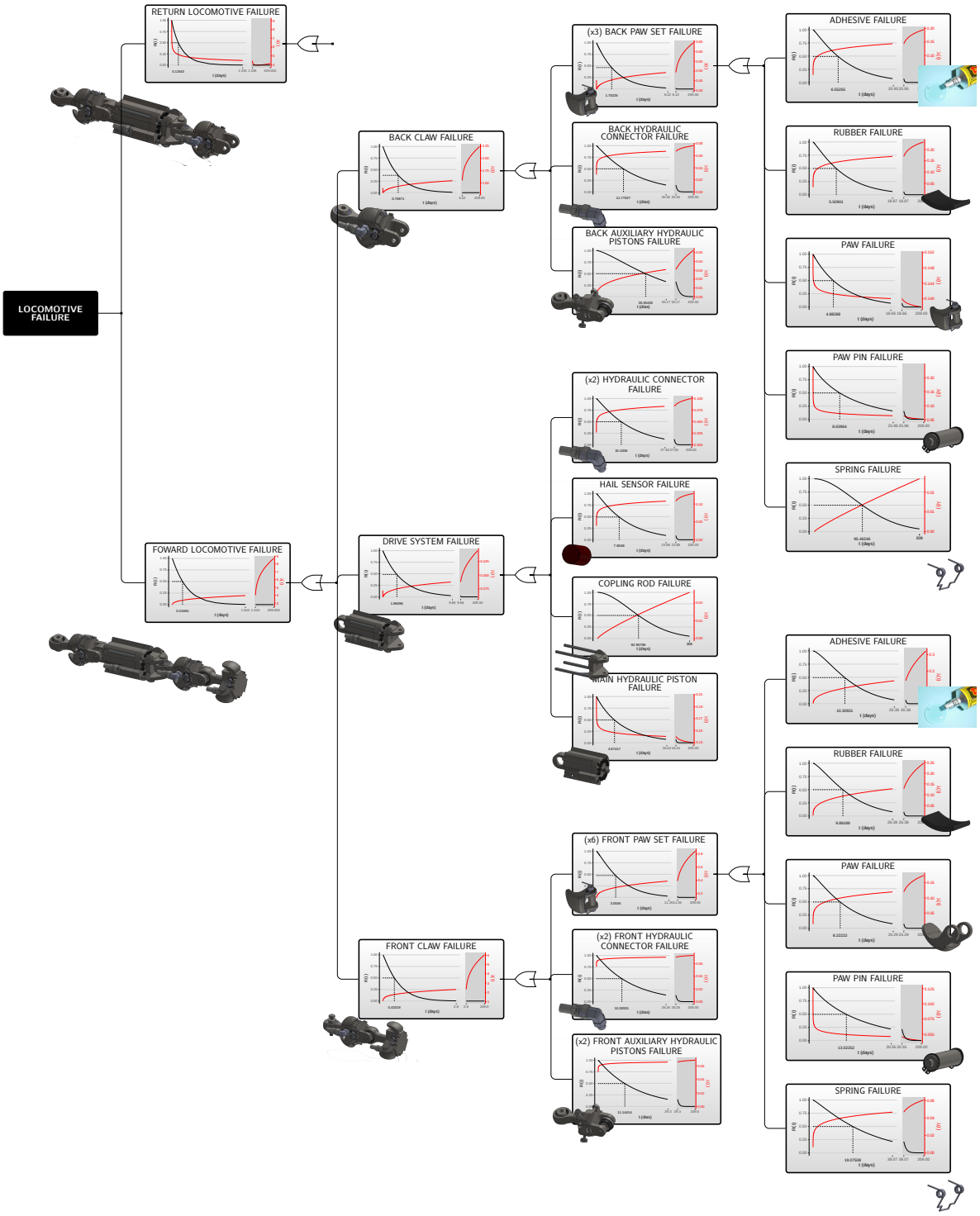


Figure 31 – First-failure time reliability (in black) and intensity (in red) functions by components, subsystems and systems, for the in-pipe robotic unit's traction system.

CHAPTER

5

CONFIDENCE ESTIMATION BY ADOPTING A PLP MODEL HIERARCHICALLY REPRESENTED WITH SERIES-PARALLEL CONNECTIONS SUBJECT TO COMPETING RISKS UNDER MR REGIME

The purpose of this study is a direct extension of the research presented in the Chapter 3. In this sense, we seek to propose a model that allows us to evaluate the failure times of a single repairable system represented hierarchically, exposed to a competing risks and MRs framework. The extension comprises the representation of complex systems through a hierarchical structure in series and/or in parallel. For this, we deduce the general form of the model, the likelihood function associated with it to obtain reliable estimates for the parameters that index the model. In addition, we present the mechanism for generating random numbers based on the presented structure, which allows obtaining point and interval estimates (via parametric bootstrap) in a more convenient way for reliability curves at any level of the system hierarchy. We illustrate the application of the model with the sequence of application presented in the Chapters 3 and 4, which deal with the reliability modeling of a robotic unit still under development, resulting from a project carried out in partnership between Petrobras and other Brazilian research centers, of which the ICMC is a part.

Contents of Chapter

5.1 Framework

In complex systems, the reliability of each component has a fraction of the responsibility for quantifying the reliability of the system, and the structure of how these components and their upper levels are arranged has an even greater impact. It is not uncommon that there is a need to add extra components as a redundancy mechanism to ensure a desired level of reliability, this is a practice studied for a long time, some interesting examples can be seen in (CANTARELLA, 1983; CHEN, 1997; ZUO, 2008; RUIZ, 2011; NOTASH, 2017; ZHANG WENPING; XU, 2018).

In the previous models we consider a minimal repair competing risk framework for a hierarchically represented system, in which failure modes are described by a power-law process, however, such model does not take into account the possibility of a parallelized structure, which will be described in this chapter.

5.2 Proposed Model

Our aim is to analyze failure times of a single repairable system subject to competing hierarchical risks. The idea is to generalize the work of Louzada *et al.* (2019) considering a series-parallel structure. That is, the systems and sub-systems are connected in series, but the components that make up the sub-system may have associations in parallel. In Figure 1 we present an illustration of the idealized structure using the fault tree. We adopt that the number of systems is p and for a system j there are k_j sub-systems. For a given subsystem jk, the number of components that compose it is r_{jk}. Therefore, if $r_{jk} = 1$ indicates a single component, but if $r_{jk} > 1$ implies that the sub-system has an association in parallel with r_{jk} components.

We emphasize that the failure of a sub-system that is composed of a single component results in the failure of the system as a whole. But if the sub-system is composed of several components connected in parallel, the system as a whole only fails at the moment when that all components in parallel fail.

The dataset of the hierarchical competitive risk model consists of the 3-tuples (t, δ, ψ), where $t > 0$ denoting the time of failure, δ is the indicator of the main cause of failure (system) and ψ is the indicator of the sub-cause (sub-system).

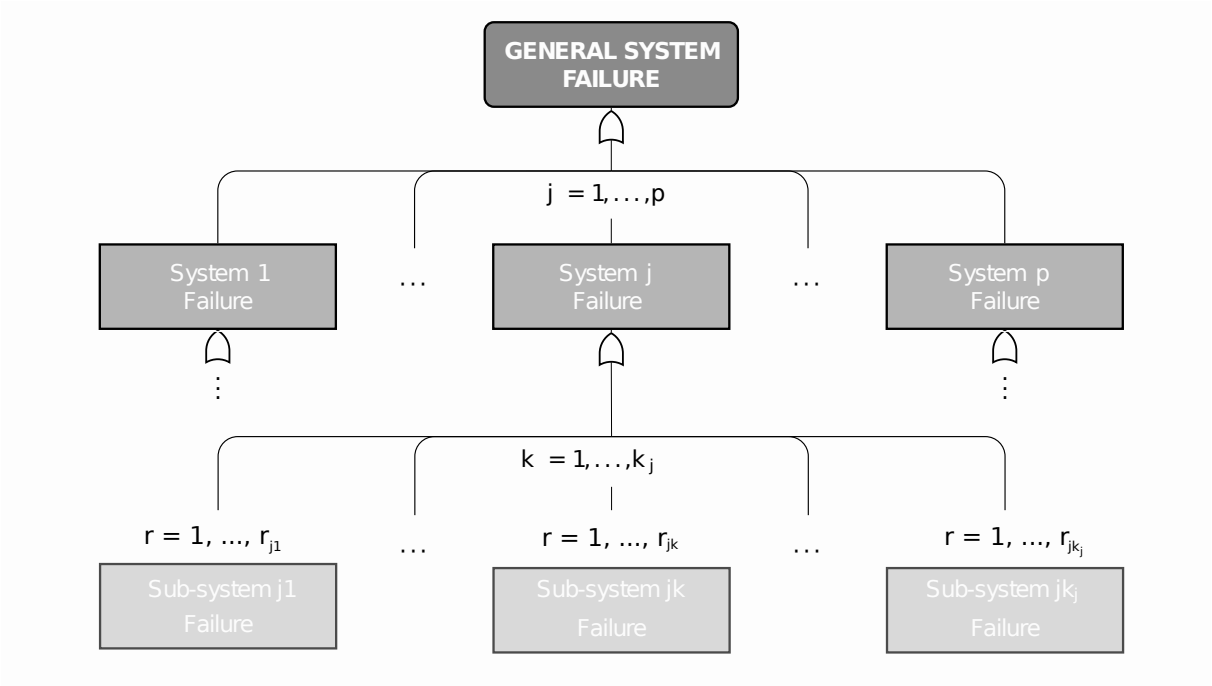


Figure 32 – Ilustração do sistema

The hierarchical competing risks model for series-parallel systems (HCRSPS) is formulated as follows: first, we assume that the failures of all components of a subsystem jk follow a NHPP with an intensity function given by

$$\lambda_{jk} = \left(\frac{\beta_{jk}}{\mu_{jk}}\right)\left(\frac{t}{\mu_{jk}}\right)^{\beta_{jk}-1},$$

for $j = 1,\ldots,p$ and $k = 1,\ldots,k_j$, with k_j the number of sub-systems for the j-th system, $\mu_{jk} > 0$ and $\beta_{jk} > 0$ are, respectively, the scale and shape parameters.

We assume that the system is observed until time T, and adopting reparametrization in terms of β_{jk} and

$$\alpha_{jk} = E(N_{jk}(T)) = \left(\frac{T}{\mu_{jk}}\right)^{\beta_{jk}},$$

where N_{jk} is the process of counting the failures of each component of the sub-system jk. This transformation results in orthogonal β_{jk} and α_{jk} parameters. From this, the intensity function for sub-systems with a single component is given by

$$\lambda_{jk}(t|\delta,\psi) = \frac{\beta_{jk}\alpha_{jk}t^{\beta_{jk}-1}}{T^{\beta_{jk}}},$$

in the same way as exposed in Louzada *et al.* (2019). While for sub-systems with r_{jk} components in parallel, the reliability is

$$R_{jk}(t|\delta,\psi) = \begin{cases} \exp\left\{-\alpha_{jk}\left(\frac{t}{T}\right)^{\beta_{jk}}\right\}, & \text{if } r_{jk} = 1, \\ 1-\left(1-\exp\left\{-\alpha_{jk}\left(\frac{t}{T}\right)^{\beta_{jk}}\right\}\right)^{r_{jk}}, & \text{if } r_{jk} > 1, \end{cases}$$

the accumulated intensity function is given by

$$\Lambda_{jk}(t|\delta,\psi)=\begin{cases}\alpha_{jk}\left(\dfrac{t}{T}\right)^{\beta_{jk}}, & \text{if } r_{jk}=1,\\ -\log\left[1-\left(1-\exp\left(-\alpha_{jk}\left(\dfrac{t}{T}\right)^{\beta_{jk}}\right)\right)^{r_{jk}}\right], & \text{if } r_{jk}>1,\end{cases}$$

and therefore, the intensity function for sub-systems with r_{jk} componentes in parallel is given by

$$\lambda_{jk}(t|\delta,\psi)=\begin{cases}\dfrac{\beta_{jk}\alpha_{jk}t^{\beta_{jk}-1}}{T^{\beta_{jk}}}, & \text{if } r_{jk}=1,\\ \dfrac{r_{jk}\left\{1-\exp\left[-\left(\frac{t}{T}\right)^{\beta_{jk}}\alpha_{jk}\right]\right\}^{r_{jk}-1}\left(\frac{\beta_{jk}t^{\beta_{jk}-1}\alpha_{jk}}{T^{\beta_{jk}}}\right)\exp\left[-\alpha_{jk}\left(\frac{t}{T}\right)^{\beta_{jk}}\right]}{1-\left\{1-\exp\left[-\alpha_{jk}\left(\frac{t}{T}\right)^{\beta_{jk}}\right]\right\}^{r_{jk}}}, & \text{if } r_{jk}>1,\end{cases}$$

on the other side, the respective cumulative intensity function calculated at time T are given by

$$\Lambda_{jk}(T|\delta,\psi)=\begin{cases}\alpha_{jk}, & \text{if } r_{jk}=1\\ -\log\left[1-(1-\exp(-\alpha_{jk}))^{r_{jk}}\right], & \text{if } r_{jk}>1.\end{cases}$$

Since the sub-systems are connected in series, the reliability function of system j is given by

$$R_j(t)=\prod_{k=1}^{k_j}R_{jk}(t),$$

using the same reasoning, we have that the overall system reliability is

$$R(t)=\prod_{j=1}^{p}R_j(t),$$

with this we can calculate reliability at any level of the hierarchy.

5.3 Inference

In this section we discuss about obtaining maximum likelihood estimators and interval estimates considering asymptotic properties. In addition, when few failures are observed, we suggest a bias correction method.

If n failures are observed in the time interval $(0,T]$, then the dataset is (t_i,δ_i,ψ_i), with $i=1,\ldots,n$, where $0<t_1<t_2<\ldots<t_n<T$ are the system failure times, δ_i and ψ_i indicate the i-th cause of failure and i-th sub-cause failure, respectively.

In this work, we use the same assumption adopted in Louzada *et al.* (2019). Meeker and Escobar (2014) presents a discussion on the correlation of failure time between components.

Adopting that failure modes act independently and are mutually exclusive, the likelihood function is given by

$$L(\boldsymbol{\eta}|\boldsymbol{t},\boldsymbol{\delta},\boldsymbol{\psi}) = \prod_{i=1}^{n}\prod_{j=1}^{p}\prod_{k=1}^{n_j}\lambda_{jk}(t_i|\delta_i,\psi_i)^{I(\delta_i=j,\psi_i=k)}\exp\left(-\sum_{j=1}^{p}\sum_{k=1}^{n_j}\Lambda_{jk}(T|\delta_j,\psi_i)\right).$$

For the subsystem that has only one component, the maximum likelihood estimator (MLE) is obtained of closed-form and is given by

$$\hat{\beta}_{jk} = \frac{n_{jk}}{\sum_{i=1}^{n} I(\delta_i = j, \psi_i = k)\log\left(\frac{T}{t_i}\right)} \qquad \text{and} \qquad \hat{\alpha}_{jk} = n_{jk},$$

we observe that the expressions obtained are the same as those found in Louzada *et al.* (2019). But when the subsystem presents r_{jk} components in parallel, the MLE is not obtained of closed-form. Parameter estimates are obtained using iterative methods. The log-likelihood function is given by

$$\begin{aligned}
\ell(\alpha_{jk},\beta_{jk}|t,\boldsymbol{\delta},\boldsymbol{\psi}) &= \sum_{i\in I^*}\lambda_{jk}(t_i|\delta_i,\psi_i) - \Lambda_{jk}(T|\delta_j,\psi_k),\\
&= \sum_{i=1}^{n} I(\delta_i = j,\psi_i = k)\left\{\log(r_{jk}) + (r_{jk}-1)\log\left[1-\exp\left(-\left(\frac{t_i}{T}\right)^{\beta_{jk}}\alpha_{jk}\right)\right]\right.\\
&+ (\beta_{jk}-1)\log(t) - \beta_{jk}\log(T) + \log(\beta_{jk}) + \log(\alpha_{jk}) - \alpha_{jk}\left(\frac{t_i}{T}\right)^{\beta_{jk}}\\
&\left.- \log\left\{1-\left[1-\exp\left(-\alpha_{jk}\left(\frac{t_i}{T}\right)^{\beta_{jk}}\right)\right]^{r_{jk}}\right\} + \log\{1-[1-\exp(-\alpha_{jk})]^{r_{jk}}\},
\end{aligned} \tag{5.1}$$

where I^* contains the indexes $i = 1,\ldots,n$ with $\delta_i = j$ and $\psi_i = k$.

Let $\widehat{\boldsymbol{\theta}} = (\widehat{\beta}_{11},\widehat{\alpha}_{11},\ldots,\widehat{\beta}_{pk_j},\widehat{\alpha}_{pk_j})$ be the MLE of $\boldsymbol{\theta} = (\beta_{11},\alpha_{11},\ldots,\beta_{pk_j},\alpha_{pk_j})$. Under some standard regularity conditions, the MLE $\widehat{\boldsymbol{\theta}}$ is consistent and follows a normal joint asymptotic distribution with mean $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$. For a large sample size, we have that $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}) = \boldsymbol{I}^{-1}(\widehat{\boldsymbol{\theta}})$, where $\boldsymbol{I}^{-1}(\widehat{\boldsymbol{\theta}})$ is the inverse expected Fisher information matrix.

For the particular case where all r_{jk} are equal to one, Louzada *et al.* (2019) obtained the inverse expected Fisher information matrix closed-form. But when $r_{jk} > 1$, it is not possible. However, the observed information matrix, $\boldsymbol{H}(\widehat{\boldsymbol{\theta}})$, is an approximation of the expected Fisher information matrix. In addition, it can be easily obtained using computational methods. Therefore, the asymptotic confidence interval for each θ_i with $(1-\gamma)100\%$ are given by

$$\left(\widehat{\theta}_i - z_{\gamma/2}\sqrt{h_{ii}^{-1}(\widehat{\boldsymbol{\theta}})};\ \widehat{\theta}_i + z_{\gamma/2}\sqrt{\widehat{h}_{ii}^{-1}(\widehat{\boldsymbol{\theta}})}\right),$$

where $\widehat{\theta}_i$ is the MLE of θ_i, $\widehat{h}_{ii}^{-1}(\widehat{\boldsymbol{\theta}})$ is the ith diagonal element of the inverse observed Fisher information matrix, and $z_{\gamma/2}$ denotes the quantile of the standard normal distribution leaving a probability to the right tail with $\gamma/2$.

5.4 Simulation Study

To evaluate the performance of the parameters estimators of the proposed model, in this section we present the results of the Monte Carlo simulation study. For this, we consider the results described by Rigdon and Basu (2000), which establish that the random variables $T_1 < T_2 < \ldots < T_n$ are distributed as n-th order statistics with cumulative distribution function given by

$$G(y) = \begin{cases} 0, & y \geq 0 \\ \Lambda(t)/\Lambda(T), & 0 < t \geq T \\ 1, & t > T \end{cases}$$

This allowed us to establish the scheme for generating random values described by Algorithm 4.

We considered some scenarios, contemplating five values for the parameter $\beta = \{0.5, 0.75, 1.0, 1.25, 1.5\}$ and $\alpha = \{5, 10, 15, 20, 25\}$, in a context with $r = \{1, 2, 3, 4, 5\}$. The simulations were programmed using the R software (R Core Team, 2020), considering $B = 5000$ Monte Carlo replicas and the criteria used to evaluate the performance of the estimators were the MRE, RMSE and 95% CP, calculated using the following expressions

$$\text{MRE}(\hat{\theta}_i) = \frac{1}{B}\sum_{j=1}^{B}\frac{\hat{\theta}_i^{(j)}}{\theta_i}, \qquad \text{RMSE}(\hat{\theta}_i) = \sqrt{\frac{1}{B}\sum_{j=1}^{B}\left(\hat{\theta}_i^{(j)} - \theta_i\right)^2},$$

and

$$\text{CP}(\hat{\theta}_i) = \frac{1}{B}\sum_{j=1}^{B}\mathbb{1}\left(\theta_i \in (a_i^{(j)}, b_i^{(j)})\right) \qquad \text{where} \qquad \begin{aligned} a_i^{(j)} &= \hat{\theta}_i^{(j)} - 1.96 \times \text{SE}(\hat{\theta}_i^{(j)}) \\ b_i^{(j)} &= \hat{\theta}_i^{(j)} + 1.96 \times \text{SE}(\hat{\theta}_i^{(j)}) \end{aligned}$$

where θ_i is the i-th component of the vector $\boldsymbol{\theta} = (\alpha, \beta)$ and $\hat{\theta}_i$ is the corresponding maximum likelihood estimator. The function $\mathbb{1}(\cdot)$ is the indicator function, which returns the value one, if its argument is true and returns zero, otherwise. In addition $\text{SE}(\hat{\theta}_i^{(j)})$ denotes the standard error of the estimator $\hat{\theta}_i^{(j)}$.

Input:

T, p

$\boldsymbol{k} = (k_1, \ldots, k_p)$

$\boldsymbol{\beta} = (\beta_{11}, \ldots, \beta_{pk_p})$

$\boldsymbol{\alpha} = (\alpha_{11}, \ldots, \alpha_{pk_p})$

$\boldsymbol{r} = (r_{11}, \ldots, r_{pk_p})$

Output:

$\{(\boldsymbol{t}, \boldsymbol{\psi})\} = \{(\boldsymbol{t}_{11}, \boldsymbol{\psi}_{11}), \ldots, (\boldsymbol{t}_{pk_p}, \boldsymbol{\psi}_{pk_p})\}$ where $(\boldsymbol{t}_{jk}, \boldsymbol{\psi}_{jk}) = \{(t_{1jk}, \psi_{1jk}), \ldots, (t_{n_{jk}jk}, \psi_{n_{jk}jk})\}$

Procedure:

for $j := 1$ **to** p **do**
  for $k := 1$ **to** k_j **do**
    $$n_{jk} \sim \text{Poisson}(\alpha_{jk})$$
    for $i := 1$ **to** n_{jk} **do**
      $$U_{ijk} \sim \text{Uniform}(0, 1)$$
      if $r_{jk} = 1$ **then**
        $$t_{ijk} = T\, U_{ijk}^{1/\beta_{jk}}$$
      else
        $$t_{ijk} = \left(-\frac{T^{\beta_{jk}}}{\alpha_{jk}} \log \left\{ 1 - \left[1 - \exp\left(U_{ijk} \log\left(1 - (1 - \exp(-\alpha_{jk}))^{r_{jk}} \right) \right) \right]^{1/r_{jk}} \right\} \right)^{1/\beta_{jk}}$$
      end
    end
  end
end

Algorithm 3: Generator of random numbers from the proposed model.

