Reliability and Maintenance Optimisation in the Age of Data-Centric Engineering

Jordan Lee Oakley

Thesis submitted for the degree of
Doctor of Philosophy

School of Mathematics, Statistics & Physics
Newcastle University
Newcastle upon Tyne
United Kingdom

February 2024

I wholeheartedly dedicate this thesis to Olive Oakley, my mother, and to Terence Patrick Oakley, my late father, for their love and support.

Acknowledgements

Thank you to the Engineering and Physical Sciences Research Council (EPSRC), Centre for Doctoral Training in Cloud Computing and Big Data (grant number EP/L015358/1) for providing the funding for me to do my PhD. Thank you to my supervisors, Dr Matthew Forshaw, Dr Pete Philipson and Dr Kevin Wilson for their guidance, expertise and support. Thank you to my teachers at the Centre for Doctoral Training in Cloud Computing and Big Data. Thank you to my teachers and supervisors during my MMATH here at Newcastle and my research masters at Durham University. Thank you to my friends and family for their encouragement. Finally, an extra thank you to two family members, Olive Oakley, my mother, and Terence Patrick Oakley, my late father, for their love, support, encouragement, and for always believing in me. I wholeheartedly dedicate this thesis to my parents.

Abstract

Data-centric engineering is drastically impacting all areas of engineering and industry. The goal of data-centric engineering is to make data science, mathematics, statistics, and machine learning fundamental to engineering practice, leading to engineered products that are more intelligently designed, monitored and maintained, more reliable, more cost efficient, and safer to use.

Reliability and maintenance optimisation are two important research areas being impacted by data-centric engineering. The large amounts of dynamic data being collected by data-driven technology has the potential to provide accurate real-time information about the state of products, allowing for the health (or reliability) of products to be continuously monitored. The continuous monitoring of the reliability of products allows us to more accurately plan maintenance, reducing costs and increasing safety.

This thesis has two main contributions. One is in the field of reliability for hard-disk drives with automatic data-collecting devices and one is in the field of condition-based maintenance for complex continuously monitored multi-component systems with dependencies.

In Part I of this thesis, we propose a novel way to model the survival probabilities and failure times of hard drives, using data collected by SMART (an automatic data-collecting device). We define critical states for hard drives using data collected by SMART and model hard drive failure times using multi-state models. Using the proposed multi-state models, we seek to concretely define the impact of critical attributes on the failure time of a hard drive.

In Part II of this thesis, we propose a novel condition-based maintenance policy for continuously monitored multi-component systems subject to economic and stochastic dependence. More specifically, we propose a novel loss-based utility function, that is incorporated in a Bayesian sequential decision framework, to decide which components are to be maintained at maintenance opportunities for continuously monitored multi-component systems that are subject to economic and stochastic dependence.

Contents

List of Algorithms

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is concerned with reliability and maintenance optimisation in the age of data-centric engineering. Data-driven technology is impacting all aspects of engineering and industry. Modern technological developments, such as smart chips, sensors, and monitoring systems, have changed data-collection processes. More and more products are being produced with automatic data-collecting devices that track how and under which environments the products are being used. There is a tremendous amount of dynamic data being collected and hence enormous potential for such data to provide more accurate information about the state of products and systems. More accurate information has the potential to improve safety and to reduce costs. For these reasons, the field of data-centric engineering is a pioneering area for research and is shaping all aspects of engineering and technology.

Reliability and maintenance optimisation are two important research areas being impacted by data-centric engineering. Reliability is defined as the probability that an item, product, or system can perform a required function, under given environmental and operating conditions, for a specified period of time (Hamada et al., 2008; Meeker et al., 2022). Maintenance is concerned with the process of preventively repairing or correctively replacing a system, or a subset of a system, in order for the system to be able to perform a required function, under given environmental and operating conditions, for a specified period of time (Ben-Daya et al., 2016). Maintenance decision making consists of deciding the times to perform maintenance and deciding which parts of a system to repair or replace (Keizer et al., 2017; de Jonge and Scarf, 2019). These decisions depend on the reliability of the components of the system and the dependencies between components (Keizer et al., 2017; de Jonge and Scarf, 2019).