Since the value of $\alpha = \Lambda(T)$ denotes the average number of observed failures and, therefore, more information to be used in the estimation process. In this sense, according to the criteria used, it is expected to observe that the values of MRE are close to zero, with an increase of α, while the values of RMSE should approach zero in the case of the estimate of β and $\sqrt{\alpha}$ in the case of the estimate of α (since $\alpha \approx N(T) \sim \text{Poisson}(\Lambda(T))$). On the other hand, with numerous experimental repetitions (represented by B) using a 95% confidence level, we expect that the relative frequency of the intervals that cover the true values should be close to 95%.

The results obtained for the evaluation of the estimators of β and α are shown in Figures 33 and 34, respectively. It is possible to see the empirical values for the three adopted criteria (MRE, RMSE and CP, in the lines of matrix plot), while the scenarios are organized in the columns of matrix plot (for β), in the horizontal axis (in the case of α) and colors (for r).

The behavior of both estimators showed the expected behavior. The highlight is the estimator of β which, in very small sample contexts ($\alpha < 10$), presents significantly large values for the MRE (indicating an overestimation of the true value of β) and also for RMSE (indicating some degree of uncertainty).

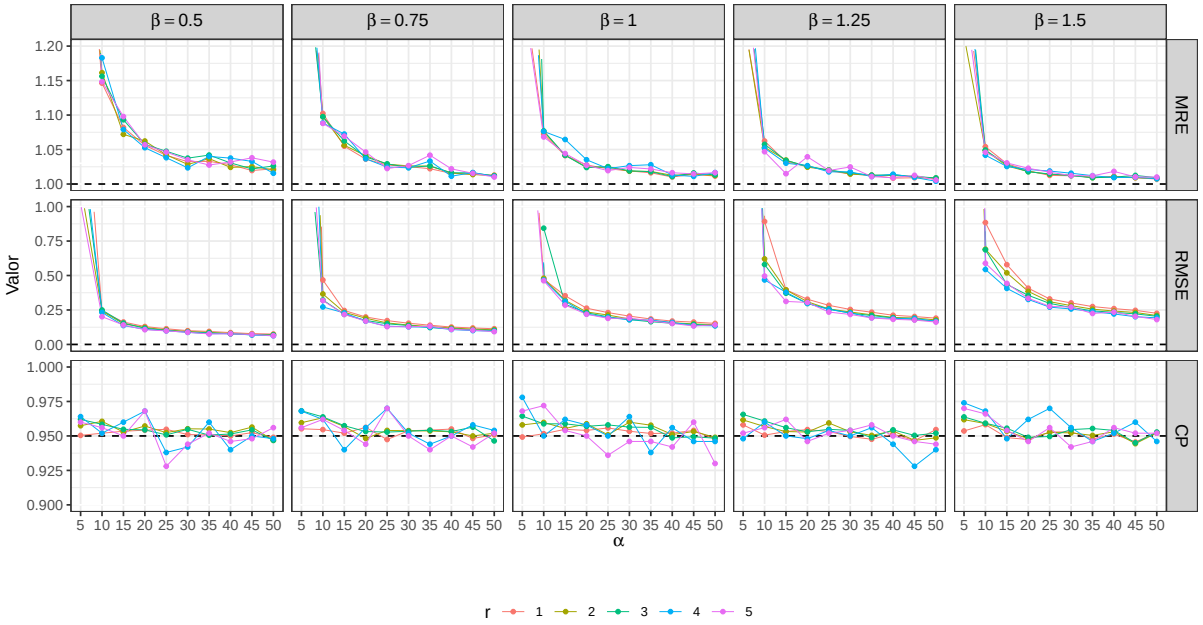


Figure 33 – Computer simulation results for the β estimator, considering the scenarios $\beta = \{0.5, 0.75, 1.0, 1.25, 1.5\}$, $\alpha = \{5, 10, 15, 20, 25\}$ and $r = \{1, 2, 3, 4, 5\}$, with 5000 Monte Carlo replicates.

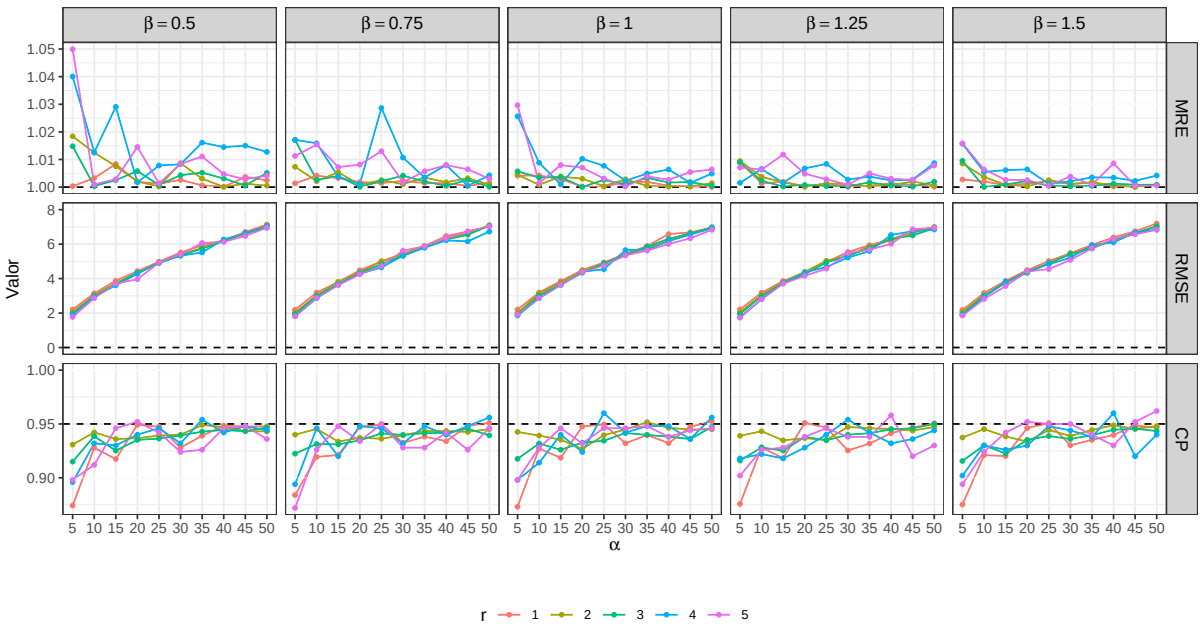


Figure 34 – Computer simulation results for the α estimator, considering the scenarios $\beta = \{0.5, 0.75, 1.0, 1.25, 1.5\}$, $\alpha = \{5, 10, 15, 20, 25\}$ and $r = \{1, 2, 3, 4, 5\}$, with 5000 Monte Carlo replicates.

Another interesting question to be investigated in the simulation study is how estimators behave in the hierarchical context. To assess whether we are able to recover reliability measures over time, we have structured six scenarios of hierarchical systems ordered according to the increase in complexity. As an example, the first proposed structure (Figure 35a) presents a general system composed of two systems which, in turn, are composed of one (with 4 components in parallel and $\beta = 1.5$) and two subsystems (with 2 and 1 components in parallel and $\beta = 0.75$ and 0.5, respectively). In the second structure (Figure 35b), we kept the previous configuration and added a subsystem (with 3 components and $\beta = 1.75$) to the second system. The logic was maintained for the other structures until the sixth structure (Figure 35f) totaled 20 components.

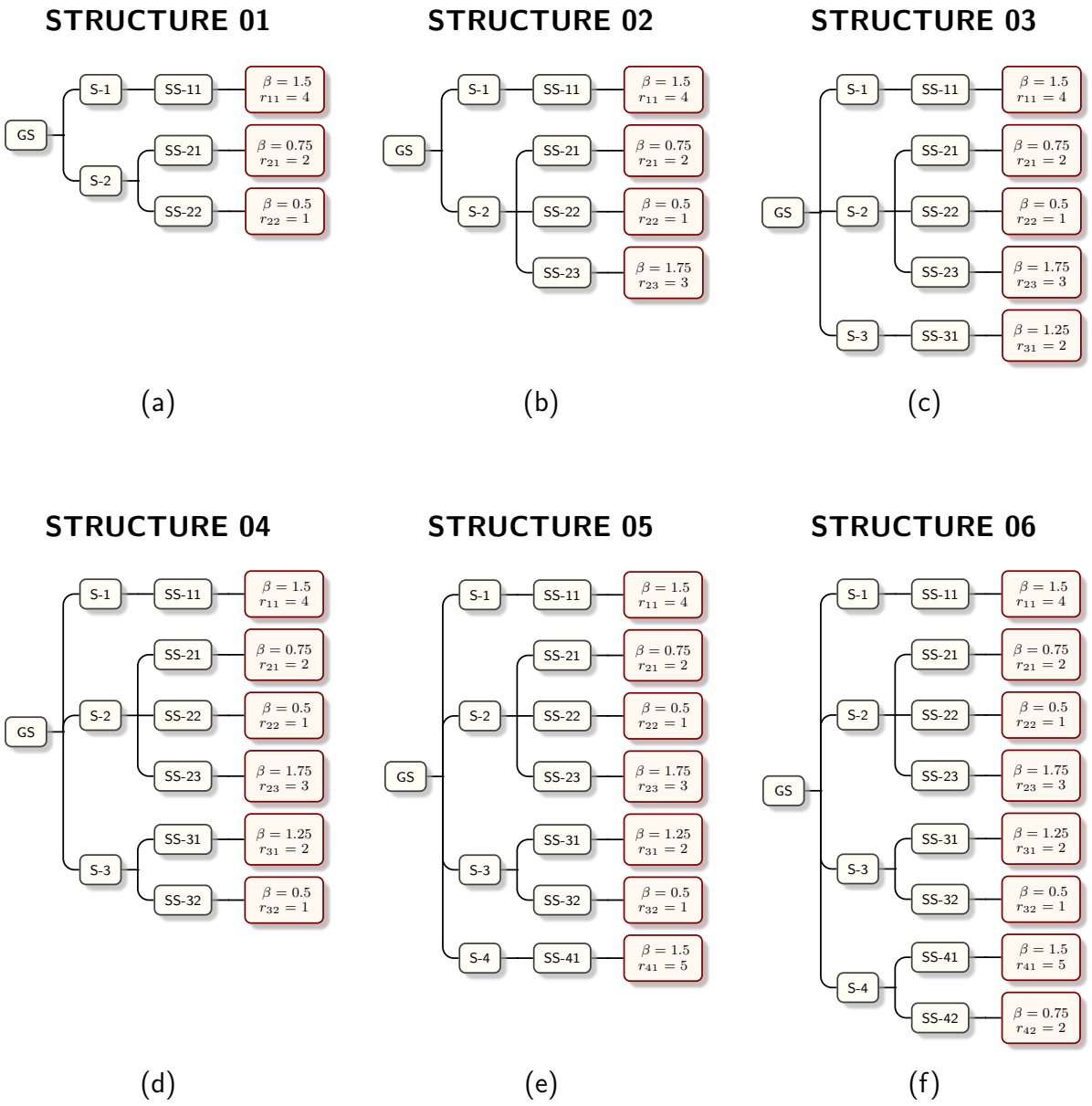


Figure 35 – Scenarios proposed according to the increase in complexity from (a) to (f).

We compared the estimated reliability in 100 equally spaced time points ($\boldsymbol{t}^* = (t_1, \ldots, t_{100})$, from 0 to T) with the actual reliability. We determine the ratio $\hat{R}(t)/R(t)$ in each time $t \in \boldsymbol{t}^*$ considered and the results obtained for the scenarios exposed above, considering several scenarios of $\alpha = (5, 10, \ldots, 50)$, are shown in Figure 36. In it we can see that the increase in complexity has the expected effect on the reliability estimation, since occurs the propagation of errors made in the individual estimates of each component. On the other hand, the increase of α (and, therefore, the average number of occurrences) causes the estimates to become closer to reality, as expected.

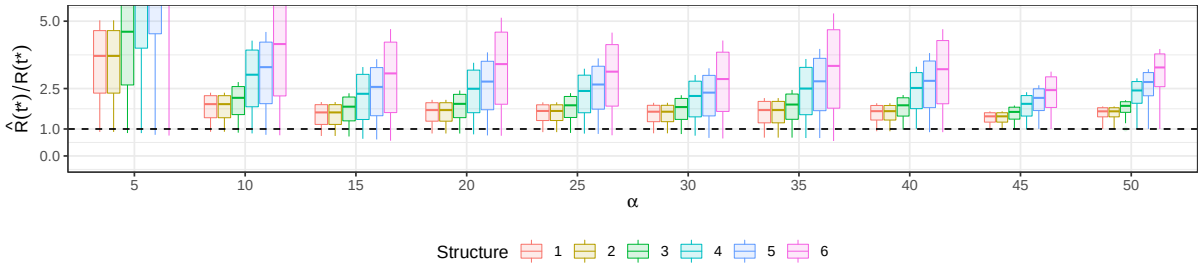


Figure 36 – Comparison between the real and estimated reliability according to the increase in both, the complexity of the system (Figure 35) and the average number of failures until T ($\mathbb{E}(N(T)) = \alpha$).

5.5 Application

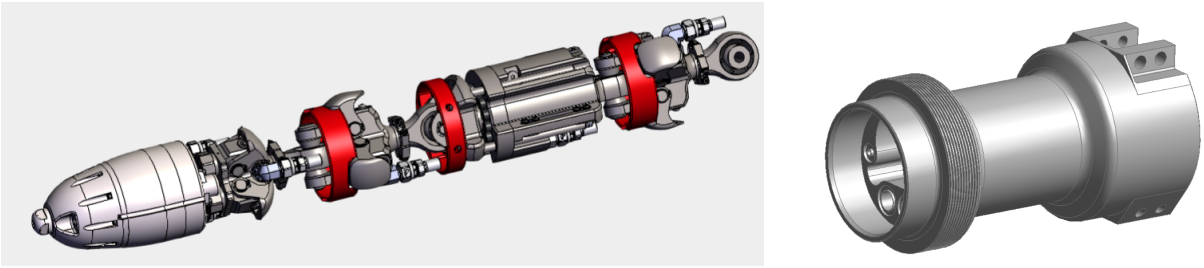
In this section, we continue the studies carried out on a robotic unit already presented in previous works such that Louzada *et al.* (2019), the Annelida project. Annelida is an innovation project whose development is carried out in partnership between Petrobras (a Brazilian oil and gas company) and six brazilian research institutes. It is a robot built to travel inside oil ducts in order to remove hydrates and paraffins that prevent the oil and gas flow, causing damage to the production lines. Among the stages already developed in the project, one of them consisted of studying the main fatigue mechanisms that could interrupt the operation of the robotic unit. The purpose of this section is to expose the results from CAE simulation, associating them with the perspective of statistical reliability.

5.5.1 Numerical simulations

We used numerical simulations to study the fatigue life of some Annelida's components. To evaluate this phenomenon, an extreme case was studied: the maximum load ($20000N$) was applied to the system, and the life of two different components were evaluated.

During the robot movement, paws act as anchor points, while the hydraulic cylinder has the role of creating the movement through successive expansion and contraction. This generates a cyclic load in the paws and other structural components, that makes it susceptible to fatigue failure. This means that a crack can start in a region of concentrated stress and

propagate, until a point of mechanical failure (BOARDMAN, 1990). In this case, the frontal paws (Figure 37a) are particularly problematic, because they face the bigger workload. The other system studied was a support in the interface between electrical cables (Figure 37b).



(a) (b)

Figure 37 – Annelida's frontal motion system (a) and; support located between electrical cables and the robot body (b).

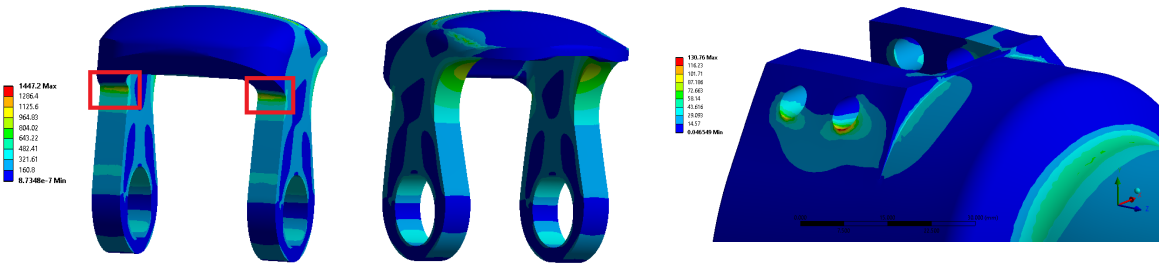
The simulations were performed using the commercial software ANSYS Workbench 14.5 and they proceeded as: (i) a static structural analysis was performed to obtain principal stress and strain; (ii) a strain-life ($\varepsilon - N$) curve was used to analyze the component life and the Morrow correlation was also used to compensate for the fact that the load is not completely reversible.

In the simulations, we also accounted for the variations in material properties. Normally, mean values are used in evaluations such as this, however (PARK; PARK; KIM, 2008; PARK; PARK; KIM, 2010) showed that some properties, such as tensile strength and elasticity modulus are well described by a normal distribution, and their standard deviation can be up to 5% of the mean value. Therefore, we assumed a distribution of the same type for the parameters associated with the $\varepsilon - N$ curve:

- Strength coefficient and Strength exponent;
- Ductility coefficient and Ductility exponent;
- Cyclic strength component and Cyclic strain hardening exponent;

This leads to different values of life depending on the material properties or a distribution of possible life time for each component. The material of the paws is AISI 4340 steel, quenched and tempered - physical properties obtained from (BAUCCIO *et al.*, 1993) and (DURAN; COSTA; JUNIOR, 2018). While the support is made of the alluminum alloy 7570-T651 - physical properties obtained from (INCE; GLINKA, 2011).

Figure 38a shows images of the front and back of one paw and the stresses calculated. The regions highlighted in the figure are the points with the highest stress and also the lower life. The same results are shown for the support in Figure 38b. In this case, there is a peak stress in the vicinity of one pinned connection.



(a) (b)

Figure 38 – Calculated stress in the paws - red squares highlight the points with higher stress concentration and lower stress life.

5.5.2 Reliability Model

The studied system corresponds to the interaction between the components with greater exposure to mechanical fatigue. The system hierarchy can be seen in Figure 42. According to the results of the CAE simulation study, the stress that some systems will be subjected to is not enough to cause any damage that arouses concern, so their lifetimes are not influenced by the fatigue mechanism. In this context, in the Figure 42, systems and components in red will be disregarded in the reliability study exposed here.

The Forward Locomotive, unlike the Return Locomotive, is exposed to a high level of stress with high concentration on the front legs (Figure 38a), which make up the Claw System of the robotic unit. The assumption we make here is that this system has two Front Claws Set acting in parallel and one Back Claw Set (whose low stress suffered was disregarded in the study), thus, the first configuration of interest is given by $(j = 1, k = 1, r_{11} = 2)$. The Umbilical Cable System has a Load Cell in which a support at the interface between the electrical cables, Load Cell Suport in the Figure 38b, is also subject to worrying mechanical stress, its configuration is $(j = 2, k = 1, r_{21} = 1)$.

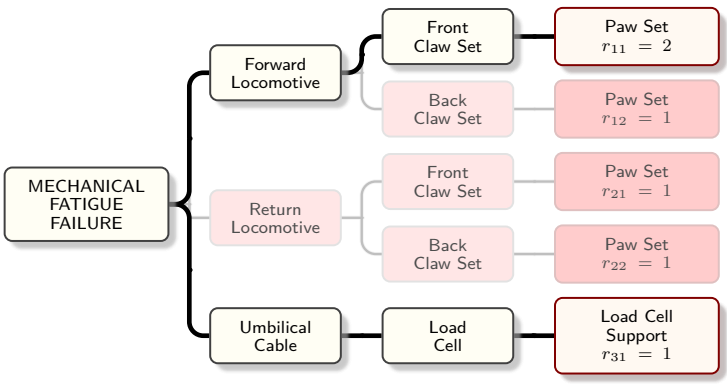


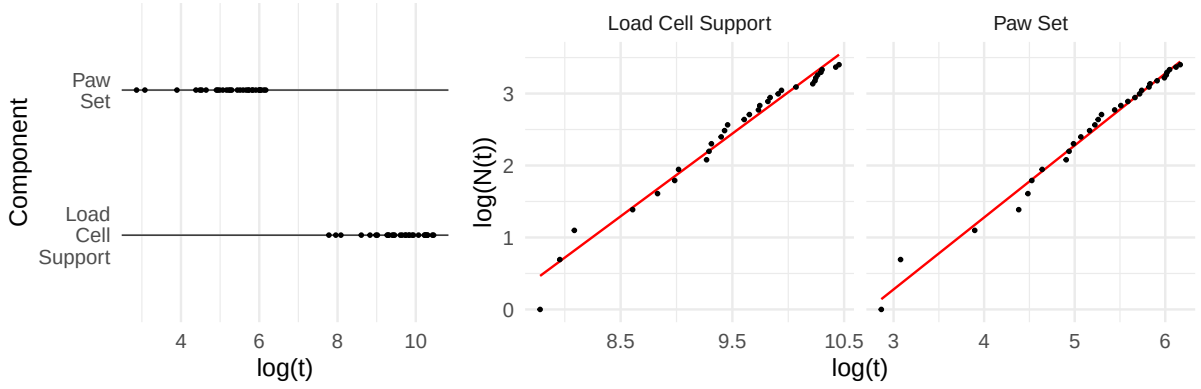
Figure 39 – Failure system associated with stress caused by mechanical fatigue. In red fill, there are the systems/ subsystems/components whose mechanical efforts were not enough to significantly reduce the life time, in the simulation study. The others showed failure behaviors influenced by mechanical stress.