Maintenance costs are a major part of the total operating costs for all manufacturing or production plants (Zio and Compare, 2013). Mobley (2002) states that maintenance costs can represent between 15 and 60 percent of the cost of goods produced, and that up to a third of these costs may be due to unnecessary or poorly executed maintenance. In both the process industry and the chemical industry over a quarter of the total workforce deal with maintenance operations (Waeyenbergh and Pintelon, 2002). The operations and maintenance costs for offshore wind farms contribute a quarter of the life-cycle costs, making maintenance one of the largest cost components for offshore wind farms (Snyder and Kaiser, 2009; Irawan et al., 2017). Hard disk drives (HDDs) are a key driving factor behind enabling the use of the large amounts of data collected by data-driven technology. However, HDDs are not only the most frequently replaced hardware components of a data centre, they are also the main reason behind server failures (Vishwanath and Nagappan, 2010), which can result in large amounts of data loss; making accurate maintenance planning a major concern for data centres.

For these reasons, research in the field of maintenance planning is of critical importance and has been extensively studied over the past decades. Reviews of maintenance strategies have been written by McCall (1965), Wang (2002), Van Horenbeek et al. (2013), Keizer et al. (2017), and de Jonge and Scarf (2019). Research in the field of reliability is of equal importance, since maintenance optimisation depends crucially on the reliability of the system being studied.

Continuous monitoring of a system's health, using data collected by sensors or monitoring systems, is playing an increasingly important role in the fields of reliability and maintenance optimisation. The advances in sensor technologies have greatly accelerated the use of real-time monitoring and condition assessment for manufacturing and production systems. Condition-based maintenance (CBM) has been shown to be an effective way to minimise maintenance costs, improve operational safety and reduce the frequency and severity of in-service system failures (de Jonge and Scarf, 2019). Prognostic methods based on sensor information for reliability or remaining useful life (RUL) prediction have been extensively studied in recent years (Pang et al., 2021; Zhang et al., 2021; Li et al., 2021). It has been shown that incorporating prognostic reliability information in maintenance planning can help make more informed maintenance decisions for single-component systems (Zhou et al., 2007; You et al., 2010; Fauriat and Zio, 2020; Zheng et al., 2020). On the other hand, CBM for complex multi-component systems with dependencies is an underexplored area. Modern systems are becoming more and more complex and their operational environments are often dynamic. Many complex systems consist of a large number of interconnected components with dependencies.

This thesis makes two main contributions. One is in the field of reliability for HDDs with automatic data-collecting devices and one is in the field of CBM for complex multi-component systems with dependencies. The first contribution, described in detail in Part I, provides a novel way to model the survival probabilities and failure times of hard drives using dynamic data collected by the drives. The methodology allows us to specify the impact of critical attributes on the failure time of a hard drive. The work described in Part I has been accepted for part of a special issue on degradation and maintenance, modelling and analysis in Applied Stochastic Models in Business and Industry (ASMBI).

The second contribution, described in detail in Part II, provides a novel loss-based utility (or reward) function, that is incorporated in a Bayesian sequential decision framework, to decide which components are to be maintained at maintenance opportunities for continuously monitored multi-component systems that are subject to economic and stochastic dependence. The work described in Part II was published in January 2022 as part of a special issue on maintenance planning in Reliability Engineering & System Safety (Oakley et al., 2022).

1.1 Part I

In Part I, we propose a coherent and novel way to model the survival probabilities and failure times of hard drives, which allows us to examine the impact of critical attributes on hard drive survival probabilities and failure times.

A recent study based on data from Microsoft reports that $76-95\%$ of all failed components in data centres are hard drives (Manousakis et al., 2016). HDDs are the main reason behind server failures (Vishwanath and Nagappan, 2010). Consequently, research in hard drive failure prediction is critically important and has been extensively studied over the past decades. Predicting drive failures before they occur can inform us to take action in advance.

Most HDDs are equipped with a monitoring system named SMART (Self-Monitoring, Analysis, and Reporting Technology). The primary function of SMART is to detect and report various indicators of drive reliability, with the intention of anticipating imminent hardware failures. At present, SMART is implemented inside most modern hard drives. However, as Murray et al. (2005) and Lu et al. (2020) reported, SMART alone does not lead to accurate predictions of failures. Moreover, in addition to whole-drive failures that make an entire drive unusable, modern drives can exhibit latent sector errors, reallocated sector counts, and many other read/write errors. The effect of such errors on the failure rates of hard drives is poorly understood (Ma et al., 2015). Empirical observations show that the failure rates of hard drives increase after their first scan error (Pinheiro et al., 2007). First errors in reallocations and probational counts are also strongly correlated to higher failure probabilities (Pinheiro et al., 2007). Pinheiro et al. (2007) discuss, but do not present, predictive models using SMART attributes; concluding that it is unlikely that SMART data alone can be effectively used to build models that predict failures of individual drives. We seek to provide a novel way of modelling the survival probability of a hard drive using data collected by SMART and to use these predictive models to concretely provide a measure of how much critical SMART attributes weaken a drive. For example, how is the survival probability of a hard drive, over a forecast horizon of interest, impacted when a hard drive records a reallocated sector count, or multiple reallocated sector counts? We seek to provide predictive models to answer these questions.

Statistical and machine learning models have been proposed based on SMART attributes to improve failure prediction accuracy. Most papers approach hard drive failure prediction from a classification point of view, classifying drives as failed or not failed within a specific time horizon. Moreover, as mentioned by Lu et al. (2020), most papers classify hard drive failures using a prediction horizon of a few hours, days, or weeks. It could be of value to provide accurate models that can predict hard drive failure over longer time intervals. It would be valuable to provide estimates, with uncertainty, of the survival probabilities and RUL of drives. Lu et al. (2020) compare the prediction quality of machine learning models with different groups of SMART, performance and location features. The machine learning models classify hard drive failures 2-15 days in advance. Chhetri et al. (2022) use machine learning methods and knowledge graphs to predict hard drive failures. Drives are classified one day in advance. Botezatu et al. (2016) use regularised greedy forest classifiers to predict hard drive failures 10-15 days in advance. Shen et al. (2018) use part-voting random forests to predict hard drive failures. They compare part-voting random forests to classification trees and recurrent neural networks. The classifications are made one week in advance.

Far fewer papers approach hard drive failure prediction from a probabilistic or RUL point of view. Mittman et al. (2019) propose a hierarchical model to obtain the lifetime distributions of hard drives from different brands. Their approach borrows strength across brands with many observed failures to help make inferences for those brands with few failures. Their approach does not incorporate the (covariate) attributes collected by SMART. Chaves et al. (2018) obtain the RUL estimates of HDDs using SMART attributes and a Bayesian Network. dos Santos Lima et al. (2017) present a RUL estimation approach for hard drives using Long Short-Term Memory (LSTM) networks. Our goal is to propose a novel way to model the survival probabilities and failure times of hard drives, using data collected by SMART, and to concretely define the impact of critical attributes on hard drive failure times. Probabilistic models are advantageous in answering the question "How much do critical attributes weaken drives?" as we will be able to directly compare probabilities of survival/failure of drives, over forecast horizons of interest, to concretely answer this question. In addition, when multiple hard drives fail simultaneously, a RAID (multiple hard drives grouped together to decrease the risk of data loss) may lose data (Ma et al., 2015). Probabilistic models can provide the probability of data loss, which would be of value to data centres. Even if a classification model predicts that no individual drive is going to fail over a chosen forecast horizon, the probability of data loss may still be alarming to data centres.

As mentioned by Lu et al. (2020), SMART attributes do not always have strong predictive capabilities for hard drive failures over longer prediction horizons (i.e., predicting drive failure several weeks or months before the actual failure instead of a few hours or days in advance). This is primarily because the values of SMART attributes do not change frequently enough during the period leading up to the failure, and the change is often noticeable only a few hours before the actual failure, especially in hard-to-predict cases. In Part I, we define critical attributes and critical states for hard drives using SMART attributes. We utilise the probabilities of changes in critical attributes and the age of a hard drive to predict the probability of drive failure over time. We treat changes in critical attributes as drives entering critical states. Using the age of the drive and the state of the drive allows us to identify drives at risk of failure, even when the values of SMART attributes have been stationary for a long time. Part I presents a general framework for modelling reliability field data collected from SMART for a population of hard drives, treating changes in critical attributes as drives entering critical (non-terminal) states. In this setting, the data collected by SMART can be considered semi-competing risks data.