The results obtained in Section 5.5.1 are given in terms of the "number of cycles performed by the systems until they reach stress levels that cause a failure". For convenience, we use a time scale in days of operation, considering that each cycle occurs in approximately three seconds. We also consider the occurrence of 30 failures, in this context, the last failure observed for the "Paw Set" system was approximately 475.99 days while the other was close to 34695.47, which already indicates a large discrepancy regarding the susceptibility of the systems to the mechanical stress. The simulated failure times were as follows:

Table 22 – Failure times (in days) obtained from the CAE simulation study, for mechanical fatigue.

Component	Failure Times
Paw Set	17.54, 21.7, 49.22, 79.99, 88.54, 92.45, 103.65, 135.14, 139.52, 146.5, 158.89, 175.02, 185.52, 192.4, 199.63, 230.76, 247.24, 267.1, 288.89, 303.99, 310.98, 337.17, 341.32, 368.85, 399.59, 408.14, 412.05, 423.24, 454.63, 475.99
Load Cell Support	2390.8, 2852.05, 3248.4, 5472.43, 6837.88, 7972.64, 8254.21, 10604.59, 10842.3, 11074.24, 12088.16, 12459.65, 12802.44, 14849.97, 15554.1, 16872.99, 17072.6, 18363.99, 18733.43, 20128.33, 20730.72, 23617.84, 27429.3, 27986.42, 28190.14, 28654.82, 29508.89, 29813.8, 33686.72, 34695.47

When we observe the failure times by component (Figure 40a), we notice the great discrepancy between the occurrence times for the first thirty simulated failures. While the last occurrence of the Paw Set occurred at approximately 476 days, the first occurrence of the Load Cell Support component occurred at approximately 2391 days, which expresses the great distinction between reliability. In addition, an important assumption in the construction of the proposed model is that the intensity of failures is governed by a power-law. An indication that this law is appropriate to adjust the intensity of failures can be captured by Duane's graphs (Figure 40b) for each component. A linear behavior, like the one observed, suggests that the failure behavior is appropriately captured by the model proposed here.



(a) (b)

Figure 40 – (a) Descriptive graph for fatigue failure times for each component and (b) Duane plot for each component.

Since, in the context of the CAE simulation, the number of simulated failures was pre-defined for each component, we are considering failure truncation. In this sense, it is sufficient that the value T be replaced by t_{30} in Equation (5.1) and, therefore, once estimated, $\hat{\alpha}$ will denote the estimative of the expected number of failures until t_{30}. The Table 23 shows the estimatives for the model parameters, based on 5000 non-parametric bootstrap replicates. We opted for the bootstrap approach to more conveniently construct the intervals for the functions $\Lambda(t)$ and $R(t)$ for each time t.

Table 23 – Point and interval estimates obtained for the parameters of the model described in the Section 5.2 applied to the CAE simulated data in the Section 5.5.1.

			Bootstrap IC95%	
	Estimative	Std. Error	Lower	Upper
β_{11}	0.989	0.132	0.781	1.27
α_{11}	30.8	0.379	30.4	31.1
β_{31}	1.08	0.174	0.784	1.48
α_{31}	29.7	1.05	25.6	30.1

The graphs in Figure 41 show the estimate for the expected number of failures overlapping the observed number of failures (Figures 41 a and c) from which we derive indications of the quality of the model's fit. On the other hand, in Figures 41 b and d, the reliability associated with the next failure time $(T_{n+1} \mid t_n, \ldots, t_1)$ is exposed, and from them, we can extract the median time until the next occurrence. Particularly, for the system Paw Set, the estimated median time is close to 10.8 (7.52–14.57), while for the system Load Cell Support, the estimate is close to 743.27 (459.65–1065.31).

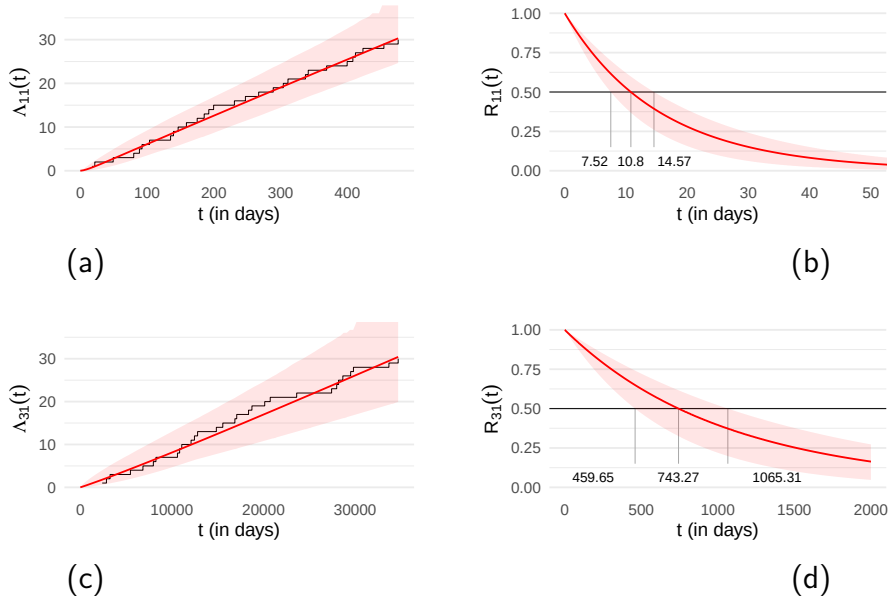


Figure 41 – Point and interval (IC95%) estimatives for the functions $\Lambda_{11}(t), \Lambda_{31}(t), R_{11}(t)$ and $R_{31}(t)$ and for the median times to the next failure.

A next step would be to assess how the reliability of the components $C11$ and $C31$, together, make up the reliability of the system associated with "Mechanical Fatigue Failure" (S0). The result of this interaction is shown in Figure 42. We can see that the high reliability estimated for the $C31$ component has little influence on the overall system, so that its reliability is mostly dependent on the reliability of the $C11$ component. Thus, considering only the stress caused by mechanical fatigue, the median system failure time is close to 10.6 (7.6–14.2).

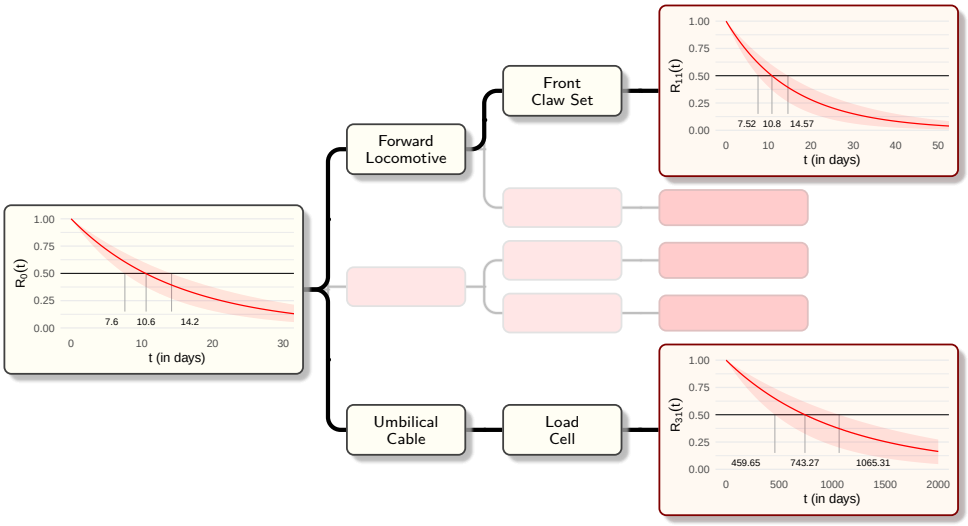


Figure 42 – Reliability results for the failure system associated with stress caused by mechanical fatigue.

It is also important to note that these results consider the exposure of the robotic unit to a stress level much higher than the levels it will be exposed to in practice. Such results can be used to set target levels for the acceptable limits of reliability of the studied systems, so that in later stages the proposed practical tests can guarantee this confidence index in the equipment.

CHAPTER

6

FRAILTY MODEL FOR MULTIPLE REPAIRABLE SYSTEMS HIERARCHICALLY REPRESENTED SUBJECT TO COMPETING RISKS

In this chapter, we propose a statistical model to understand the behavior of failure times associated with groups of repairable systems hierarchically represented, under a competing risks framework, with the consideration of the existence of unobserved heterogeneity that acts individually on the systems of each group, and also the possibility of IRs whose initial failure rate is in the form of the power-law. To illustrate the use of the proposed model, we consider a database with the failures of 38 agricultural machines categorized in five different groups. As an illustration, we understand that the tractor fleet corresponds to the farm's agricultural system, so that the need for intervention in this system occurs with the failure of any unit, individually, in a serial structure, of competing risks. On the other hand, the understanding of the time in which all the machinery that makes up the fleet will have required some intervention is obtained by analyzing the results under a parallel structure.

Contents of Chapter

6.1 Introduction

6.1.1 Practical Motivation

Due to the continuous search for minimizing the use of natural resources and reducing operating costs today, it is of great interest to organizations to optimize the maintenance of equipment and machines in operation. The prolonged and effective operation of equipment to support agricultural operations, such as tractors and implements, is a vital requirement to achieve cost reduction. One of the factors that directly impacts these operations is the proper maintenance planning.

Preventive maintenance is a way to keep the system operating at peak performance, avoiding possible failures during the operation of agricultural tractors, the object of study in this work. On the other hand, there is also corrective maintenance, which occurs only after the unexpected failure of the machine, in order to make it productive again, which may cause it to stop longer.

From an operational point of view, the need to know and control the greatest number of possible failures that can compromise the execution of the work is imposed and, uniting with an efficient management of these indicators, it is possible to prolong the operations activity in maximum performance and without being affected by defects or prolonged shutdowns. Thus, although preventive maintenance is the most suitable for reducing costs and increasing productivity, in real cases, there is a very high incidence of corrective measures.

The preventive maintenance model is more efficient when equipped with the prediction of breakage, informing when the machines will present possible failures. In order to use such a tool there is a need to perform initial diagnostics, being able to obtain the status of each part of the machine, when they will fail and how their replacement is programmed before the break. Therefore, the theme of the project in question will involve the use of reliability models so that it is possible to investigate and define the necessary action to avoid defects in agricultural tractors during operation.

The success of harvesting fruits, vegetables and grains depends fundamentally on the agricultural processes carried out during planting during the ripening period. A good harvest

is a function of operations involving plant nutrition, weed control and spraying, for example. Operations such as pruning and clearing, on the other hand, are responsible for maintaining the orchard. Thus, for the execution of these operations, rural producers make use of tractors, carrying them out mechanically.

Considering this, it is reasonable that tractors depend on preventive and corrective maintenance so that they can carry out agricultural activities on an approximately continuous basis. Therefore, it is necessary that there are specific processes and methodologies of reliability analysis in view of the useful life of agricultural machines, aiming at preventive maintenance to guarantee the operation with the least possible number of machine defects.

6.1.2 Statistical Background

We understand a system as a hierarchical group with one or more levels of subsystems and/or components. With a well-defined structure, the relationship between components, subsystems and the system in general is highlighted, bringing light to the importance of parts of the system over the whole. This allows quantifying levels of damage, costs and impacts caused by external sources of variation (LIEPING; ZHE, 2009). In a classic definition, a so-called repairable system is a system that, after the occurrence of a failure, that prevents it from performing one or more functions satisfactorily, can undergo a repair action that restores it, without the need for its complete replacement (ASCHER; FEINGOLD, 1984; RIGDON; BASU, 2000).

Statistical models to understand the event recurrence process have been used in numerous industrial problems, such as natural gas regulating and metering stations (BA-HOOTOROODY *et al.*, 2020), traction systems of trains (NAVAS; SANCHO; CARPIO, 2017), aero engines (SHARMA; RAI, 2020) and other machine tools (PENG *et al.*, 2018; WANG; PAN, 2021). In the context of complex repairable systems, it is important to consider the occurrence of failures in terms of the hierarchical representation given to the system, considering the multiplicity of failure modes, such as the work (ALMEIDA, 2019; MULLOR; MULERO; TROTTINI, 2019; SYAMSUNDAR; NAIKAN; WU, 2020; YAN, 2019).

The representation in groups of systems/subsystems/components in series, takes us to a context of competing risks (WU; SCARF, 2017; TODINOV, 2015; SOMBOONSAVATDEE; SEN, 2015), on the other hand, groups in parallel give us freedom for other types of interpretation. Once the registration of the subsystem and/or component that caused the failure is commonly performed, it makes no sense to fail to consider this in the statistical analyzes. Good examples of this are (LANGSETH; LINDQVIST, 2006) in the industrial context and (TULI *et al.*, 2000) in the context of child health.

In general, it is assumed that failures are governed by a NHPP with the intensity described by a power-law, i.e. the PLP, a reasonably flexible model, easy to implement and

interpretable. This assumption, however, implies the hypothesis of a MR model (RAUSAND, 2003), which may not be the most realistic choice, in this sense, a class of IR models brings the possibility of considering a intermediate level of repair, somewhere between MR and PR. This idea is well exposed in works such as those of (KIJIMA; MORIMURA; SUZUKI, 1988; DOYEN; GAUDOIN, 2004; TOLEDO *et al.*, 2015).

Doyen and Gaudoin (2004) were responsible for proposing two of the main models for IR, the ARA and ARI model classes, whose repair efficiency is quantified by one of the model parameters. In addition, a constant characterizes the models in terms of the used memory order, which refers to the maximum number of previous information involved in the model's calculations.

Another common assumption is the absence of any unobserved heterogeneity among the causes that compete for the failure of the general system. An immediate consequence of this assumption is the disregard of effects that, although not observed, can have effects on the behavior of failure times (WIENKE, 2010). The literature that adequately addresses this issue is related to frailty models. A review of these models can be found at (HOUGAARD, 1995; ANDERSEN *et al.*, 1996; WIENKE, 2010). From the perspective of multiplicative frailty, we introduce a multiplicative factor (a random variable) that inflates, deflates or preserves the failure intensity function.

6.1.3 Overview

The next sections of this chapter are organized as follows: In the Section 6.2 we present the modeling framework, exposing its notation, assumptions and hierarchical representation. In the section 6.3 we present the deduction of the likelihood function that summarizes the proposed model, suggest the maximum likelihood method for estimating the parameters and obtaining the errors associated with the estimates. In the Section 6.4 we present the method of generating random values from the model, the criteria used to evaluate the quality of the estimators and the results obtained in the simulation process. In Section 6.5 we illustrate the use of the method with the analysis of a real data set regarding the monitoring of the reliability of a tractor fleet on a farm. Finally, in Section 6.6 we make some final considerations and suggest some steps to expand this study in terms of modeling.

6.2 Proposed Model

The methodology proposed in this work follows the works already discussed in the Sections (3, 4 and 5), taking new steps to some studies carried out by Almeida (2019), DAndrea (2020) and Lindqvist (2006). In the context previously developed, we consider the study of the reliability of repairable systems, with the function of the Powers Law intensity, hierarchically represented, under the assumption of competing risks and subject to MRs and arranged in the

form of a serial and/or parallel hierarchy. The next step is with the consideration of multiple systems that can share unobservable factors that cause heterogeneity in the behavior of the occurrence of system failures.

In this case, we aim to analyze the behavior of the failure times of a cluster of repairable systems that, together, represent a unit that serves some activity of practical interest. In this context, from the perspective of a context of competitive risks, the failure of each of these systems represents a necessary intervention on the unit. If we want to understand probabilistically the time in which the entire fleet will be repaired at least once, we can operate in a context of systems in parallel, in which case the unit will no longer be serviced (in any case) only if all units fail to function.

We represent this perspective hierarchically (Figure 43), considering the unit as the "General System". The second level of this representation denotes G groups (in any sense of interest) of systems that exhibit similar failure behaviors or that share other characteristics of practical interest. The g-th (with $g = 1,\ldots,G$) group comprises U_g systems, particularly the (g,u)-th (with $u = 1,\ldots,U_g$) system presents $n_{g,u}$ corrective interventions, due to the occurrence of unforeseen failures.

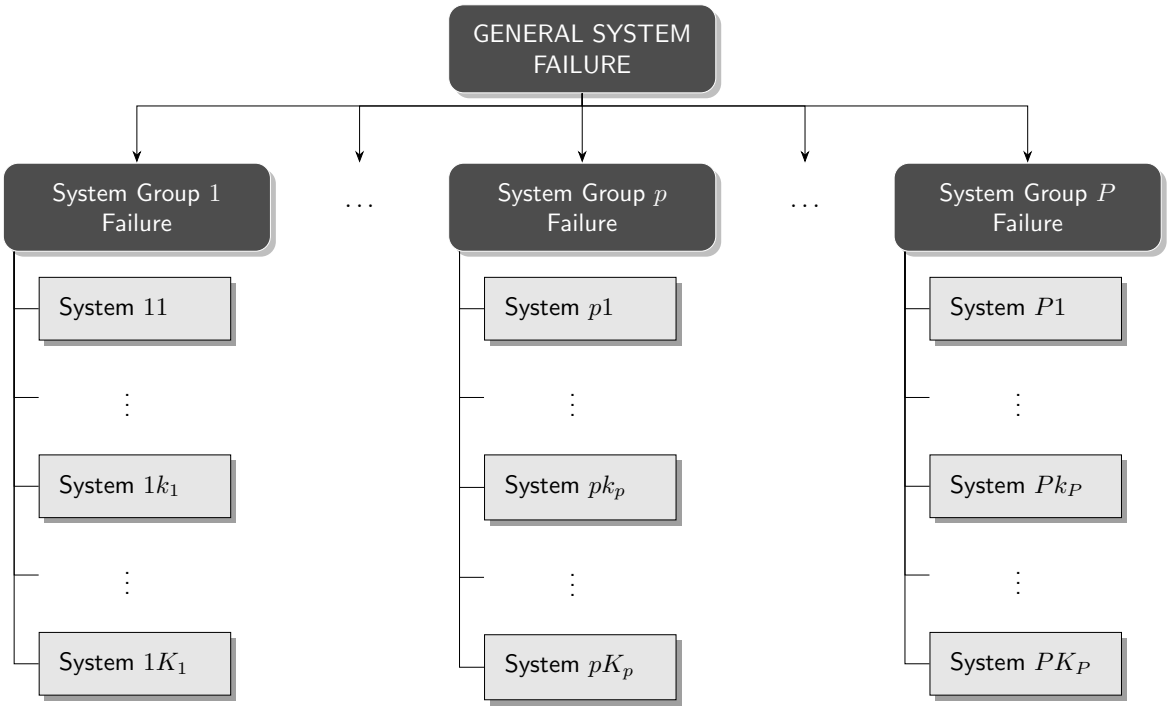


Figure 43 – Illustration of the structure explored in this study.

A dataset that can be used under this structure is denoted as

$$(\boldsymbol{t},\mathcal{G},\mathcal{U}) \qquad \text{where} \qquad \begin{aligned} \mathcal{G} &= (1,\ldots,g,\ldots,G), \\ \mathcal{U} &= (\mathcal{U}_1,\ldots,\mathcal{U}_g,\ldots,\mathcal{U}_G), \\ \boldsymbol{t} &= (\boldsymbol{t}_1,\ldots,\boldsymbol{t}_g,\ldots,\boldsymbol{t}_G), \end{aligned}$$

and yet

$$\begin{aligned}\mathcal{U}_g &= (1,\ldots,u,\ldots,U_g),\\ \boldsymbol{t}_g &= (\boldsymbol{t}_{g,1},\ldots,\boldsymbol{t}_{g,u},\ldots,\boldsymbol{t}_{g,U_g})\end{aligned} \qquad \text{where} \qquad \boldsymbol{t}_{g,u} = (t_{g,u,1},\ldots,t_{g,u,i},\ldots,t_{g,u,n_{g,u}}).$$

The 3-tuple $(\boldsymbol{t},\mathcal{G},\mathcal{U})$ denotes, respectively, the set of all recorded failure times $(\boldsymbol{t})$, the indicators the group to which the system that failed occurred $(\mathcal{G})$ and the identification of the system that required an intervention on the unit as a whole $(\mathcal{U})$.

We assume that the systems of a particular group share the same default behavior behavior. In this sense, the functions that define its reliability are associated with the g-th group: $\underline{\lambda}_g(t;\boldsymbol{\theta})$, $\underline{\Lambda}_g(t;\boldsymbol{\theta})$ and $\underline{R}_g(t;\boldsymbol{\theta})$, respectively, the failure intensity function, the accumulated failure intensity (or the mean value function) and the reliability function. The parameter vector $\boldsymbol{\theta}$ denotes the parameters that index the model and that characterize the respective behaviors. The symbols $\underline{\lambda}$, $\underline{\Lambda}$ and $\underline{R}$ try to emphasize the idea of "default behavior".

6.2.1 The frailty model framework

Even if the systems of a particular group have the same default behavior, there may be unobservable factors that all systems in that group are subject to, but which influence their respective behaviors individually. To considere this unobserved heterogeneity, we consider the model of multiplicative frailty. This model is defined by the conditioned versions of the intensity, accumulated intensity and reliability functions, generically and respectively expressed by

$$\lambda(t|z) = z\underline{\lambda}(t) \quad \Longleftrightarrow \quad \Lambda(t|z) = z\underline{\Lambda}(t) \quad \Longleftrightarrow \quad R(t|z) = e^{-\underline{\Lambda}(t)z},$$

where z denotes the frailty term, a multiplicative factor that inflates (when $z>1$), deflates $(z<1)$ or preserves $(z=1)$ the failure intensity function.

Commonly, the probability model adopted for the random variable Z, whose realizations are the frailty terms, is the Gamma model. In particular, the parameterization $\mathrm{Gamma}(1/\alpha,1/\alpha)$ is used, whose probability density function is expressed by

$$f(z) = \frac{(1/\alpha)^{1/\alpha}}{\Gamma(1/\alpha)} z^{1/\alpha-1} e^{-(1/\alpha)z},$$

where $\mathbb{E}(z)=1$ and $\mathrm{Var}(z)=\alpha$. This avoids the inclusion of an identifiability problem in the model and also brings the possibility of interpreting unobserved heterogeneity based on the parameter α (ELBERS; RIDDER, 1982).