Semi-competing risks refers to the setting where interest lies in the time-to-event for some terminal event, the observation of which may be subject to some non-terminal event(s) (Fine et al., 2001). In contrast to competing risks, where each of the outcomes under consideration is typically terminal, in the semi-competing risks setting, it is possible to observe multiple events in the same study unit, providing at least partial information on the joint distribution (Fine et al., 2001; Xu et al., 2010). Towards the analysis of semi-competing risks data, the statistical literature has focused on three broad frameworks that seek to exploit the joint information on the times to the non-terminal and terminal events (Varadhan et al., 2014). Those based on copulas (Fine et al., 2001; Peng and Fine, 2007; Hsieh et al., 2008; Lakhal et al., 2008); those framed from the perspective of causal inference (Egleston et al., 2007; Tchetgen Tchetgen, 2014); and, those based on the illness-death model (Xu et al., 2010; Liu et al., 2004; Putter et al., 2007; Lee et al., 2015, 2020, 2021). In Part I, we focus on the last of these approaches, for which the underpinning idea is that drives begin in some initial (healthy) state and may transition into the non-terminal (critical) state(s) and/or the terminal (failed) state. Analyses typically proceed through the development of models for transition-specific hazard functions, which dictate the rate at which units experience the respective events.

In the analysis of time-to-event outcomes, data are subject to left-truncation, or delayed entry, when units are enrolled into a study. In addition, units are subject to right-censoring. Left-truncation is common in large populations of units where units are required to be in a working state at an enrollment time in order to be included in a study. In this setting, sampling is biased since units are only included in the study if they are in a working state on entry. The analysis of left-truncated time-to-event data should apply statistical methods that account for this bias. Moreover, right-censoring is common in reliability studies when, due to advancements in technology, units are retired and replaced by newer technology, or because many units simply will not have failed at the end of the study. Not accounting for right-censoring will also produce biased estimates.

The contributions of Part I are as follows. We define critical attributes and transient states, named critical states, for hard drives using data collected by SMART. We model the semi-competing risks data (entry to the critical state(s) and failure) using multi-state models. The proposed multi-state models provide a coherent and novel way to model the failure times of hard drives and allow us to statistically examine the impact of critical attributes on hard drive failure times. The multi-state models utilise the probability of a change in critical attributes, the age and the state of a hard drive to predict the probability of drive failure over time. Using the age of the drive and the state of the drive allows us to identify drives at risk of failure, even when the values of critical attributes have been stationary for a long time. We illustrate how multi-state models can be used to obtain distributions of remaining life (DRLs) for hard drives using the age and state of a drive and compare our results to previous work by Mittman et al. (2019). The DRLs play an important role in health monitoring and making maintenance decisions. Our motivating example concerns a large dataset of hard drives, from data backup company Backblaze (Backblaze, 2022a), that is subject to left-truncation and right-censoring.

The structure of Part I is as follows. In Chapter 2, we introduce reliability concepts. We define commonly used functions to describe the reliability of a product and provide examples of reliability data. Moreover, we introduce censoring and truncation in the context of reliability and describe the likelihood contributions for censored observations and the likelihood adjustments for truncated observations. We conclude the chapter by introducing Bayesian inference and Bayesian computation. In Chapter 3, we introduce multi-state models and extend the standard reliability methodology described in Chapter 2. We present the general form of left-truncated and right-censored data and derive the likelihood and the DRLs for three multi-state models. In Chapter 4, we present discrimination and calibration measures to assess model performance. In Chapter 5, we provide a novel way to model the reliability of hard drives, utilising data collected by SMART, based on the methodology presented in Chapter 3. The proposed methodology enables us to identify the impact of critical attributes on hard drive failure times. We compare our results to previous work by Mittman et al. (2019) using the methods presented in Chapter 4. Appendix A.1 provides supplementary material for Part I.

1.2 PART II

In Part II, we propose a novel CBM policy for continuously monitored multi-component systems subject to economic and stochastic dependence. More specifically, we propose a novel loss-based utility (or reward) function, that is incorporated in a Bayesian sequential decision framework, to decide which components are to be maintained at maintenance opportunities for continuously monitored multi-component systems that are subject to economic and stochastic dependence.

Traditionally, three types of dependence between units are distinguished in the maintenance literature. Namely, economic dependence, stochastic dependence and structural dependence (Thomas, 1986; Laggoune et al., 2010). Economic dependence exists when the cost of maintaining or inspecting multiple units simultaneously is different from the sum of the costs of maintaining or inspecting these units separately, for instance due to a fixed set-up cost. Economic dependence has been well-studied and substantial cost savings can be made by grouping maintenance activities to reduce overall cost.