Based on this construction, we can find the marginal reliability function for failure times using the Laplace transform, since the reliability $R(t)$ is expressed by

$$R(t) = \int_0^\infty R(t|z)f(z)\,\mathrm{d}z = \int_0^\infty e^{-\underline{\Lambda}(t)z} f(z)\,\mathrm{d}z,$$

and integrals written in this way have a well-known Laplace transform, given by $Q(\underline{\Lambda}(t)) = (1+\alpha\underline{\Lambda}(t))^{-1/\alpha}$ and can be used to solve this problem (WIENKE, 2010). In this sense

$$\begin{aligned} \lambda(t) &= \frac{1}{Q(\underline{\Lambda}(t))}\frac{\mathrm{d}}{\mathrm{d}t}\Big[Q(\underline{\Lambda}(t))\Big] & & \lambda(t) = -\frac{\underline{\lambda}(t)}{1+\alpha\underline{\Lambda}(t)} \\ \Lambda(t) &= -\log Q(\underline{\Lambda}(t)) & \Longleftrightarrow \quad & \Lambda(t) = -\log\big[\,(1+\alpha\underline{\Lambda}(t))^{-1/\alpha}\big] \\ R(t) &= Q(\underline{\Lambda}(t)), & & R(t) = (1+\alpha\underline{\Lambda}(t))^{-1/\alpha}, \end{aligned} \tag{6.1}$$

The value of α represents the variance of the random variable Z and, therefore, brings indications of the existence of unobserved effects that could significantly affect the intensity of failures, making some systems more susceptible to failures than others.

6.2.2 The IR framework

The choice of the default model involves both, the choice for the general behavior of the failure intensity function (such as power-law or log-linear models for example) and the possibility of considering a repair regime, as a minimum, perfect or imperfect (LINDQVIST, 2006). In this study, we are using a ARA_m model of IRs whose initial failure rate is the power-law model. In this sense

$$\underline{\lambda}(t;\boldsymbol{\theta}) = h\left(t-(1-\rho)\sum_{j=0}^{\min\{m-1,N(t)-1\}}\rho^j T_{N(t)-j}\right), \tag{6.2}$$

where $h(t) = (\beta/\eta)(t/\eta)^{\beta-1}$. And also

$$\begin{aligned} \underline{\Lambda}(t;\boldsymbol{\theta}) &= \int_0^t \underline{\lambda}(s;\boldsymbol{\theta})\,\mathrm{d}s \\ &= \int_0^t h\left(t-(1-\rho)\sum_{j=0}^{\min\{m-1,N(t)-1\}}\rho^j T_{N(t)-j}\right)\mathrm{d}s \\ &= H\left(t-(1-\rho)\sum_{j=0}^{\min\{m-1,N(t)-1\}}\rho^j T_{N(t)-j}\right) - H\left((1-\rho)\sum_{j=0}^{\min\{m-1,N(t)-1\}}\rho^j T_{N(t)-j}\right) \end{aligned} \tag{6.3}$$

where $H(t) = (t/\eta)^{\beta}$.

Finally, based on Equations in (6.1), (6.2) and (6.3), we have the IR.ARAm.F model (Frailty Power-Law Model with ARA_m for IR) for the functions λ, Λ and R. This model has as particular cases and/or limit cases, the models shown in Table 24.

Table 24 – Some cases for proposed intensity function.

$\alpha \to 0$	$\rho = 0,\ m = 1$	PR
	$\rho = 1$	MR
	$0 < \rho < 1$	ARA$_m$ IR (IR.ARAm)
$\alpha > 0$	$\rho = 0,\ m = 1$	Frailty PR (PR.F)
	$\rho = 1$	Frailty MR (MR.F)
	$0 < \rho < 1$	Frailty ARA$_m$ IR (IR.ARAm.F)

With this construction, we established a model for each group with an arbitrary number of systems, which considers unobserved effects that all of them can share, but which also individualize them. Another interesting issue to be considered is the issue of truncation, due to failures or time. If we choose to consider a truncation by time, it is interesting to make a convenient reparametrization, in terms of a parameter μ as a function of the truncation time τ, like suggested by Oliveira, Colosimo and Gilardoni (2012). In this sense, some one could consider

$$\mu = \mathbb{E}[N(\tau)] = \left(\frac{\tau}{\eta}\right)^{\beta} \qquad \Longleftrightarrow \qquad h(t) = \mu\left(\frac{\beta}{\tau}\right)\left(\frac{t}{\tau}\right)^{\beta-1}.$$

6.2.3 Reliability for the next failure time

Once the parameters are estimated, we can use this model to describe the future reliability of the system, considering the history of failures and from the perspective of the need for any intervention calculating

$$R(t \,|\, \mathcal{H}) = \prod_{g=1}^{G} R_g(t \,|\, \mathcal{H}_g) = \prod_{g=1}^{G} \left[\prod_{u=1}^{U_g} R_{g,u}(t \,|\, \mathcal{H}_{g,u}) \right]. \tag{6.4}$$

On the other hand, if we want to estimate the reliability from the perspective of the time in which all units from all groups of the systems will already require some intervention, we calculate as

$$R(t \,|\, \mathcal{H}) = 1 - \prod_{g=1}^{G} \big(1 - R_g(t \,|\, \mathcal{H}_g)\big) = 1 - \prod_{g=1}^{G} \left[\prod_{u=1}^{U_g} \big(1 - R_{g,u}(t \,|\, \mathcal{H}_{g,u})\big) \right]. \tag{6.5}$$

In this case, both in (6.4) and in (6.5), to determine $R_{g,u}(t)$ denote the total number of failures of (g,u)-th system, for simplicity, as $n = n_{g,u}$ and calculate its reliability after the

last failure as

$$
\begin{aligned}
R_{g,u}(t \,|\, \mathcal{H}_{g,u}) &= \Pr(T_{n+1} - t_n > t \,|\, \mathcal{H}_{g,u}) \\
&= \Pr(N(t_n, t_n + t) = 0 \,|\, \mathcal{H}_{g,u}) \\
&= \exp\left\{ - \int_{t_n}^{t_n+t} \lambda_{g,u}(s)\, \mathrm{d}s \right\} \qquad \text{with} \qquad t_n < s < t_n + t < T_{n+1} \\
&= \exp\Big\{ - \Big(\Lambda_{g,u}(t_n + t) - \Lambda_{g,u}(t_n) \Big) \Big\}.
\end{aligned}
$$

6.3 Inference

The estimation process adopted in this study is the Maximum Likelihood Method (PAWITAN, 2013). We consider the information as previously described and, in this sense, the likelihood function is given by

$$
\begin{aligned}
&L(\boldsymbol{\theta}; \boldsymbol{t}, \mathcal{G}, \mathcal{U}) = \\
&= \prod_{g=1}^{G} \prod_{u=1}^{U_g} \left[\prod_{i=1}^{n_{g,u}} \lambda_g(t^*_{g,u,i}) \exp\Big\{ - \Big[\Lambda_g(t^*_{g,u,i}) - \Lambda_g(t^*_{g,u,i-1}) \Big] \Big\} \right] \exp\Big\{ - \Big[\Lambda_g(\tau^*_{g,u}) - \Lambda_g(t^*_{g,u,n_{g,u}}) \Big] \Big\} \\
&= \prod_{g=1}^{G} \prod_{u=1}^{U_g} \left[\prod_{i=1}^{n_{g,u}} \frac{\underline{\lambda}_g(t^*_{g,u,i})}{1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,i})} \exp\Big\{ \log\Big[\Big(1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,i}) \Big)^{-1/\alpha_g} \Big] - \log\Big[\Big(1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,i-1}) \Big)^{-1/\alpha_g} \Big] \Big\} \right] \times \\
&\qquad \exp\Big\{ \log\Big[\Big(1 + \alpha_g \underline{\Lambda}_g(\tau^*_{g,u}) \Big)^{-1/\alpha_g} \Big] - \log\Big[\Big(1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,n_{g,u}}) \Big)^{-1/\alpha_g} \Big] \Big\} \\
&= \prod_{g=1}^{G} \prod_{u=1}^{U_g} \left[\prod_{i=1}^{n_{g,u}} \left(\frac{\underline{\lambda}_g(t^*_{g,u,i})}{1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,i})} \right) \left(\frac{1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,i-1})}{1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,i})} \right)^{1/\alpha_g} \right] \left(\frac{1 + \alpha_g \underline{\Lambda}_g(t^*_{g,u,n_{u,g}})}{1 + \alpha_g \underline{\Lambda}_g(\tau^*_{g,u})} \right)^{1/\alpha_g} \\
&= \prod_{g=1}^{G} \prod_{u=1}^{U_g} \left[\prod_{i=1}^{n_{g,u}} \left(\frac{\left(\frac{\beta_g}{\eta_g}\right) \left(\frac{t^*_{g,u,i}}{\eta_g}\right)^{\beta_g - 1}}{1 + \alpha_g \left(\frac{t^*_{g,u,i}}{\eta_g}\right)^{\beta_g}} \right) \left(\frac{1 + \alpha_g \left(\frac{t^*_{g,u,i-1}}{\eta_g}\right)^{\beta_g}}{1 + \alpha_g \left(\frac{t^*_{g,u,i}}{\eta_g}\right)^{\beta_g}} \right)^{1/\alpha_g} \right] \left(\frac{1 + \alpha_g \left(\frac{t^*_{g,u,n_{g,u}}}{\eta_g}\right)^{\beta_g}}{1 + \alpha_g \left(\frac{\tau^*_{g,u}}{\eta_g}\right)^{\beta_g}} \right)^{1/\alpha_g} \\
&= \prod_{g=1}^{G} L_g(\boldsymbol{\theta}_g; \boldsymbol{t}_g, g, \mathcal{U}_g).
\end{aligned}
$$

Note that, once one are able to factor the likelihood function into a product of G independent factors, we can optimize the parameter vector $\boldsymbol{\theta}_g = (\eta_g, \beta_g, \alpha_g, \rho_g)$, separately. In

this sense, the log-likelihood function for a particular group g is expressed by

$$\ell_g(\boldsymbol{\theta}_g;\boldsymbol{t}_g,g,\mathcal{U}_g) = \sum_{u=1}^{U_g}\sum_{i=1}^{n_{g,u}}\left\{\log(\beta_g)-\log(\eta_g)+(\beta_g-1)\log\left(\frac{t^*_{g,u,i}}{\eta_g}\right)+\right.$$
$$+\frac{1}{\alpha_g}\log\left(1+\alpha_g\left(\frac{t^*_{g,u,i-1}}{\eta_g}\right)^{\beta_g}\right)-\left(1+\frac{1}{\alpha_g}\right)\log\left(1+\alpha_g\left(\frac{t^*_{g,u,i}}{\eta_g}\right)^{\beta_g}\right)+$$
$$\left.+\frac{1}{\alpha_g}\left[\log\left(1+\alpha_g\left(\frac{t^*_{g,u,n_{g,u}}}{\eta_g}\right)^{\beta_g}\right)-\log\left(1+\alpha_g\left(\frac{\tau^*_{g,u}}{\eta_g}\right)^{\beta_g}\right)\right]\right\}.$$

In both functions, $L_g(\boldsymbol{\theta}_g,\boldsymbol{t}_g,g,\mathcal{U}_g)$ and $\ell_g(\boldsymbol{\theta}_g,\boldsymbol{t}_g,g,\mathcal{U}_g)$, where $g=1,\ldots,G$, we have

$$t^*_{g,u,i} = t_{g,u,i}-(1-\rho_g)\sum_{j=0}^{\min(m_g-1;i-2)}\rho_g^j\, t_{g,u,i-1-j}\ ,$$

$$t^*_{g,u,i-1} = t_{g,u,i-1}-(1-\rho_g)\sum_{j=0}^{\min(m_g-1;i-2)}\rho_g^j\, t_{g,u,i-1-j}\ ,$$

$$t^*_{g,u,n_{g,u}} = t_{g,u,n_{g,u}}-(1-\rho_g)\sum_{j=0}^{\min(m_g-1;i-1)}\rho_g^j\, t_{g,u,n_{g,u}-j}\ ,$$

$$\tau^*_{g,u} = \tau_{g,u}-(1-\rho_g)\sum_{j=0}^{\min(m_g-1;i-1)}\rho_g^j\, t_{g,u,n_{g,u}-j}\ .$$

The standard errors of Maximum Likelihood Estimators (MLE) $\hat{\boldsymbol{\theta}}=(\hat{\boldsymbol{\theta}}_1,\ldots,\hat{\boldsymbol{\theta}}_g,\ldots,\hat{\boldsymbol{\theta}}_G)$, where $\hat{\boldsymbol{\theta}}_g=(\hat{\eta}_g,\hat{\beta}_g,\hat{\alpha}_g,\hat{\rho}_g)$ of the parameters vector $\boldsymbol{\theta}=(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_g,\ldots,\boldsymbol{\theta}_G)$, with $\boldsymbol{\theta}_g=(\eta_g,\beta_g,\alpha_g,\rho_g)$, can be obtained based on the asymptotic normality of MLEs. Under standard conditions of regularity, $\hat{\boldsymbol{\theta}}$ has multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and matrix of variances and covariance $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$. Both can be approximated by $\hat{\boldsymbol{\theta}}$ and $\mathcal{J}^{-1}(\hat{\boldsymbol{\theta}})$ (the inverse of the observed Fisher information matrix).

Using the point estimates $\hat{\boldsymbol{\theta}}$ and the respective standard errors $\mathrm{SE}(\hat{\boldsymbol{\theta}})$, we determine interval estimates for each component of $\boldsymbol{\theta}$. Considering a confidence of $(1-\gamma)100\%$, a confidence interval for θ_k with the expression

$$\left(\hat{\theta}_k - z_{\gamma/2}\sqrt{\mathcal{J}^{-1}_{kk}(\hat{\boldsymbol{\theta}})};\ \hat{\theta}_k + z_{\gamma/2}\sqrt{\mathcal{J}^{-1}_{kk}(\hat{\boldsymbol{\theta}})}\right),$$

where $\hat{\theta}_k$ is the MLE of θ_k, $\mathcal{J}^{-1}_{kk}(\hat{\boldsymbol{\theta}})$ is k-th diagonal element of the matrix $\mathcal{J}^{-1}(\hat{\boldsymbol{\theta}})$ and the $z_{\gamma/2}$ confidence coefficient denotes the standard normal quantile of the right tail which leaves a probability of $\gamma/2$.

In the context of an IR model, it is interesting to select the memory order, m_g, which best fits the data. The selection between models can be made based on the value of the log-likelihood function applied to $\hat{\boldsymbol{\theta}}_g$ (TOLEDO *et al.*, 2015).

6.4 Simulation Study

The evaluation of the previously exposed estimators was carried out based on 1000 Monte Carlo simulations. The generation of data in each iteration followed the process illustrated in the Algorithm 4.

Input:

m

G

$\mathcal{U} = (U_1, \ldots, U_G)$

$\boldsymbol{n} = \{(n_{1,1}, \ldots, n_{1,U_1}), \ldots, (n_{G,1}, \ldots, n_{G,U_G})\}$

$\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_G\} = \{(\eta_1, \beta_1, \alpha_1, \rho_1), \ldots, (\eta_G, \beta_G, \alpha_G, \rho_G)\}$

Output:

$$\{(\boldsymbol{t}; \mathcal{G}; \mathcal{U})\} = \{(\boldsymbol{t}_{1,1}; 1; 1), \ldots, (\boldsymbol{t}_{G,U_g}; G; U_G)\}$$

where

$$(\boldsymbol{t}_{g,u}; g; u) = \{(t_{g,u,1}; g; u), \ldots, (t_{g,u,n_{g;u}}; g; u)\}$$

Procedure:

/* Define $F_{X|t}(x; t, \boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\eta, \beta, \alpha, \rho)$ */

$$F_{X|t} := (x; t, \boldsymbol{\theta}) \to 1 - \exp\left\{ - \left(\Lambda(x+t; \boldsymbol{\theta}) - \Lambda(t; \boldsymbol{\theta}) \right) \right\}$$

for $g := 1$ **to** G **do**
  for $u := 1$ **to** U_g **do**
    for $i := 1$ **to** $n_{g,u}$ **do**
      $\omega \sim \text{Uniform}(0, 1)$ // Sample auxiliary variable ω
      $F_{X|t_{g,u,i-1}}(x; t_{g,u,i-1}, \boldsymbol{\theta}_g) - \omega = 0$ // Solve for x ($t_{g,u,0} = 0$)
      $t_{g,u,i} = t_{g,u,i-1} + x$ // Determine the i-th failure time
    end
  end
end

Algorithm 4: Generator of random numbers from the proposed model for a specific group.

We consider a combination of values for $\boldsymbol{\theta}_g = (\eta_g, \beta_g, \alpha_g, \rho_g) = (250, 1.5, 1.5, 0.65)$, and simulate values in a context that considers three orders of memory ($m_g = 1, 10, 20$), five values for the number of failure by system within each group ($n_{g,u} = 20, 60, 100, 140, 180$) and four values for the number of systems within a group ($U_g = 5, 10, 15, 20$).

The simulations were programmed using the R software (R Core Team, 2020), considering $B = 1000$ Monte Carlo replicas and the criteria used to evaluate the performance of the estimators were the mean relative estimation (MRE), square root of the mean squared error

(RMSE) and 90% coverage probability intervals (CP), calculated using the following expressions

$$\text{MRE}(\hat{\theta}_i) = \frac{1}{B}\sum_{j=1}^{B}\frac{\hat{\theta}_i^{(j)}}{\theta_i}, \qquad \text{RMSE}(\hat{\theta}_i) = \sqrt{\frac{1}{B}\sum_{j=1}^{B}\left(\hat{\theta}_i^{(j)} - \theta_i\right)^2},$$

and

$$\text{CP}(\hat{\theta}_i) = \frac{1}{B}\sum_{j=1}^{B}\mathbb{1}\left(\theta_i \in (a_i^{(j)}, b_i^{(j)})\right)$$

with

$$a_i^{(j)} = \hat{\theta}_i^{(j)} - 1.96 \times \text{SE}(\hat{\theta}_i^{(j)}) \qquad \text{and} \qquad b_i^{(j)} = \hat{\theta}_i^{(j)} + 1.96 \times \text{SE}(\hat{\theta}_i^{(j)}),$$

where θ_i is the i-th component of the vector $\boldsymbol{\theta} = (\eta, \beta, \alpha, \rho)$ and $\hat{\theta}_i$ is the corresponding maximum likelihood estimator. The function $\mathbb{1}(\cdot)$ is the indicator function, which returns the value one, if its argument is true and returns zero, otherwise. In addition $\text{SE}(\hat{\theta}_i^{(j)})$ denotes the standard error of the estimator $\hat{\theta}_i^{(j)}$.

We expect that with the increase in the number of failures ($n_{g,u}$) or the number of units in each group (U_g) the MRE criterion will be close to one, in the sense that the value The estimated value tends to approach the true value. Similarly, we expect the RMSE criterion to approach zero, indicating that the variability inherent in the estimation becomes less. On the other hand, the CP criterion must be close to the nominal value of 0.9.

The empirical results obtained can be seen in Figure 44. The measures according to each criterion present the expected behavior, the highlight is the RMSE of the estimator of η_g, however, this is due to the magnitude of the value set for this parameter (250) compared to the other values (1.5, 1.5, 0.65) for β, α and ρ, respectively.

In this sense, the results expose the asymptotic properties already known to MLEs, so that the model can be used to obtain reliable estimates of the process it is proposed to model.

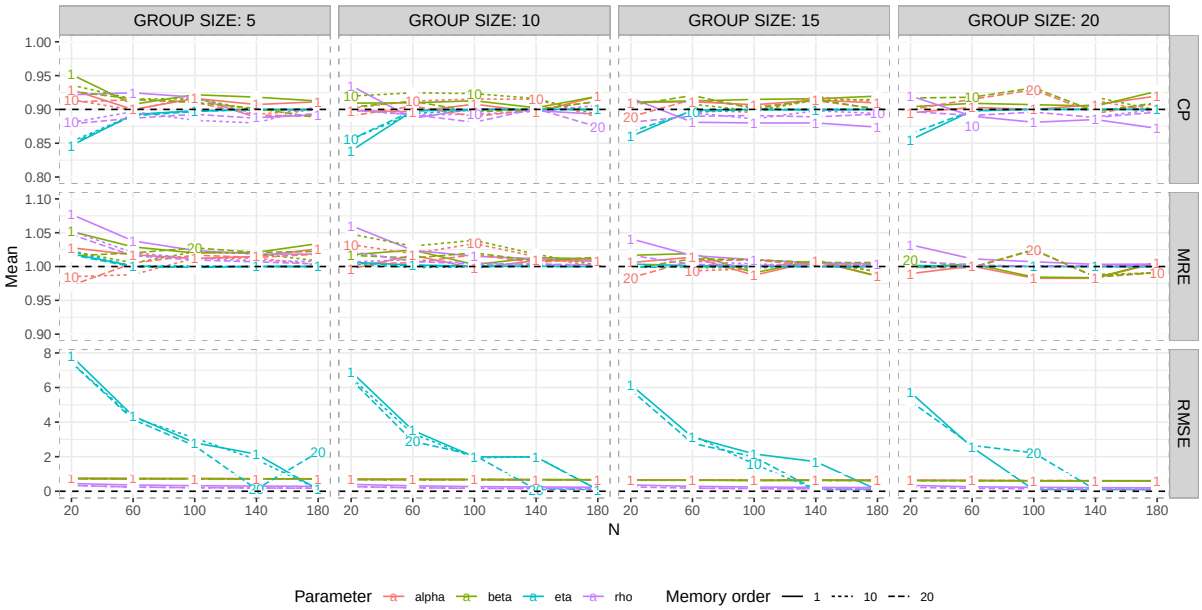


Figure 44 – Simulation results for the IR.ARA.F model, considering 1000 Monte Carlo replicas and the scenarios $\boldsymbol{\theta}_g = (\eta_g, \beta_g, \alpha_g, \rho_g) = (250, 1.5, 1.5, 0.65)$, $m_g = \{1, 10, 20\}$, $n_{g,u} = \{20, 60, 100, 140, 180\}$ and $U_g = \{5, 10, 15, 20\}$.

6.5 Application

To illustrate the use of the proposed model, we considered a dataset with the failures of a tractor fleet with agricultural machines categorized in five different groups. As an illustration, we understand that the performance of all machines, together, corresponds to the agricultural system studied, the farm fleet, so that the need for intervention in this system occurs with the failure of any unit within any group, individually, in a serial structure, or competing risks. On the other hand, the understanding of the time when all the machinery will have required some intervention is obtained by analyzing the results under a parallel structure.

6.5.1 Dataset

The record of maintenance interventions were obtained during the period between 2017 and 2020 on tractors on a farm in the state of São Paulo (Brazil), and the database maintenance orders are for 38 small agricultural tractors. Orders are divided into classes (checklist, preventive, tire repair, tractor and corrective suitability) and subclasses (accident, suitability, natural wear, operational failure, scheduled or preventive maintenance, quality of maintenance and quality of material), which describes the cause of equipment breakage.

The results of the preliminary evaluation of work orders (Table 25) shown the contribution of the total duration of orders measured in hours and the number of work order requests.

Table 25 – Total and average volume of hours and number of database maintenance work orders.

Class	Mean duration (hours)	Total hours (%)	Quantity (%)
Corrective	37.03	58246.78 (89.52)	1573 (60.78)
Adequacy trator	36.47	510.53 (9.05)	14 (36.75)
Tires	8.74	410.60 (0.78)	47 (1.82)
Preventive	6.20	5891.58 (0.63)	951 (0.54)
Check List	5.24	15.72 (0.02)	3 (0.11)
Total	25.14	65075.22 (100.00)	2588 (100.00)

We can notice that there is a prevalence of the total duration of maintenance hours in the corrective class with a participation of 89.52% and, for the preventive class this value is approximately 9%. In the maintenance order quantity indicators, it can be noted that there is a reasonably high issue of preventive service requests, of almost 37%, but still does not exceed the amount of corrective order issues, with a participation of approximately 61% . The prevalence of corrective maintenance over preventive maintenance is related to the low level of operational planning of agricultural fleets, making the tractors return to operation as quickly as possible.

In this study we will focus on some particular aspects of the problem, in order to illustrate the methodology previously proposed. In this sense:

- we will only consider the corrective maintenance records of the data set;
- we will only consider the times when the tractors were in operation until they failed. In other words, we consider that the repair measures took place instantly;
- the agricultural fleet represents the system under study, as a whole;
- the groups of machines denote groupings according to some characteristic of practical interest. Conceptually, the intervention in any group implies the intervention of the system, which will be counted as a failure associated with that group;
- tractors denote the component elements of each group, they compete with each other to demand an intervention on the fleet. Conceptually, the occurrence of a tractor failure implies intervention in a group that, in turn, implies system intervention. The failure will be registered as a result of a specific group and tractor;
- tractors within the same group exhibit the same basic behavior, but share intrinsic characteristics that make them unique within their group.

6.5.2 Reliability model

Given the previous considerations, we made a descriptive outline in terms of reliability (Figure 45). In the top two lines of we observe the behavior of the failures over time, considering

each component unit, within each group. In the first line, we have an indication of how the failures are distributed over the operation time. The failure are more concentrated in groups 3 and 5, this can also be seen in the graphs of the second line, in which the growth in the number of failures presents a more intense behavior in these groups. On the other hand, groups 2 and 4 show the slowest growth. In the third line, Duane plots are exposed for each unit in each group, in which one can see an approximately linear behavior, at least for most cases, indicating that the power-law model is enough to represent the basic behavior for failure intensity.

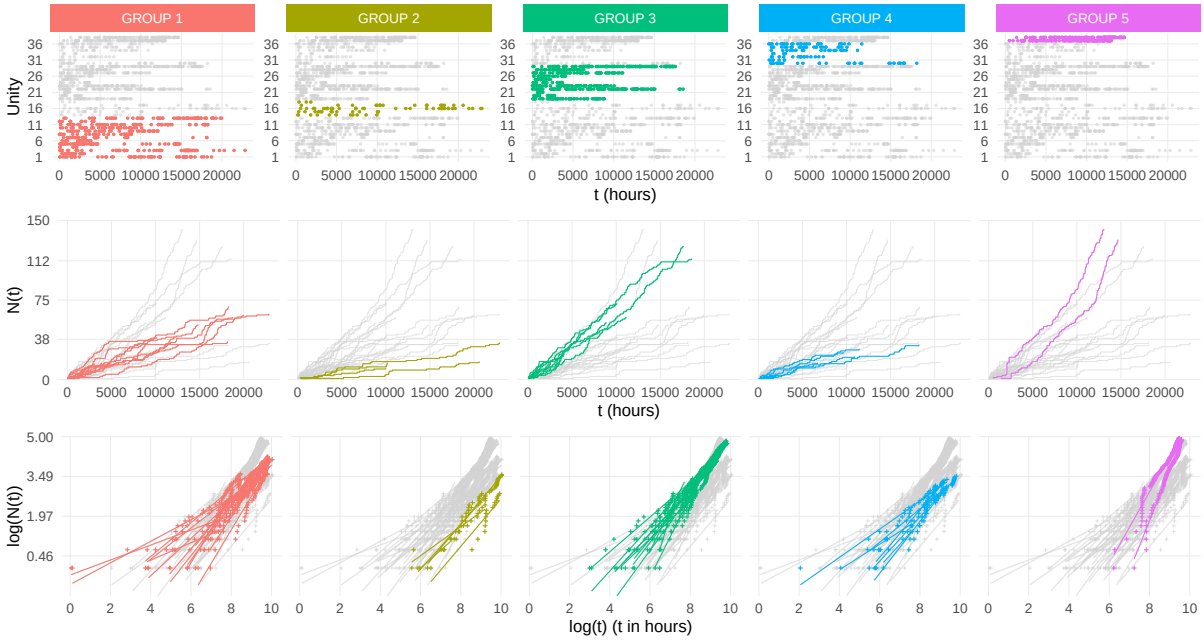


Figure 45 – Descriptive graphs for failure times (on the top line of plots) and cumulative failures by group and unit (middle line). Duane plots by group and unit (bottom line).

Considering the models shown in Table 24, we made a selection among them using the highest log-likelihood value applied in parameter estimates as criteria. The normalized results for each group can be seen in Figure 46 and, the models chosen were IR.ARA8.F (GROUP 1), IR.ARA8.F (GROUP 2), IR.ARA4 (GROUP 3), IR.ARA1.F (GROUP 4), IR.ARA15 (GROUP 5).

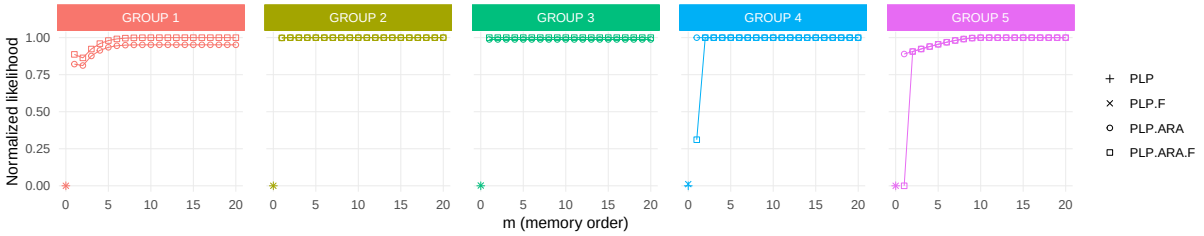


Figure 46 – Normalized likelihood as a criterion for model selection.

According to the estimates for the models (Table 26), we have evidences that there may be unobserved effects that influence the failure times in the systems of groups 1 and

3, given that $\alpha_1 = 0.222\,(0.108)$ and $\alpha_3 = 0.0295\,(0.0052)$. On the other hand, with respect to the repair effect, no one express failure behaviors very close to a minimum repair model, since $\rho << 1.0$ for everyone, the group 1 presented an evident intermediate effect of the repair actions, $\rho_1 = 0.3215\,(0.1169)$. On the other hand, the groups 2 and 4 presented repair effect close to perfect repair sinse $\rho_2 \approx \rho_4 \approx 0$.

Table 26 – Estimatives to the frailty imperfect repair models.

				90% Interval	
		Estimative	Standard Error	Lower	Upper
IR.ARA8.F	α_1	0.2220	0.1080	0.1236	0.3707
	β_1	0.8161	0.1442	0.6144	1.0854
	η_1	121.1883	28.3588	59.4364	176.3932
	ρ_1	0.3215	0.1169	0.16052	0.4816
IR.ARA1	β_2	0.8036	0.0614	0.7039	0.9161
	η_2	710.8643	0.0025	710.8609	710.8671
	ρ_2	3.77e-9	6.6912e-10	1.3183e-9	3.9548e-9
IR.ARA1.F	α_1	0.0295	0.0052	0.0216	0.0444
	β_3	0.8507	0.0673	0.7512	0.9712
	η_3	135.8359	0.0107	135.8193	135.8507
	ρ_3	1.1371e-9	2.2023e-10	4.0094	1.2028
IR.ARA1	β_4	0.8054	0.0629	0.7066	0.9135
	η_4	394.5440	0.01104	394.5387	394.5517
	ρ_4	1.5787e-11	1.8339e-12	1.6109e-11	1.6109e-11
IR.ARA10	β_5	0.5067	0.0415	0.4426	0.5803
	η_5	61.5044	0.4130	61.4779	61.5228
	ρ_5	4.3495e-6	1.3775e-6	1.7039e-6	5.1116e-6

In order to have a perception of the quality of the adjustment, we can observe the estimated mean value function over the accumulation of failures observed in the systems (upper line of Figure 47) and also the number of failures estimated by the fitted model and the MCF obtained by nonparametric method (bottom line of Figure 47). From these results, we have evidence that the model is quite adherent to the behavior observed in the actual data.

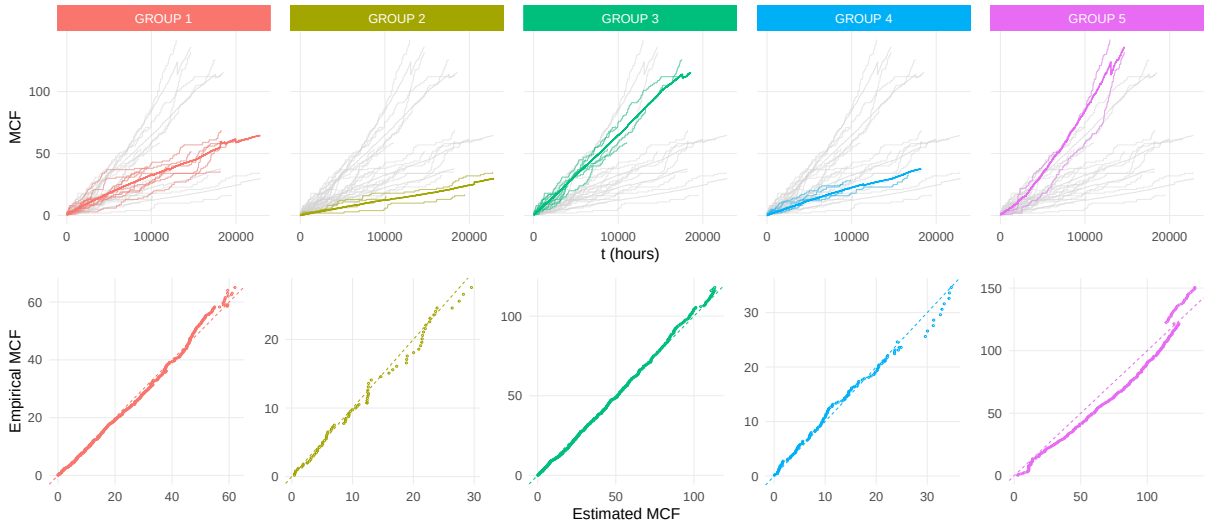


Figure 47 – Results for estimated mean value function by group over the observed accumulated failures per unit (on the top line plots). Observed mean number of failures versus estimated (bottom line).

Although the estimation process is done on the mean value function considering all units within a group, it is possible to visualize the behavior of the estimated function on each unit specific history. These results can be seen in the Appendix A.

After adjusting, selecting and validating the models, we use the estimates and the failure history of the groups to estimate the reliability function that describes the common behavior of the units in each group. These results are shown in Figure 48 and, from them, we calculated some measures of practical interest, such as the mean and median next failure time. The most reliable group is group 2, with an estimated median time close to 568 (403,796) hours, followed by group 4 with 382 (214,658) hours. On the other hand, the group with the lowest estimate for the median time is the 5 with only 41 (33,52) hours of operation.

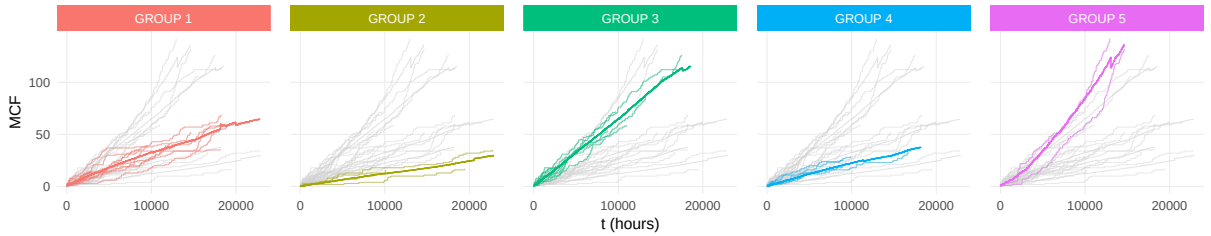


Figure 48 – Next failure reliability, by group. The thick and solid line denotes the estimated reliability. Transparent margins are 90% confidence intervals. The vertical lines indicate the median next failure time with the respective 90% confidence limits.

Although we have estimated a general behavior of the reliability function for the units in each group, we keep in mind that each unit has its own failure history and that the group's reliability carries the effect of each one. Thus, considering the time scale (the data-time recorded in the data set) of failures occurrence, the results for reliability at each level of the hierarchy

proposed in Figure 43 calculated based on Equation (6.4) (in the case of Figure 49) and in the Equation (6.5) (in the case of Figure 50).

From the results shown in Figure 49, we have that the reliability of the general system (after its last failure) is very close to zero, since the reliability of many units is quite low. The group with the lowest reliability is group 3, and within this group the most critical units are units 20, 21, 24 and 25. Group 1 is the second least reliable, with emphasis on the criticality of units 2, 4, 6 and 8. The third group is 4 with critical units 31, 32 e 33, followed by 2 with critical units 14, 15 e 18 and the group 5.

On the other hand, the results in Figure 50 show us the reliability associated with the next failure of all units. In this sense, we understand that the group that will probably take the longest time for all units to fail is group 1, followed by groups 2, 3, 4 and 5.

Although we have made a more general interpretation of the reliability curves in the forecast horizon as a whole. In practical contexts, an interpretation in terms of data-time allows planning and maintenance schedules considering the work schedule of the company for the fleet. These results provide reliability guides for both the general system and the units that compose it, and can be updated with each new failure occurrence, generating new estimates based on the updated failure history.

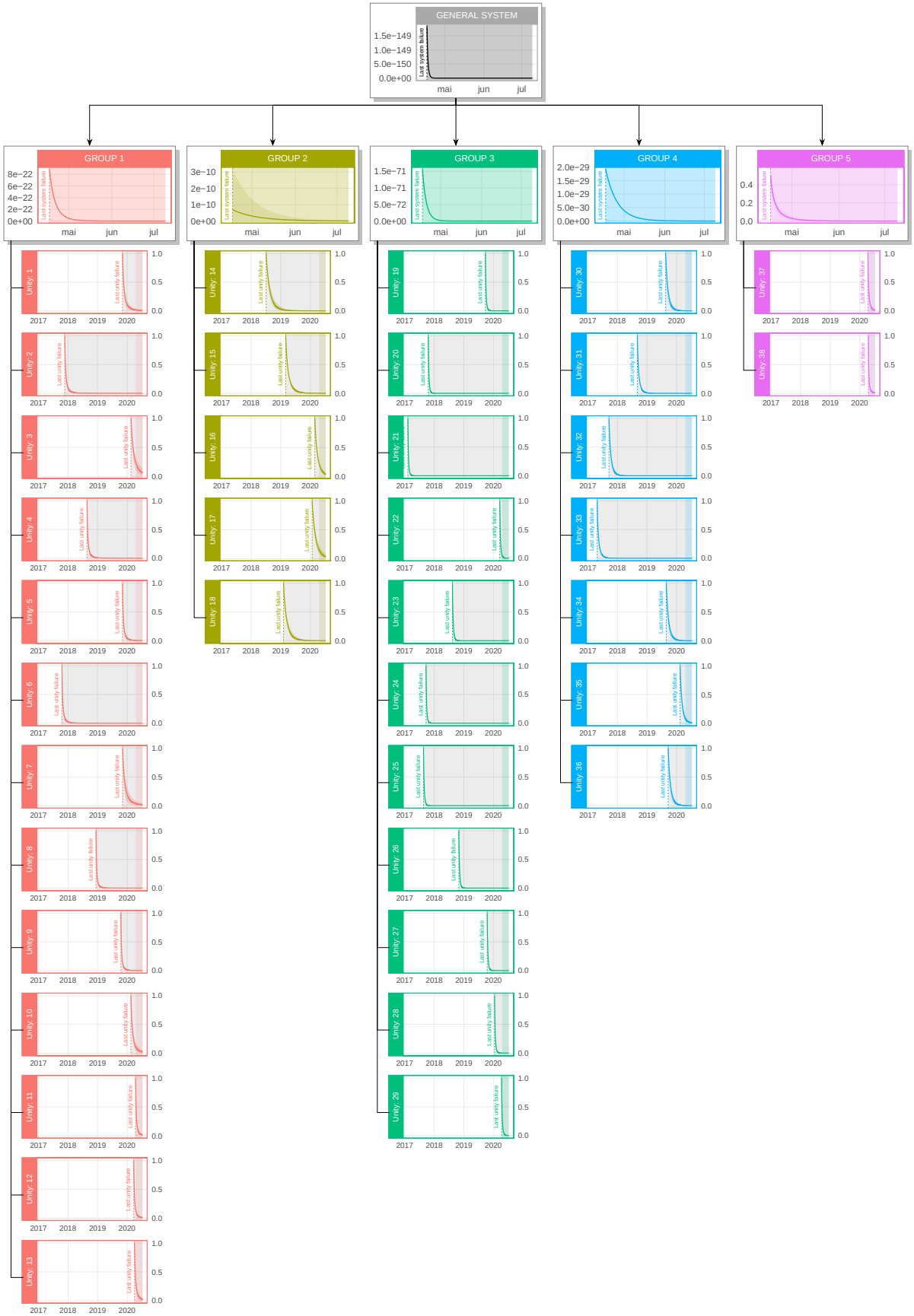


Figure 49 – Next failure reliability (thick and solid line), by hierarchical level in a serial structure. Transparent margins around thick and solid line are 90% confidence intervals. The transparent gray band (for the units) is the horizon after the last unit failure. The colored transparent band is the horizon after the last general system failure.

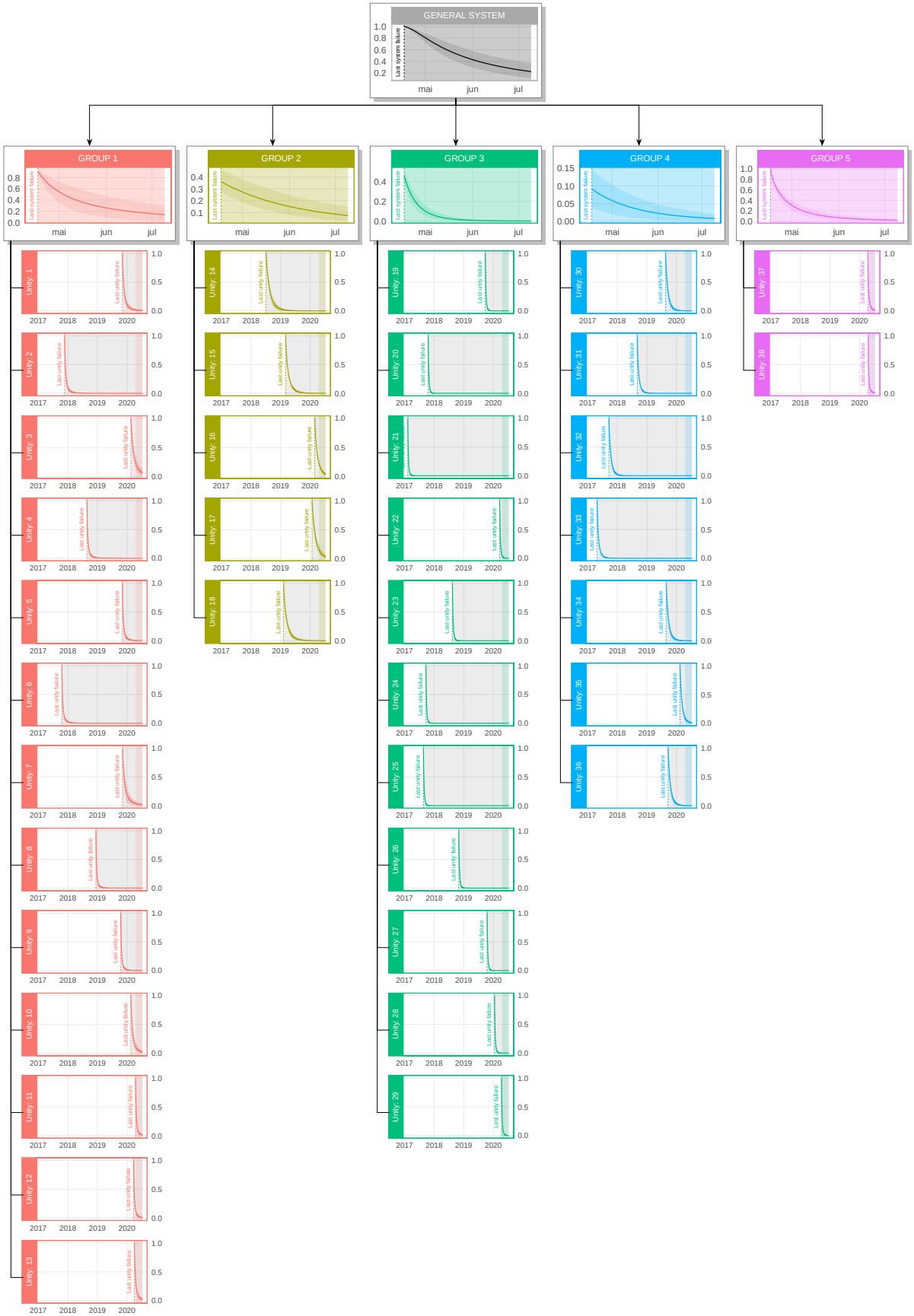


Figure 50 – Next failure reliability (thick and solid line), by hierarchical level in a parallel structure. Transparent margins around thick and solid line are 90% confidence intervals. The transparent gray band (for the units) is the horizon after the last unit failure. The colored transparent band is the horizon after the last general system failure.