Martinod et al. (2018) implement an opportunistic maintenance policy in multi-component systems with economic dependence. A stochastic optimisation model is used to reduce the long-term total maintenance cost of complex systems. Do and Bérenguer (2020) define an importance measure of a group of components as its ability to improve the system reliability during a mission given the current conditions (states or degradation levels) of the components. An extension of the proposed importance measure is investigated to incorporate economic dependence. Liu et al. (2017) develop a maintenance policy for multi-component systems subject to hidden failures. Components are assumed to suffer from hidden failures, which can only be detected at inspection. A common cost is incurred at each inspection time, which can be shared when multiple inspections are carried out simultaneously. Arts and Basten (2018) study a periodic maintenance policy and a CBM policy in which the scheduled maintenance downtime can be coordinated between components. Liu et al. (2021) analyse a finite-horizon CBM policy for a two-unit system with dependent degradation processes. A set-up cost is incurred whenever maintenance is performed. The set-up cost can be shared when multiple maintenance actions are carried out simultaneously.

Stochastic dependence arises when the state of a component influences the deterioration processes or lifetime distributions of other components or when components are subjected to common-cause failures. This is often observed for redundant mechanical systems where the degradation of a component leads to internal force redistribution, which can overload other components. Component failures in load-sharing systems increase the workload on the remaining components, and consequently the failure-rates of these components also increase. Keizer et al. (2018) consider a parallel system that is subject to both stochastic dependence through load sharing and economic dependence through maintenance set-up costs. The parallel system is formulated as a Markov decision process, and the optimal replacement decisions that minimise the long-run average cost per unit time are obtained. Brown et al. (2022) develop and evaluate a load-sharing system with spatial dependence and proximity effects. If a component fails, its load is taken up by its working spatial neighbours in close proximity. Liu et al. (2016) develop two reliability models for assessing the reliability of load-sharing systems with continuously degrading components. The proposed models are used to formulate preventive maintenance policies. The system load is distributed equally between all working components. Component failures increase the load on the remaining working components, and consequently the degradation rates of these components also increase. Zhang et al. (2018) consider two versions of a two-component system with failure interactions. Failure of the first component either causes a random amount of damage to the other component, or it results in failure of the other component with a certain probability. Three preventive maintenance policies are analysed numerically. Common-mode stochastic dependence is considered by Liu et al. (2020). Components in systems subject to common-mode stochastic dependence follow similar deterioration or failure patterns when operating in a common environment. An increase in the degradation of one component is usually accompanied with a degradation increase in the other components. Liu et al. (2020) present a life cycle cost model for systems subject to multiple dependent degradation processes and environmental influence. The degradation dependence between components is modelled using copulas.

Keizer et al. (2017) make a distinction between structural dependence from a technical point of view and structural dependence from a performance point of view. Structural dependence from a technical point of view exists, for instance, if maintenance of a unit requires other units or subsystems to be dismantled as well. This may induce deterioration or failure of these other units. Nguyen et al. (2015) propose a predictive maintenance policy with multi-level decision making for multi-component systems with complex structures. Selecting optimal components is based on a cost-based improvement factor taking into account the predictive reliability of the components, the economic dependence, and the location of the components in the system. A failure of one or several components can cause other components to transition to an idle state.

Structural dependence from a performance point of view exits when the performance of a system depends on the configuration of its units, and when it is not just the sum of the performance of the individual units. Components are dependent through the physical structure of the system. Wu et al. (2016) consider structural dependence from a performance point of view. Systems with an arbitrary structure are considered, and an importance measure is introduced to determine which units should be preventively maintained when a unit fails. Heuristic decision rules are optimised based on simulation.

Keizer et al. (2017) identify resource dependence as a fourth type of dependence that can exist. This type of dependence exists, for example, if limited repair workers are responsible for the maintenance activities of various units or systems, if a single, finite stock of spare parts is used for the replacement of multiple units, or if time windows during which maintenance can be carried out have limited lengths. Rasmekomen and Parlikad (2013) use simulation-based optimisation to analyse a system consisting of a number of parallel machines. A single maintenance crew can only maintain one machine at a time. Liu et al. (2014) propose a value-based preventive maintenance policy for multi-component systems with continuously degrading components. Only one maintenance crew is available and only one component can be maintained at a time. A yield-cost importance measure is proposed to determine which component should be maintained when the system reliability reaches a reliability threshold. Diallo et al. (2018) develop a two-stage heuristic algorithm to select components for maintenance, and to determine the degrees of repair for serial k-out-of-n systems that operate for consecutive missions interspersed with finite breaks during which only a selected set of component repairs or replacements can be carried out due to limited time.

As a result of these dependencies, maintenance optimisation for multi-component systems is a challenging problem. It combines the dependent failure processes of components, the resource dependence and the structural dependence with the combinatorial optimisation problem regarding the grouping of maintenance activities.

The research in Part II is motivated by gaps in the maintenance literature. There are only a small number of papers in the literature that consider stochastic dependence (Keizer et al., 2017; de Jonge and Scarf, 2019). The majority of studies on multi-component systems consider a single type of dependence, implying that ample research opportunities exist that incorporate multiple dependencies (de Jonge and Scarf, 2019). Moreover, as noted by de Jonge and Scarf (2019), only a limited number of studies take parameter uncertainty into account. We contribute to all of these areas in Part II of this thesis. Furthermore, we highlight the benefits of sequential maintenance decisions over one-step ahead decisions.

It is common in the maintenance literature to make maintenance decisions (scheduling maintenance times and deciding which components to repair or replace at scheduled times) by minimising the cost per unit time. In Part II of this thesis we propose a loss-based utility (or reward) function, Λ, for multi-component systems with economic dependence (through a fixed set-up cost) and stochastic dependence (through failure-based load sharing). The utility, Λ, is a combination of interpretable penalties that encapsulate the costs of economic and stochastic dependence.

There are potential benefits to using a loss-based utility function in practice. First, writing the loss of each type of dependence separately shows the cost of each type of dependence. It could be beneficial, in terms of design or maintenance planning, to understand which dependence is the most expensive. For example, if we find that the cost of load-sharing is magnitudes larger than other costs, it may prompt us to add more redundancy to a system or to a particular subsystem; or if the cost of resource dependence is magnitudes larger than other costs, it may allow us to schedule more workers or to buy more spare parts in advance. The above insights may not be as clearly drawn when using the cost-per-unit-time utility. Separating the costs of each form of dependence may be of value to the maintenance literature, especially for large complex systems with multiple dependencies. Second, we will see that the rewards and losses of a sequence of maintenance actions are more easily defined and calculated using a loss-based utility function compared to the cost-per-unity time utility.

In Part II, we define penalties for systems with economic dependence (through a fixed set-up cost) and failure-based load sharing dependence (a form of stochastic dependence). The ideas proposed in Part II could be used to extended loss-based utility functions to systems with resource and structural dependence. In addition, the losses due to each type of dependence can be tailored to application. Through simulation studies we will compare Λ to the cost per unit time utility. We will compare a random-threshold approach to a fixed-threshold approach and an expected failure time approach to highlight the importance of incorporating all uncertainty when making maintenance decisions.

The structure of Part II is as follows. Chapter 6 provides the reliability and maintenance concepts required for the subsequent chapters. Chapter 6 introduces system reliability for multi-component systems and defines economic and stochastic dependence in the context of maintenance. Chapter 6 also introduces degradation processes and degradation thresholds. In Chapter 7, we illustrate the penalties incurred by multi-component systems as a result of economic dependence, through a fixed set-up cost, and stochastic dependence, through failure-based load sharing. We then propose a novel CBM policy that incorporates a loss-based utility function, which is a combination of interpretable penalties that encapsulate the costs of economic and stochastic dependence, in a sequential Bayesian decision framework. In Chapter 8, we implement the CBM policy proposed in Chapter 7. The CBM policy is compared to alternative policies, using simulation studies, to highlight the advantages of our sequential CBM policy. Appendix A.2 provides supplementary material for Part II.

Part I

Chapter 2

Bayesian Reliability Analysis

In this chapter we introduce concepts of reliability. We begin by defining reliability and provide reasons to collect reliability data. Next, we introduce the probability density function, the cumulative distribution function, the reliability or survival function, and the hazard function. We then provide examples of reliability data: including pass/fail, failure count, failure age, and degradation data. We describe censoring and truncation in the context of reliability. We present the likelihood contributions for censored observations and the likelihood adjustments for truncated observations. We conclude the chapter by introducing Bayesian inference and Bayesian computation.