6.6 Concluding Remarks

In this chapter we continued the studies started in the Sections (3, 4 and 5), taking new steps to some studies carried out by Almeida (2019), DAndrea (2020) and Lindqvist (2006). We present a frailty model that considers a hierarchical level for groups of repairable systems, maintaining the assumptions of competing risks, in an ARAm IR regime with a PLP intensity function. We deduce the likelihood function and suggest the maximum likelihood method to obtain the estimators of the model parameters. The properties of these estimators were evaluated in a simulation study, which returned the expected results.

The proposed model and the analysis carried out as an example express a way of assessing the next failure time reliability, of a general system hierarchically represented composed by groups of repairable units, in an analytical and graphical way at any level of the proposed hierarchy. So that the exposed methodology can be easily applied to the daily life of the industry as an auxiliary tool for decision making.

The following steps of this modeling can consider any number of hierarchical levels, new formats for the intensity function associated with the first failure, the extension of the analysis to the failure modes associated with the units and the modeling of their possible dependence via shared frailty models.

CHAPTER

7

CONCLUDING REMARKS AND FURTHER RESEARCH

In this thesis we introduced some new statistical models for repairable systems. In all of them, we assumed a framework for hierarchical competing risks under the assumption that the failure modes act independently. We also assumed that the repairs are minimal (in Chapters 3, 4 and 5) and imperfect (Chapter 6), and the failure intensity follows a PLP model after a convenient reparametrization. Under a classical (in Chapters 3, 5 and 6) and bayesian (Chapter 4) framework, we proposed some bias-corrected estimators and confidence intervals for the model parameters (in Chapters 3 and 4). In all the performances were investigated using a simulation study. When was the case, the simulation results revealed that the bias-corrected provide better estimates, mainly for the PLP shape parameters, than the usual estimators. Besides, for the models from Chapters 3 and 4, the exact confidence intervals for parameters were obtained. In Chapter 6 we present a frailty model that considers a hierarchical level for groups of repairable systems, maintaining the assumptions of competing risks, in an ARAm IR regime.

To illustrate the use of the methodologies, we used a series of real problems, especially from the project that motivated this research. For systems in which there was no practice data, we used a process to simulate failure times using the information available about the system and we analyze the first-failure time reliability of a real early-stage project related to an in-pipe robot traction system, and also a real example on a BOP system (in Chapter 3), two systems (at any hierarchical level), namely, the pressure vessels set and the traction system (in Chapter 4). To illustrate the proposed methodology in Chapter 5 we consider the fatigue system of the robotic unit already studied in the previous chapters, in this context, the data were obtained with the physical simulation of failure times in terms of system fatigue. In the example where we had pratical data available, the tractor fleet problem (Chapter 6) we assessing the next failure time reliability. we understand that the tractor fleet corresponds to

the farm's agricultural system.

As future works, we intend to evaluate the quality of these estimators in a context with outliers. In addition, we wish to evaluate the change in reliability based on the increase in redundancy of some subsystems. We also intend to model the dependence among the failure modes via shared frailty models, new formats for the first-time intensity function, the extension of the analysis to also consider repair times in the analysis.

BIBLIOGRAPHY

ALMEIDA, M. P. **Statistical inference for non-homogeneous Poisson process with competing risks: a repairable systems approach under power-law process**. 111 p. Phd Thesis (PhD Thesis) — Universidade de São Paulo, 2019. Citations on pages 94, 133, 134, and 151.

ANDERSEN, P. K.; BORGAN, O.; GILL, R. D.; KEIDING, N. **Statistical Models Based on Counting Processes**. [S.l.]: Springer New York, 1996. ISBN 0387945199. Citations on pages 27, 58, and 134.

ASCHER, H.; FEINGOLD, H. **Repairable systems reliability: modeling, inference, misconceptions and their causes**. [S.l.]: Marcel Dekker, 1984. 223 p. Citations on pages 26, 33, 34, 39, 41, 46, 95, and 133.

ATWOOD, C. L. Parametric estimation of time-dependent failure rates for probabilistic risk assessment. **Reliability Engineering & System Safety**, Elsevier Science, v. 37, 1992. Citation on page 46.

AVEN, T. Optimal replacement under a minimal repair strategy, a general failure model. **Advances in Applied Probability**, Cambridge Univ Press, v. 15, n. 01, p. 198–211, 1983. Citation on page 62.

AVEN, T.; JENSEN, U. A general minimal repair model. **journal of applied probability**, Cambridge Univ Press, v. 37, n. 01, p. 187–197, 2000. Citation on page 62.

BAHOOTOROODY, A.; ABAEI, M. M.; ARZAGHI, E.; SONG, G.; De Carlo, F.; PALTRINIERI, N.; ABBASSI, R. On reliability challenges of repairable systems using hierarchical bayesian inference and maximum likelihood estimation. **Process Safety and Environmental Protection**, v. 135, p. 157–165, 2020. ISSN 0957-5820. Citation on page 133.

BAIN, L. **Statistical analysis of reliability and life-testing models: theory and methods**. [S.l.]: Routledge, 2017. Citation on page 63.

BARLOW, R.; HUNTER, L. Optimum preventive maintenance policies. **Operations research**, INFORMS, v. 8, n. 1, p. 90–100, 1960. Citation on page 62.

BAUCCIO, M. *et al.* **ASM metals reference book**. [S.l.]: ASM international, 1993. Citation on page 125.

BERGER, J. **On the development of reference priors (with discussion). Bayesian Statistics IV, JM Bernardo, et. al**. [S.l.]: Oxford University Press, Oxford, 1992. Citations on pages 95 and 100.

BERGER, J. O.; BERNARDO, J. M. Estimating a product of means: Bayesian analysis with reference priors. **Journal of the American Statistical Association**, Taylor & Francis, v. 84, n. 405, p. 200–207, 1989. Citation on page 95.

_____. Ordered group reference priors with application to the multinomial problem. **Biometrika**, Oxford University Press, v. 79, n. 1, p. 25–37, 1992. Citation on page 95.

_____. Reference priors in a variance components problem. In: **Bayesian analysis in statistics and econometrics**. [S.l.]: Springer, 1992. p. 177–194. Citations on pages 95 and 100.

BERGER, J. O.; BERNARDO, J. M.; SUN, D. *et al.* Overall objective priors. **Bayesian Analysis**, International Society for Bayesian Analysis, v. 10, n. 1, p. 189–221, 2015. Citation on page 101.

BERNARDO, J. M. Reference posterior distributions for bayesian inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 113–147, 1979. Citations on pages 95 and 100.

_____. Reference analysis. **Handbook of statistics**, Elsevier, v. 25, p. 17–90, 2005. Citations on pages 95, 99, and 101.

BOARDMAN, B. Fatigue resistance of steels. **ASM International, Metals Handbook. Tenth Edition**, v. 1, p. 673–688, 1990. Citation on page 125.

BROWN, M.; PROSCHAN, F. Imperfect repair. **Journal of Applied probability**, JSTOR, p. 851–859, 1983. Citation on page 35.

CANTARELLA, R. The reliability of periodically repaired n 8722; i/n parallel redundant systems. **IEEE Transactions on Computers**, IEEE, C-32, p. 597–598, 1983. ISSN 0018-9340. Available: <http://doi.org/10.1109/tc.1983.1676284>. Citation on page 116.

CHAN, J.; SHAW, L. Modeling repairable systems with failure rates that depend on age and maintenance. **IEEE Transactions on Reliability**, v. 42, n. 4, p. 566–571, Dec 1993. ISSN 1558-1721. Citations on pages 49 and 50.

CHEN, I.-R. Effect of parallel planning on system reliability of real-time expert systems. **IEEE Transactions on Reliability**, IEEE, v. 46, p. 81–87, 1997. ISSN 0018-9529,1558-1721. Available: <http://doi.org/10.1109/24.589932>. Citation on page 116.

CORBERA, S.; OLAZAGOITIA, J. L.; LOZANO, J. A. Multi-objective global optimization of a butterfly valve using genetic algorithms. **ISA transactions**, Elsevier, v. 63, p. 401–412, 2016. Citation on page 76.

CORDEIRO, G. M.; KLEIN, R. Bias correction in arma models. **Statistics & Probability Letters**, Elsevier, v. 19, n. 3, p. 169–176, 1994. Citation on page 72.

COX, D. R. Regression models and life-tables. **Journal of the Royal Statistical Society. Series B (Methodological)**, [Royal Statistical Society, Wiley], v. 34, n. 2, p. 187–220, 1972. ISSN 00359246. Citation on page 58.

COX, D. R.; LEWIS, P. A. W. **The statistical analysis of series of events**. [S.l.]: Methuen & Co Ltd, 1966. Citation on page 47.

COX, D. R.; REID, N. Parameter orthogonality and approximate conditional inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 1–39, 1987. Citations on pages 63, 99, and 100.

COX, D. R.; SNELL, E. J. A general definition of residuals. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 248–275, 1968. Citation on page 72.

CROW, L. H. Reliability analysis for complex, repairable systems. **Reliability and Biometry**, v. 13, n. 6, p. 379–410, 1974. Citation on page 26.

_____. Confidence interval procedures for the weibull process with applications to reliability growth. **Technometrics**, v. 24, n. 1, p. 67–72, 1982. Citation on page 26.

CROWDER, M. J. **Classical Competing Risks**. [S.l.]: Chapman and Hall/CRC, 2001. Citations on pages 25, 26, 64, and 95.

DANDREA, A. M. E. **Classe de modelos de fragilidade com efeito do acúmulo de reparos em múltiplos sistemas reparáveis**. 132 p. Phd Thesis (PhD Thesis) — Universidade de São Paulo, 2020. Citations on pages 134 and 151.

DOYEN, L.; GAUDOIN, O. Classes of imperfect repair models based on reduction of failure intensity or virtual age. **Reliability Engineering & System Safety**, Elsevier BV, v. 84, n. 1, p. 45–56, apr 2004. Citations on pages 26, 35, 50, 51, 54, 62, and 134.

DRÆGEBØ, E. **Reliability Analysis of Blowout Preventer Systems - A comparative study of electro-hydraulic vs. all-electric BOP technology**. Master's Thesis (Master's Thesis) — Norwegian University of Science and Technology, Norway, 2014. Citation on page 87.

DUANE, J. T. Learning curve approach to reliability monitoring. **IEEE Transactions on Aerospace**, v. 2, n. 2, p. 563–566, 1964. Citation on page 49.

DURAN, J. A. R.; COSTA, D. J. R. d.; JUNIOR, L. C. d. A. R. Numerical stress-life curves for the aisi 4340 steel using two sets of materials properties and different bi-axial stress ratios. **Latin American Journal of Solids and Structures**, SciELO Brasil, v. 15, n. 11, 2018. Citation on page 125.

ELBERS, C.; RIDDER, G. True and spurious duration dependence: The identifiability of the proportional hazard model. **The Review of Economic Studies**, Oxford University Press (OUP), v. 49, n. 3, p. 403, jul 1982. Citations on pages 59 and 136.

ENGELHARDT, M.; BAIN, L. J. On the mean time between failures for repairable systems. **IEEE Transactions on Reliability**, v. 35, n. 4, p. 419–422, 1986. Citation on page 26.

FERREIRA, P. H.; RAMOS, E.; RAMOS, P. L.; GONZALES, J. F.; TOMAZELLA, V. L.; EHLERS, R. S.; SILVA, E. B.; LOUZADA, F. Objective bayesian analysis for the lomax distribution. **Statistics & Probability Letters**, Elsevier, v. 159, p. 108677, 2020. Citation on page 95.

FINKELSTEIN, M. Minimal repair in heterogeneous populations. **Journal of Applied Probability**, Cambridge Univ Press, v. 41, n. 01, p. 281–286, 2004. Citation on page 62.

FIRTH, D. Bias reduction of maximum likelihood estimates. **Biometrika**, Oxford University Press, v. 80, n. 1, p. 27–38, 1993. Citations on pages 72 and 73.

FU, J.; XU, A.; TANG, Y. Objective bayesian analysis of pareto distribution under progressive type-ii censoring. **Statistics & Probability Letters**, Elsevier, v. 82, n. 10, p. 1829–1836, 2012. Citation on page 95.

GAUDOIN, O.; YANG, B.; XIE, M. Confidence intervals for the scale parameter of the power-law process. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 35, n. 8, p. 1525–1538, 2006. Citation on page 63.

GUTHRIE, W. F. **NIST/SEMATECH e-Handbook of Statistical Methods (NIST Handbook 151)**. [S.l.]: National Institute of Standards and Technology, 2020. Citation on page 47.

HASUNUMA, S.; OKI, S.; MOTOMATSU, K.; OGAWA, T. Fatigue life prediction of carbon steel with machined surface layer under low-cycle fatigue. **International Journal of Fatigue**, Elsevier, v. 123, p. 255–267, 2019. Citation on page 80.

HOKSTAD, P. The failure intensity process and the formulation of reliability and maintenance models. **Reliability Engineering & System Safety**, Elsevier BV, v. 58, n. 1, p. 69–82, oct 1997. Citation on page 49.

HOUGAARD, P. Frailty models for survival data. **Lifetime Data Analysis**, Springer Science and Business Media LLC, v. 1, n. 3, p. 255–273, 1995. Citations on pages 27, 58, and 134.

HØYLAND, A.; RAUSAND, M. **System reliability theory: models and statistical methods**. [S.l.]: John Wiley & Sons, 2009. Citation on page 64.

INCE, A.; GLINKA, G. A modification of morrow and smith–watson–topper mean stress correction models. **Fatigue & Fracture of Engineering Materials & Structures**, Wiley Online Library, v. 34, n. 11, p. 854–867, 2011. Citation on page 125.

JEFFREYS, H. An invariant form for the prior probability in estimation problems. In: THE ROYAL SOCIETY. **Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences**. [S.l.], 1946. v. 186, n. 1007, p. 453–461. Citation on page 72.

JIANG, T.; LIU, Y. Parameter inference for non-repairable multi-state system reliability models by multi-level observation sequences. **Reliability Engineering & System Safety**, Elsevier, v. 166, p. 3–15, 2017. Citations on pages 25 and 94.

KANG, S. G.; LEE, K.; LEE, W. D. Noninformative priors for the generalized half-normal distribution. **Journal of the Korean Statistical Society**, Elsevier, v. 43, n. 1, p. 19–29, 2014. Citation on page 95.

KIJIMA, M. Some results for repairable systems with general repair. **Journal of Applied probability**, JSTOR, p. 89–102, 1989. Citations on pages 34 and 35.

KIJIMA, M.; MORIMURA, H.; SUZUKI, Y. Periodical replacement problem without assuming minimal repair. **European Journal of Operational Research**, Elsevier, v. 37, n. 2, p. 194–203, 1988. Citations on pages 26, 35, and 134.

KIM, D. H.; LEE, W. D.; KANG, S. G. Non-informative priors for the inverse weibull distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 84, n. 5, p. 1039–1054, 2014. Citation on page 95.

KIMURA, T.; TANAKA, T.; FUJIMOTO, K.; OGAWA, K. Hydrodynamic characteristics of a butterfly valve—prediction of pressure loss characteristics. **ISA transactions**, Elsevier, v. 34, n. 4, p. 319–326, 1995. Citation on page 76.

KYPARISIS, J.; SINGPURWALLA, N. D. Bayesian inference for the weibull process with applications to assessing software reliability growth and predicting software failures. In: ELSEVIER SCIENCE PUBLISHERS B. V. (NORTH HOLLAND), AMSTERDAM. **Com- puter Science and Statistics: Proceedings of the Sixteenth Symposium on the Interface (ed. L. Billard)**. [S.l.], 1985. p. 57–64. Citation on page 26.

LANGSETH, H.; LINDQVIST, B. H. Competing risks for repairable systems: a data study. **Journal of Statistical Planning and Inference**, Elsevier, v. 136, n. 5, p. 1687–1700, 2006. Citations on pages 26, 94, and 133.

LIEPING, Y.; ZHE, Q. Failure mechanism and its control of building structures under earthquakes based on structural system concept. **Journal of Earthquake and Tsunami**, v. 3, n. 4, p. 249–259, 2009. Citations on pages 25, 94, and 133.

LINDQVIST, B. H. On the statistical modeling and analysis of repairable systems. **Statistical Science**, Institute of Mathematical Statistics, v. 21, n. 4, p. 532–551, nov 2006. Citations on pages 134, 137, and 151.

LIU, J.; SONG, B.; ZHANG, Y. Competing failure model for mechanical system with multiple functional failures. **Advances in Mechanical Engineering**, v. 10, n. 5, p. 1–16, 2018. Citation on page 26.

LOUZADA, F.; ALEGRÍA, L. F. A.; COLOMBO, D.; MARTINS, D. E.; SANTOS, H. F. L.; CUMINATO, J. A.; RODRIGUEZ, O. M. H.; TOMAZELLA, V. L. D.; FERREIRA, P. H.; RAMOS, P. A. M.; NIAKI, S. R. A.; GONZATTO, O. A. J.; PERISSINI, I. C. A repairable system subjected to hierarchical competing risks: Modeling and applications. **IEEE Access**, v. 7, p. 171707–171723, 2019. Citations on pages 93, 95, 96, 106, 112, 116, 117, 118, 119, and 124.

MALIK, M. A. K. Reliable preventive maintenance scheduling. **A I I E Transactions**, Informa UK Limited, v. 11, n. 3, p. 221–228, sep 1979. Citations on pages 54 and 55.

MAZZUCHI, T. A.; SOYER, R. A bayesian perspective on some replacement strategies. **Reliability Engineering & System Safety**, Elsevier, v. 51, n. 3, p. 295–303, 1996. Citation on page 62.

MEEKER, W. Q.; ESCOBAR, L. A. **Statistical Methods for Reliability Data**. [S.l.]: Wiley, 2014. (Wiley Series in Probability and Statistics). ISBN 9781118625972. Citations on pages 67 and 118.

MULLOR, R.; MULERO, J.; TROTTINI, M. A modelling approach to optimal imperfect maintenance of repairable equipment with multiple failure modes. **Computers & Industrial Engineering**, Elsevier, v. 128, p. 24–31, 2019. Citations on pages 94 and 133.

NAVAS, M. A.; SANCHO, C.; CARPIO, J. Reliability analysis in railway repairable systems. **International Journal of Quality & Reliability Management**, Emerald Publishing Limited, 2017. Citation on page 133.

NOTASH, L. Solutions of interval systems for under-constrained and redundant parallel manipulators. **Journal of Mechanisms and Robotics**, 2017. ISSN 1942-4302. Available: <http://doi.org/10.1115/1.4037803>. Citation on page 116.

OLIVEIRA, M. D. de; COLOSIMO, E. A.; GILARDONI, G. L. Bayesian inference for power law processes with applications in repairable systems. **Journal of Statistical Planning and Inference**, Elsevier, v. 142, n. 5, p. 1151–1160, 2012. Citations on pages 63, 98, 99, and 138.

PARK, S.; PARK, N.; KIM, J. A statistical study on tensile characteristics of stainless steel at elevated temperatures. In: IOP PUBLISHING. **Journal of Physics: Conference Series**. [S.l.], 2010. v. 240, n. 1, p. 012083. Citation on page 125.

PARK, S.-H.; PARK, N.-S.; KIM, J.-H. A study on tensile characteristics of aisi 304 stainless steel under room and elevated temperatures. **Journal of the Korean Society of Propulsion Engineers**, The Korean Society of Propulsion Engineers, v. 12, n. 5, p. 35–42, 2008. Citation on page 125.

PAWITAN, Y. **In All Likelihood**: Statistical modelling and inference using likelihood. [S.l.]: Oxford University Press, 2013. ISBN 0199671222. Citation on page 139.

PENG, W.; SHEN, L.; SHEN, Y.; SUN, Q. Reliability analysis of repairable systems with recurrent misuse-induced failures and normal-operation failures. **Reliability Engineering System Safety**, v. 171, p. 87–98, 2018. ISSN 0951-8320. Citation on page 133.

PHAM, H.; WANG, H. Imperfect maintenance. **European Journal of Operational Research**, Elsevier BV, v. 94, n. 3, p. 425–438, nov 1996. Citation on page 49.

PINTILIE, M. **Competing risks: a practical perspective**. [S.l.]: John Wiley & Sons, 2006. Citations on pages 26 and 95.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2018. Available: <https://www.R-project.org/>. Citation on page 103.

_____. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2020. Available: <https://www.R-project.org/>. Citations on pages 120 and 141.

RAMOS, P. L.; MOALA, F. A.; ACHCAR, J. A. Objective priors for estimation of extended exponential geometric distribution. **Journal of Modern Applied Statistical Methods**, v. 13, n. 2, p. 12, 2014. Citation on page 95.

RAUSAND, H. **System Reliability Theory**. 2. ed. [S.l.]: John Wiley & Sons, 2003. ISBN 047147133X. Citations on pages 26 and 134.

RAUSAND, M.; HOYLAND, A. **System Reliability Theory**: Models, statistical methods, and applications. [S.l.]: John Wiley & Sons, 2004. 664 p. ISBN 047147133X. Citations on pages 34 and 44.

REIS, R. C. P.; COLOSIMO, E. A.; GILARDONI, G. L. Hierarchical modelling of power law processes for the analysis of repairable systems with different truncation times: An empirical bayes approach. **Brazilian Journal of Probability and Statistics**, v. 33, n. 2, p. 374–396, 2019. Citation on page 26.

RIGDON, S. E.; BASU, A. P. The power law process: A model for the reliability of repairable systems. **Journal of Quality Technology**, v. 21, n. 4, p. 251–260, 1989. Citation on page 26.

_____. **Statistical methods for the reliability of repairable systems**. [S.l.]: Wiley New York, 2000. Citations on pages 34, 62, 63, 70, 77, 95, 103, 120, and 133.

ROSS, S. **Stochastic processes**. [S.l.]: Wiley, 1996. 510 p. (Wiley series in probability and statistics: Probability and statistics). ISBN 9780471120629. Citations on pages 36 and 39.

RUIZ, F. B. H. M.-P. J. M. On optimal allocation of redundant components for series and parallel systems of two dependent components. **Journal of Statistical Planning and Inference**, Elsevier Science, v. 141, p. 3094–3104, 2011. ISSN 0378-3758. Available: <http://doi.org/10.1016/j.jspi.2011.03.031>. Citation on page 116.

SHARMA, G.; RAI, R. N. Risk-based threshold on intensity function of repairable systems: A case study on aero engines. **International Journal of Reliability, Quality and Safety Engineering**, World Scientific Pub Co Pte Lt, p. 2150020, nov 2020. Citation on page 133.

SLIMACEK, V.; LINDQVIST, B. H. Nonhomogeneous poisson process with nonparametric frailty. **Reliability Engineering & System Safety**, Elsevier BV, v. 149, p. 14–23, may 2016. Citation on page 58.

SOMBOONSAVATDEE, A.; SEN, A. Statistical inference for power-law process with competing risks. **Technometrics**, Taylor & Francis, v. 57, n. 1, p. 112–122, 2015. Citations on pages 27, 63, 65, 68, 93, 95, 96, and 133.

SONG, X. G.; WANG, L.; BAEK, S. H.; PARK, Y. C. Multidisciplinary optimization of a butterfly valve. **ISA transactions**, Elsevier, v. 48, n. 3, p. 370–377, 2009. Citation on page 76.

SURESH, S. **Fatigue of materials**. [S.l.]: Cambridge university press, 1998. Citation on page 80.

SYAMSUNDAR, A.; NAIKAN, V.; WU, S. Alternative scales in reliability models for a repairable system. **Reliability Engineering & System Safety**, Elsevier, v. 193, p. 106599, 2020. Citations on pages 94 and 133.

THOMPSON, W. **Point process models with applications to safety and reliability**. [S.l.]: Chapman and Hall, 1988. 146 p. ISBN 9780412014819. Citation on page 39.

THOMPSON, W. A. On the foundations of reliability. **Technometrics**, Taylor & Francis, v. 23, n. 1, p. 1–13, 1981. Citation on page 43.

TIBSHIRANI, R. Noninformative priors for one parameter of many. **Biometrika**, Biometrika Trust, v. 76, n. 3, p. 604–608, 1989. Citations on pages 99 and 100.

TODINOV, M. **Reliability and risk models: setting reliability requirements**. [S.l.]: John Wiley & Sons, 2015. Citation on page 133.

TOLEDO, M. L. G. de; FREITAS, M. A.; COLOSIMO, E. A.; GILARDONI, G. L. ARA and ARI imperfect repair models: Estimation, goodness-of-fit and reliability prediction. **Reliability Engineering & System Safety**, Elsevier BV, v. 140, p. 107–115, aug 2015. Citations on pages 26, 134, and 140.

TOMAZELLA, V. L.; CANCHO, V. G.; LOUZADA, F. Bayesian reference analysis for the poisson-exponential lifetime distribution. **Chilean Journal of Statistics**, v. 4, n. 1, p. 99–113, 2013. Citation on page 95.

TULI, S.; DRAKE, J.; LAWLESS, J.; WIGG, M.; LAMBERTI-PASCULLI, M. Risk factors for repeated cerebrospinal shunt failures pediatric patients with hydrocephalus. **Journal of Neurosurgery**, v. 92, n. 1, p. 31–38, 2000. Citations on pages 26, 94, and 133.

VAUPEL, J. W.; MANTON, K. G.; STALLARD, E. The impact of heterogeneity in individual frailty on the dynamics of mortality. **Demography**, Duke University Press, v. 16, n. 3, p. 439, aug 1979. Citation on page 58.

VESELY, W. Incorporating aging effects into probabilistic risk analysis using a taylor expansion approach. **Reliability Engineering & System Safety**, Elsevier Science, v. 32, 1991. Citation on page 46.

WANG, B. X.; XIE, M.; ZHOU, J. X. Generalized confidence interval for the scale parameter of the power-law process. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 42, n. 5, p. 898–906, 2013. Citation on page 63.

WANG, H.; PHAM, H. **Reliability and Optimal Maintenance**. [S.l.]: Springer London, 2006. Citation on page 62.

WANG, Z.; PAN, R. Point and interval estimators of reliability indices for repairable systems using the weibull generalized renewal process. **IEEE Access**, Institute of Electrical and Electronics Engineers (IEEE), v. 9, p. 6981–6989, 2021. Citation on page 133.

WIENKE, A. **Frailty models in survival analysis Andreas Wienke.** Boca Raton: Taylor Francis, 2010. 298 p. p. (Chapman Hall/CRC biostatistics series). Citations on pages 26, 27, 58, 59, 134, and 137.

WU, S.; SCARF, P. Two new stochastic models of the failure process of a series system. **European Journal of Operational Research**, Elsevier, v. 257, n. 3, p. 763–772, 2017. Citations on pages 64 and 133.

YAN, T. **Nonhomogeneous Poisson Process Models with a Generalized Bathtub Intensity Function for Repairable Systems**. Phd Thesis (PhD Thesis) — Ohio University, 2019. Citations on pages 94 and 133.

YUSOF, M. A. B.; ABDULLAH, N. H. B. Failure mode and effect analysis (FMEA) of butterfly valve in oil and gas industry. **Journal of Engineering Science and Technology**, Taylors University, v. 11, p. 9–19, 2016. Citations on pages 64, 76, and 77.

ZHANG WENPING; XU, D. W.-H. Fault analysis and fault-tolerant design for parallel redundant inverter systems. **IEEE Journal of Emerging and Selected Topics in Power Electronics**, p. 1–1, 2018. ISSN 2168-6777,2168-6785. Available: <http://doi.org/10.1109/JESTPE.2018.2844092>. Citation on page 116.

ZUO, M. H. H. Z. T. Reliability-redundancy allocation for multi-state series-parallel systems. **IEEE Transactions on Reliability**, IEEE, v. 57, p. 303–310, 2008. ISSN 0018-9529,1558-1721. Available: <http://doi.org/10.1109/tr.2008.920871>. Citation on page 116.

APPENDIX

A

MEAN VALUE FUNCTION BY UNITY (CHAPTER 6)

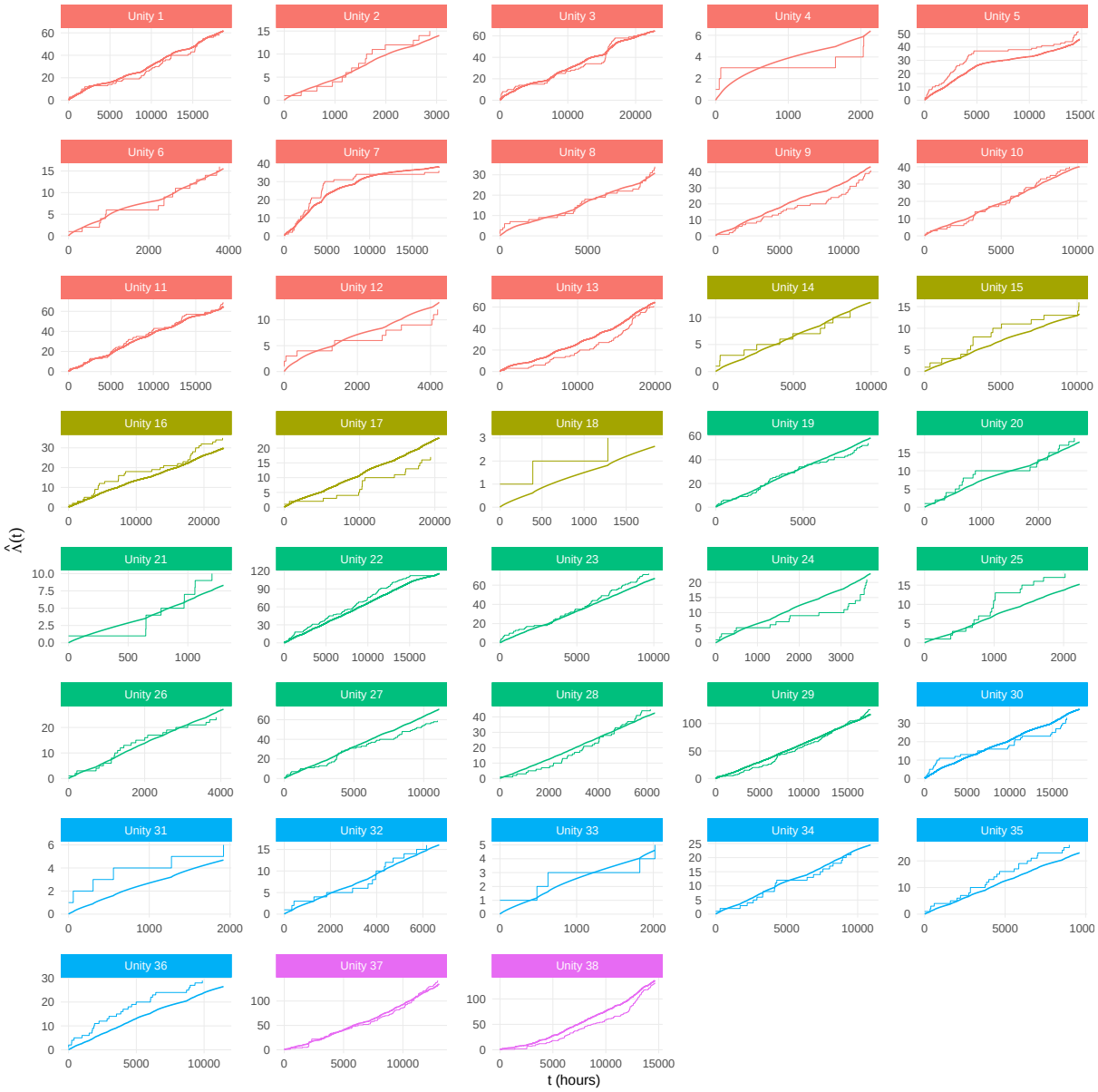


Figure 51 – Results for estimated mean value function with history of each unit over the observed accumulated failures per unit.

APPENDIX

B

GENERATED DATA AND GRAPHS (CHAPTER 4)

B.1 Generated Failure Times for Pressure Vessel Set

Table 27 – Failure data for the in-pipe robot's pressure vessel set. FT=Failure Time, FM=Failure Mode.

FM	FT
1.1.1	78.52, 100.74, 138.52, 194.66, 64.09, 193.1, 200.4, 153.93, 148.44, 26.76, 65.1, 58.1, 158.41, 103.12, 172.3, 124.86, 163.59, 207.75, 102.32, 173.55, 198.84, 66.53, 152.35, 45.17, 78.89, 103.52, 8.65, 102.78, 188.53, 94.32, 121.96, 143.26, 124.09, 60.43, 181.71, 155.24, 176.31, 40.4, 164.61, 108.46, 180.67, 151.56, 174.46, 134.97, 130.74, 175.51, 13.04
1.1.2.1	16.27, 122.49, 90.11, 14.09, 191.51, 191.4, 9.46, 158.67, 66.3, 84.62, 85.25, 23.98, 138.15, 15.74, 53.33, 164.49, 201.59, 22.02, 61.39, 4.16, 111.98, 49.84, 159.66, 11.92, 42.22, 70.79, 11.77, 44.03
1.1.2.2	28.82, 164.81, 72.37, 60.52, 118.95, 119.45, 20.67, 53.76, 113.59, 125.31, 99.35, 97.85, 104.11, 109.15, 178.57, 169.85, 18.26, 141.45, 185.34, 50.76, 40.48, 2.02, 21.48, 15, 42.2, 161.11, 118.43, 188.24, 109.84
1.1.2.3	115.38, 1.11, 53.59, 50.28, 166.19, 46.88, 146.08, 187.31, 197.2, 11.44, 152.88, 52.02, 16.2, 198.36, 78.8, 87.16, 202.29, 114.98, 200.24, 154.46, 143.86, 208.21, 98.11, 94.6, 129.32, 170.1
1.1.2.4	30.22, 132.67, 188.29, 46.09, 13.85, 136.36, 96.97, 161.73, 198.12, 14.78, 43.93, 88.76, 52.79, 103.9, 41.87, 30.52, 66.86, 181.15, 102.95, 170, 181.73, 140.97, 32.25, 34.91, 19.65, 86.46, 77.34, 200.48
1.1.2.5	133.08, 197.21, 62.95, 87.34, 172.34, 204.86, 201.05, 163.7, 113.79, 17.63, 140.44, 193.08, 25.06, 69.57, 165.09, 61.13, 115.45, 147.17, 37.13, 151.61, 201.06, 176.4, 30.71, 57.76, 159.91, 73.9, 82.38, 192.79, 65.1, 98.3, 116.23, 32.3, 200.72, 163.17, 34.77, 195.92, 88.6, 16.33, 206.43
⋮	⋮
1.2.1	207.29, 105.82, 42.52, 29.27, 73.71, 175.95, 94.26, 204.67, 55.48, 117.64, 56.92, 70.96, 172.79, 37.14, 116.57, 33.78, 136.34, 6.3, 206.8, 89.41, 150.24, 84.91, 208.49, 194.31, 207.26, 27.99, 148.08, 123.52, 204.51, 98.75, 157.21, 78.13, 60.31, 60.17, 102.17, 184.89, 124.93, 175.71, 183.51, 117.22, 177.17, 102.7, 170.62, 113.39, 194.03, 90.02, 33.15, 179.91, 193.12, 203.81, 138.56, 164
1.2.2.1	65.93, 19.42, 149.04, 109.42, 37.57, 126.89, 32.29, 187.97, 140.53, 107.55, 63.96, 5.41, 58.81, 83.46, 10.48, 186.61, 0.01, 27.96, 29.92, 77.99, 21.71, 53.86, 100.17, 20.44
⋮	⋮
1.11.1	96.64, 200.06, 97.07, 9.99, 20.34, 53.48, 149.82, 31.09, 33.33, 83.16, 111.18, 78.13, 170.53, 123.02, 159.99, 105.51, 195.78, 43.25, 101.43, 104.51, 38.35, 147.37, 30.76, 30.71, 158.36, 89.5, 205.04, 172.05, 38.19, 101.54, 26.03, 202.24, 150.82, 201.92, 175.19, 175.11, 165.33, 38.18, 59.92, 13.45, 135.01, 149.91, 103.11, 127.89, 84.01, 208.04
1.11.2.1	145.91, 41.99, 99.45, 126.07, 3.36, 9.66, 79.96, 32.72, 205.25, 119.29, 78.81, 200.44, 154.69, 66.06, 7.9, 163.97, 81.99, 117.03, 13.31, 0.19, 14.61, 1.61, 165.48, 29.62, 102.84, 195.58, 3.03
1.11.2.2	187.58, 34.93, 4.33, 12.32, 138.12, 149.23, 4.05, 36.81, 142.37, 12.12, 23.88, 3.15, 130.9, 163.25, 108.15, 93.94, 7.35, 195.78, 132.64, 114.92, 129.28, 105.64, 84.55, 93.61, 125.84, 121.58, 112.14, 79.88, 76.43, 138.03, 115.44, 185.25, 65.01, 166.73, 167.5, 8.97, 169.94
1.11.2.3	6.47, 169.29, 6.4, 206.52, 14.37, 167.12, 194.73, 122.92, 98.07, 115.53, 69.74, 10.5, 169.85, 125.84, 70.29, 27.03, 192.16, 197.51, 182.73, 201.5, 134.8, 159.86, 9.32, 52.69, 37.55, 141.41, 134.49, 86.21, 12.66, 39.97, 1.98, 167.49, 79.1, 70.14
1.11.2.4	15.63, 133.32, 24.94, 4.71, 115.3, 42.73, 9.17, 95.82, 158.47, 175.76, 182.61, 26.35, 58.47, 193.1, 14.06, 56.57, 40.83, 31.97, 16.47, 97.55, 122.38, 148.14, 181.85, 149.52, 127.72, 8.69, 3.39, 109.74, 70.07, 90.58, 182.92, 45.04, 194.27, 103.21, 13.6, 1.37, 0.78, 117.43
1.11.2.5	207.1, 176.61, 129.04, 94.96, 144.09, 87.18, 96.44, 179.07, 86.58, 192.25, 61.7, 49.66, 54.34, 192.84, 188.19, 66.2, 97.49, 65.02, 114.36, 82.18, 32.69, 69.36, 40.34, 113.19, 86.85, 45.7, 25.58, 179.42, 42.75, 7.49, 126.47, 55.88, 88.38, 37.73

B.2 Duane Plots for Pressure Vessel Set failure modes

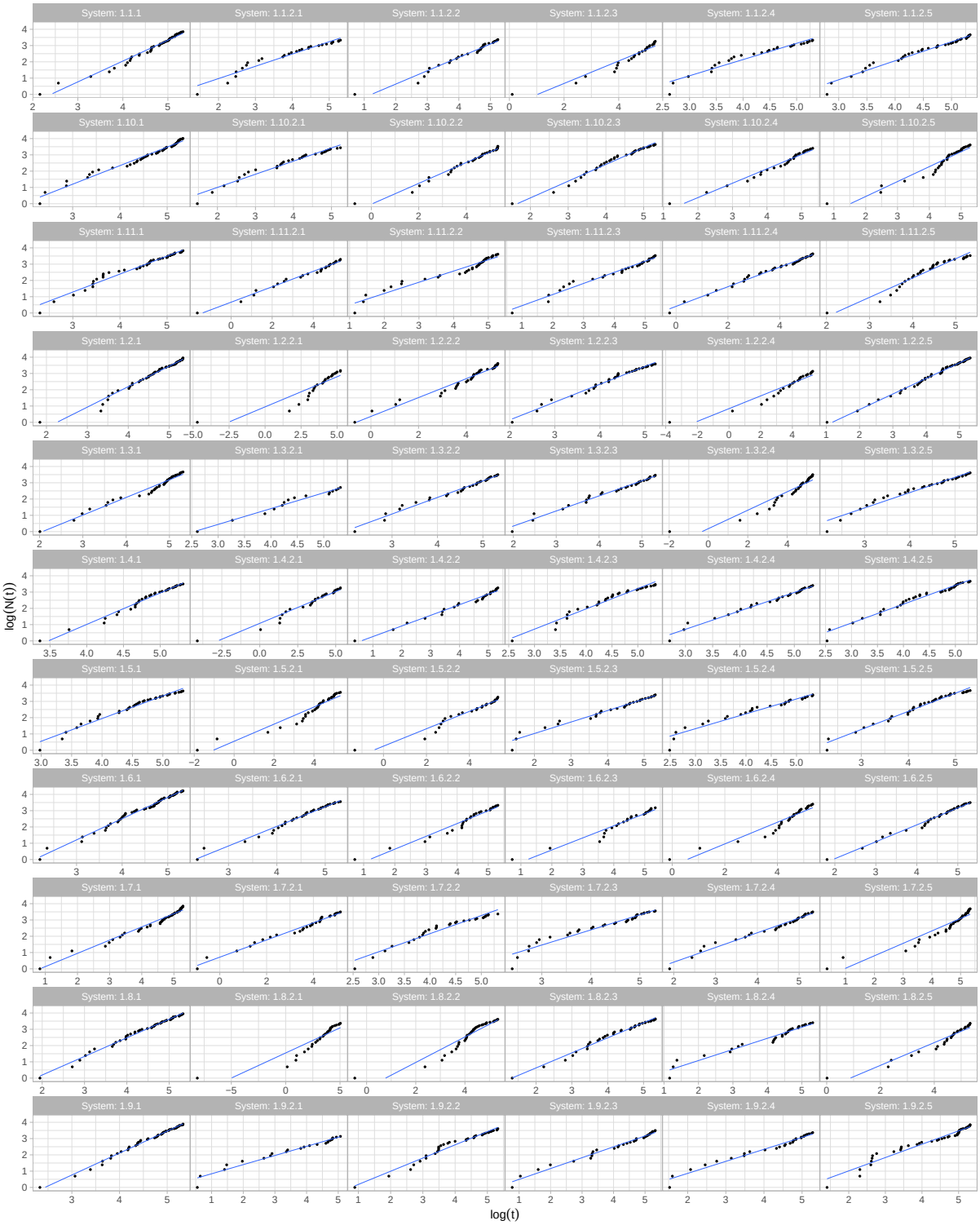


Figure 52 – Duane plots for the failure modes related to the in-pipe robot's pressure vessel system.

B.3 Number of observed and estimated failures per component

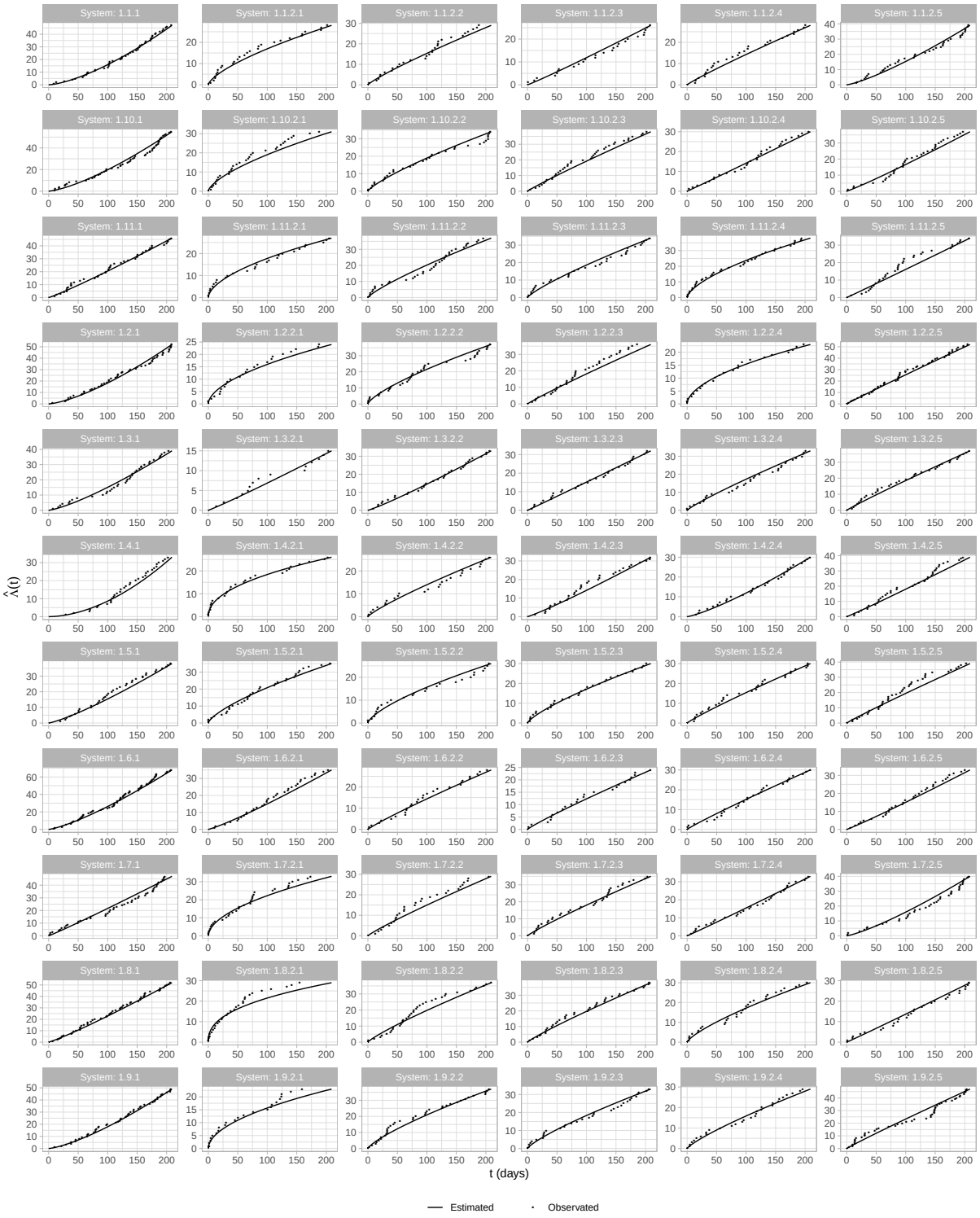


Figure 53 – Number of observed and estimated failures per component, for the in-pipe robot's pressure vessel system.

B.4 FTA with FMEA index for Traction System

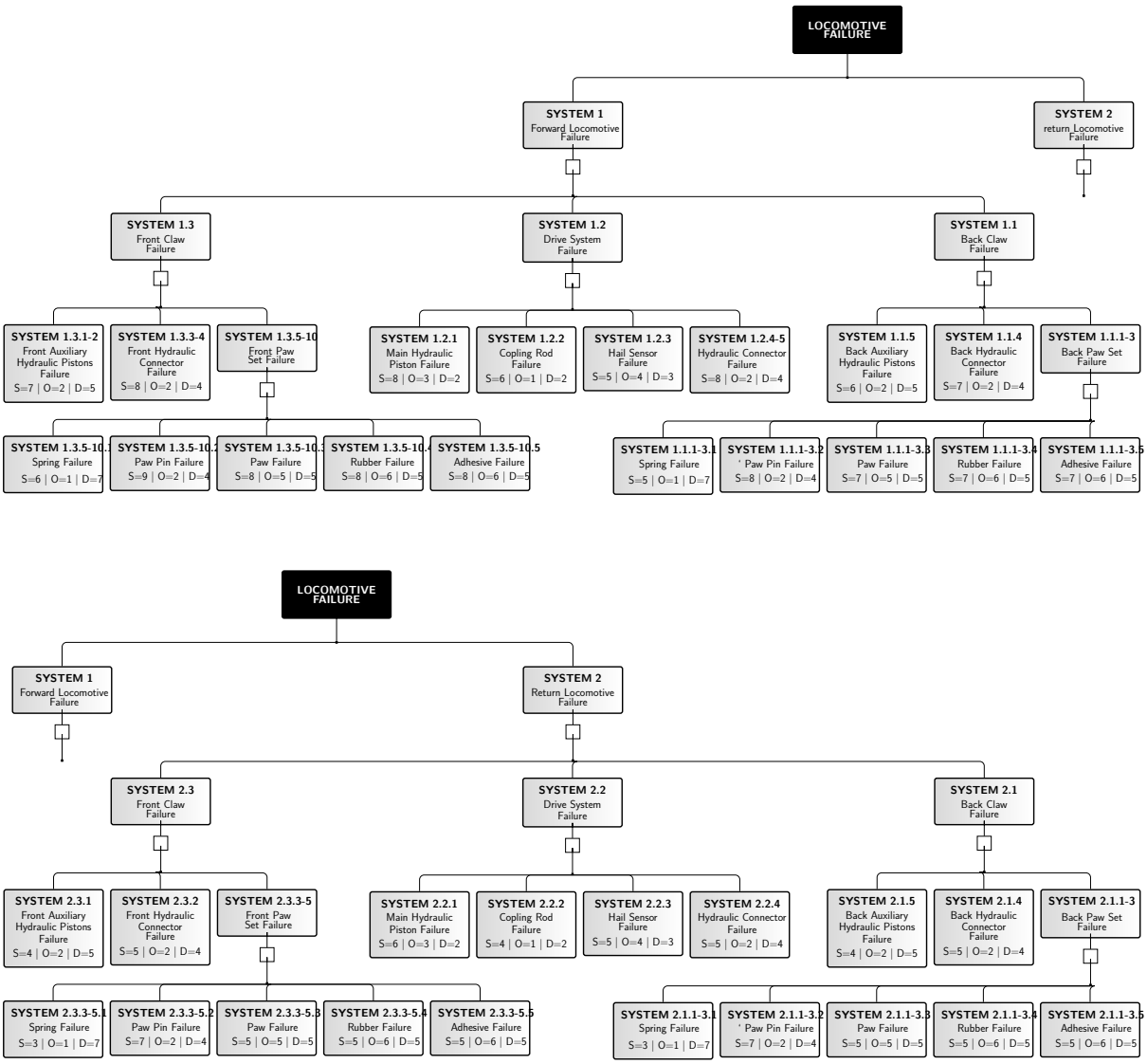


Figure 54 – FTA (with FMEA indices) for the in-pipe robot's traction system.

B.5 Generated Failure Times for Traction System

Table 28 – Failure data for the in-pipe robot's traction system. FT=Failure Time, FM=Failure Mode.

FM	FT
1.1.1.1	176.09, 107.53, 165.17
1.1.1.2	166.4, 86.33, 152.82, 153.65, 71.53, 28.49, 103.93, 189.07, 14.05, 173.56, 14.35, 66.54
1.1.1.3	159.1, 72.54, 143.68, 144.61, 57.91, 19.24, 90.57, 185.37, 8.26, 167.33, 8.47, 53.12, 179.16, 70.77, 130.08, 144.52, 37.15, 93.39, 139.13, 63.04, 86.58, 89.36, 102.25, 111.1, 190.72, 122.86, 56.74, 65.76, 71.29
1.1.1.4	165.76, 85.05, 152.01, 152.85, 70.24, 27.54, 102.7, 188.75, 13.42, 173.02, 13.71, 65.26, 183.36, 83.28, 139.69, 152.76, 48.17, 105.42, 147.91, 75.49, 98.85, 101.54, 113.86, 122.17, 193.36, 133.07, 69.03, 78.25, 83.8, 129.49, 197.76, 99.09, 99.21, 2.67, 158.08, 162.12, 88.86, 77.63, 148.11, 112.3, 91.79, 43.94, 76.51
⋮	⋮
1.2.1	157.84, 70.34, 142.12, 143.08, 55.79, 17.95, 88.4, 184.73, 7.52, 166.25, 7.71, 51.04, 178.36, 68.58, 128.3, 142.98, 35.33, 91.23, 137.5, 60.88, 84.4, 87.19, 100.15, 109.08, 190.21, 120.97, 54.63, 63.59
1.2.2	175.06, 105.12, 163.85
1.2.3	162.64, 79.01, 148.09, 148.98, 64.24, 23.33, 96.9, 187.18, 10.72, 170.36, 10.97, 59.33, 181.4, 77.24, 135.16, 148.89, 42.71, 99.67, 143.78, 69.45, 92.97, 95.71, 108.33, 116.92, 192.13
⋮	⋮
1.3.1	168.12, 89.84, 154.98, 155.79, 75.06, 31.16, 107.25, 189.92, 15.87, 175.02, 16.19, 70.06, 184.83, 88.09
1.3.2	168.12, 89.84, 154.98, 155.79, 75.06, 31.16, 107.25, 189.92, 15.87, 175.02, 16.19
1.3.3	166.4, 86.33, 152.82, 153.65, 71.53, 28.49, 103.93, 189.07, 14.05, 173.56, 14.35, 66.54, 183.76, 84.57, 140.64
⋮	⋮
2.1.1	173.2, 100.84, 161.46, 162.18, 86.35, 40.43
2.1.2	171.47, 96.98, 159.24, 160, 82.36, 37.03, 113.93, 191.58
2.1.3.1	178.54, 113.45, 168.34, 168.97
2.1.3.2	168.12, 89.84, 154.98, 155.79, 75.06, 31.16, 107.25, 189.92, 15.87, 175.02, 16.19, 70.06, 184.83, 88.09, 143.17, 155.71
2.1.3.3	156.28, 67.68, 140.19, 141.17, 53.24, 16.46, 85.74, 183.92, 6.68, 164.91, 6.86, 48.56, 177.36, 65.92, 126.1, 141.07, 33.18, 88.59, 135.48, 58.28, 81.73, 84.53, 97.58, 106.59, 189.58, 118.65, 52.1, 60.97
⋮	⋮
2.2.1	164.6, 82.76, 150.54, 151.4, 67.95, 25.9, 100.51, 188.16, 12.35, 172.03, 12.62, 63, 182.63, 80.99, 137.99, 151.32, 46.07, 103.25, 146.37, 73.19, 96.63
2.2.2	177.22, 110.24, 166.63, 167.29, 96.2
⋮	⋮
2.3.1	173.2, 100.84, 161.46, 162.18, 86.35, 40.43, 117.5, 192.43, 22.58, 179.32
2.3.2	171.47, 96.98, 159.24, 160, 82.36, 37.03, 113.93, 191.58, 20.05, 177.86, 20.41, 77.36
2.3.3.1	177.22, 110.24, 166.63, 167.29, 96.2
2.3.3.2	168.12, 89.84, 154.98, 155.79, 75.06, 31.16, 107.25, 189.92, 15.87, 175.02, 16.19, 70.06, 184.83, 88.09, 143.17, 155.71
2.3.3.3	151.32, 59.72, 134.11, 135.15, 45.74, 12.41, 77.67, 181.33, 4.56, 160.63, 4.7, 41.29, 174.16, 58, 119.22, 135.05, 27.04, 80.54, 129.11, 50.58, 73.64, 76.44, 89.66, 98.91, 187.54, 111.42, 44.65, 53.17, 58.5, 107.27, 193.51, 73.88, 74.01
⋮	⋮
2.3.5.4	160.21, 74.52, 145.06, 145.98, 59.84, 20.45, 92.52, 185.94, 8.97, 168.28, 9.19, 55.01, 179.87, 72.75, 131.66, 145.89, 38.82, 95.33, 140.59, 65, 88.55, 91.32, 104.14, 112.9, 191.16, 124.53, 58.66, 67.73, 73.27, 120.7, 196.16, 88.8, 88.92, 1.4, 151.73, 156.19, 78.37, 67.12, 140.8, 102.5
2.3.5.5	160.21, 74.52, 145.06, 145.98, 59.84, 20.45, 92.52, 185.94, 8.97, 168.28, 9.19, 55.01, 179.87, 72.75, 131.66, 145.89, 38.82, 95.33, 140.59, 65, 88.55, 91.32, 104.14, 112.9, 191.16, 124.53, 58.66, 67.73, 73.27, 120.7, 196.16, 88.8, 88.92, 1.4, 151.73, 156.19, 78.37, 67.12, 140.8, 102.5

B.6 Duane Plots for Traction System failure modes

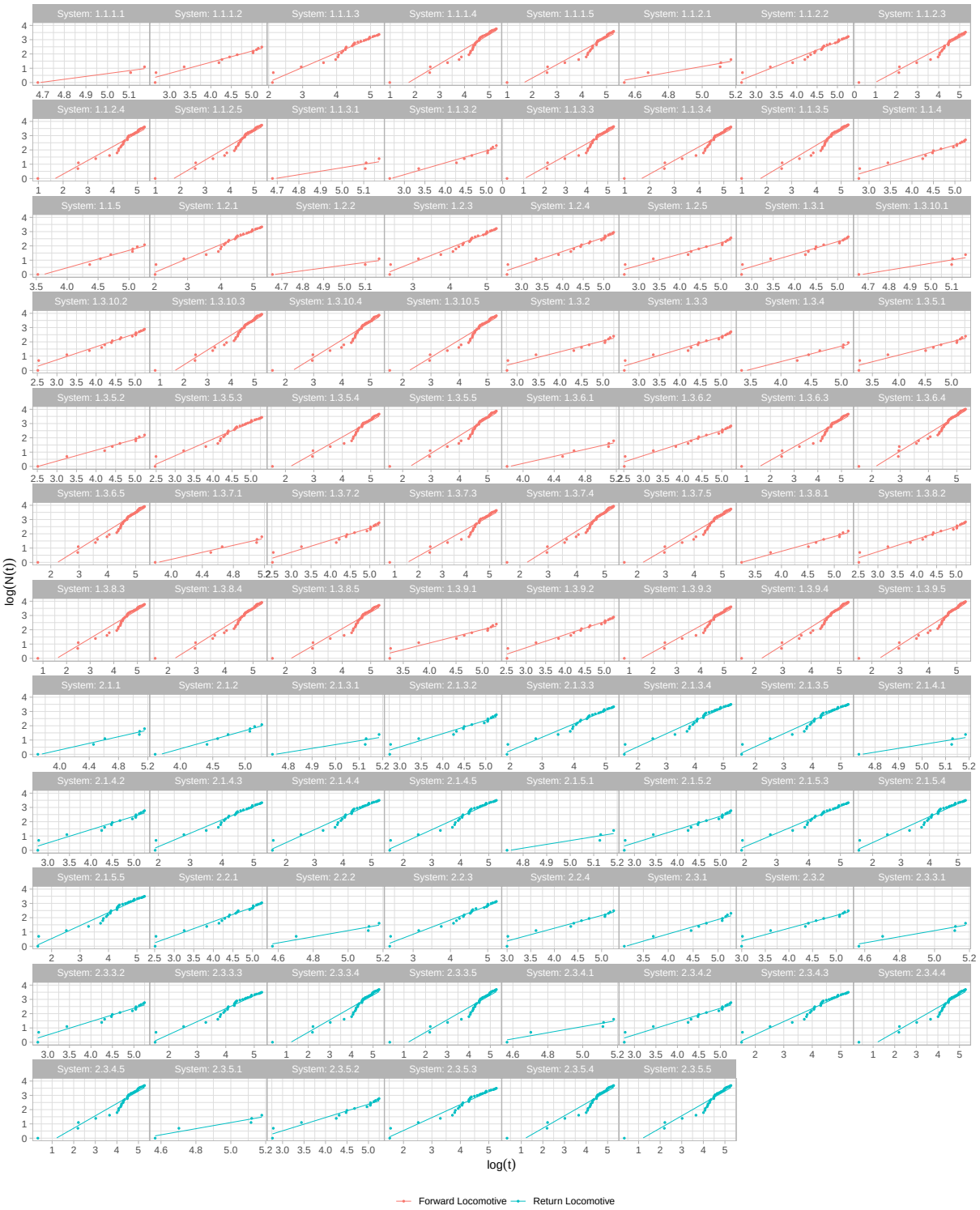


Figure 55 – Duane plots for the failure modes related to the in-pipe robot's traction system.

B.7 Number of observed and estimated failures per component, for the in-pipe robot's traction system.

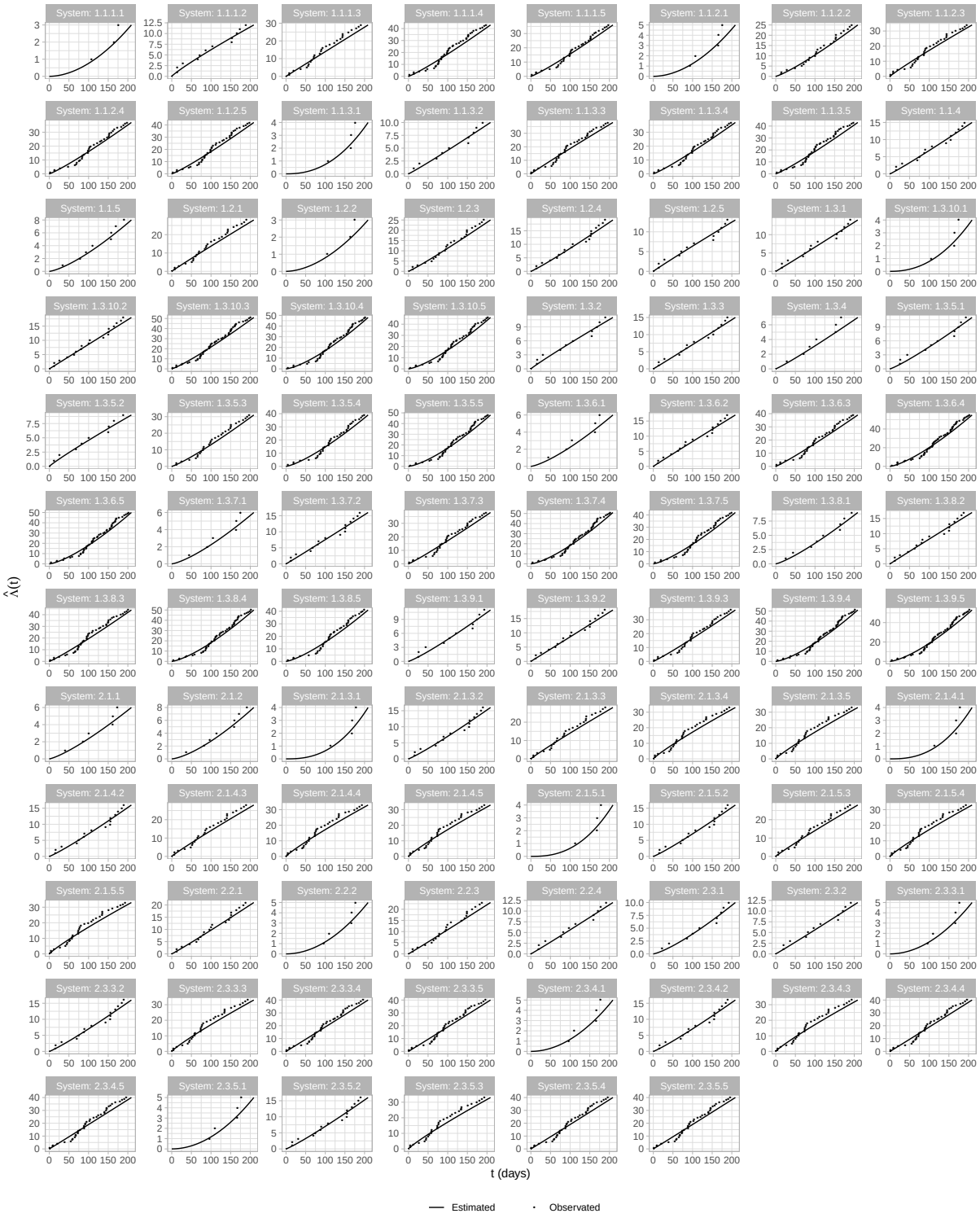


Figure 56 – Number of observed and estimated failures per component, for the in-pipe robot's traction system.

B.8 Structure of the systems illustrated in scenarios 1 to 6

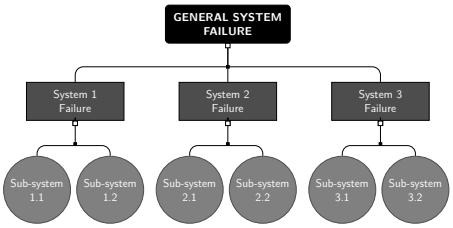


Figure 57 – Simulated failure structure for Scenarios 1, 2 and 3.

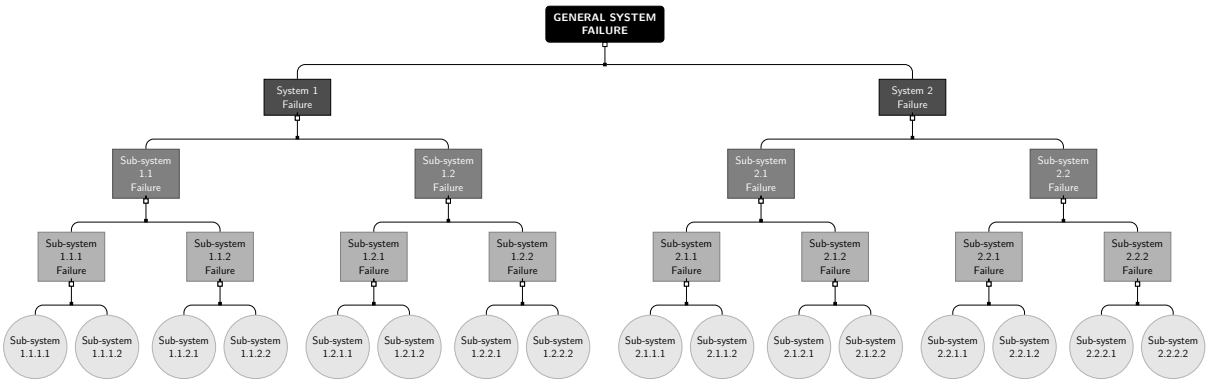


Figure 58 – Simulated failure structure for Scenarios 4, 5 and 6.


ICMC
USP
SÃO CARLOS