2.1 Introduction

2.1.1 Quality and reliability

Rapid advances in technology, development of highly sophisticated products, intense global competition, and increasing customer expectations have put new pressures on manufacturers to produce high-quality products (Meeker et al., 2022). Customers expect products to be reliable and safe. Systems, vehicles, machines, devices, and so on should, with high probability, perform their intended function under usual operating conditions, over an extended period of time (Meeker et al., 2022).

The International Organization for Standardization (ISO) defines reliability as "the ability of an item to perform a required function, under given environmental and operating conditions and for a stated period of time" (Hamada et al., 2008; Meeker et al., 2022). Assessing, or improving, the reliability of products requires methods for predicting and assessing various aspects of product reliability. In most cases this will involve the collection of reliability data from studies, such as laboratory tests or designed experiments, of materials, devices, items and components, tests on early prototype units, monitoring of units in the field, or analysis of warranty data (Meeker et al., 2022).

2.1.2 Reasons for collecting reliability data

There are many reasons for collecting reliability data. Examples include:

1. Assessing characteristics of items over a warranty period or the product's lifetime.
2. Predicting product reliability.
3. Predicting product warranty costs.
4. To aid with the design of a system or to assess the effect of a proposed design change.
5. Tracking the product in real-time to provide information on causes of failure and methods of improving product reliability.
6. Tracking the product in real-time to inform maintenance decisions.
7. Tracking the product in real-time to assess how different environmental and operating conditions impact product lifetime.

In Part I of this thesis we will be concerned with predicting hard drive reliability and assessing the impact of critical attributes on the failure times of drives, that is, bullet points 2 and 7 above.

2.2 Failure-time distribution functions

Much of reliability analysis focuses on modelling the failure time distribution of an item; which involves a random variable, say T. For example, if we make use of a hard drive to store data until the hard drive fails, the age of the drive at failure, T, is a random variable, taking positive real values.

We can specify the properties of a random variable using different representations, all of which contain equivalent information. Each representation is useful in specific contexts. These representations include the probability density function (PDF), the reliability or survival function, the cumulative distribution function (CDF), and the hazard function.

For a positive continuous random variable, T, taking values on the positive real line, the PDF is a function, $f(t)$, that satisfies

$$f(t) \geq 0,\ 0 \leq t < \infty, \tag{2.2.1}$$

and

$$\int_0^\infty f(t)dt = 1. \tag{2.2.2}$$

We can define the survival function of T, also known as the reliability function. The survival function is defined as

$$S(t) = \Pr(T > t) = \int_t^\infty f(x)dx, \tag{2.2.3}$$

where $f(t)$ is the PDF. The survival function takes values in the interval $[0, 1]$.

The CDF of T is defined as

$$F(t) = \Pr(T \leq t) = \int_0^t f(x)dx. \tag{2.2.4}$$

The CDF is the complement of the survival function.

Another way to specify the properties of a random variable is the hazard function, also called the instantaneous failure rate function. Suppose that we are interested in the probability that an item will fail in the time interval $[t, t + \Delta t]$, given that the item is working at time t; from probability theory, we can write this as

$$\Pr(t < T \leq t + \Delta t \mid T > t) = \frac{\Pr(t < T \leq t + \Delta t)}{\Pr(T > t)} = \frac{F(t + \Delta t) - F(t)}{S(t)}. \tag{2.2.5}$$

If we want to know the instantaneous failure rate, we divide by the length of the interval, Δt, and let $\Delta t \to 0$. This gives

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T \leq t + \Delta t \mid T > t)}{\Delta t}. \tag{2.2.6}$$

We call $\lambda(t)$ the hazard function. The hazard function can be thought of as an item's propensity to fail in the next short interval of time, given that the item has survived to time t.

For positive random variables, Table 2.1 summarises the mathematical relationships between the PDF, $f(t)$, the CDF, $F(t)$, the survival function, $S(t)$, and the hazard function, $\lambda(t)$.

	$f(t)$	$F(t)$	$S(t)$	$\lambda(t)$
$f(t)$	$f(t)$	$\frac{d}{dt}F(t)$	$-\frac{d}{dt}S(t)$	$\lambda(t)\exp\left[-\int_0^t \lambda(s)ds\right]$
$F(t)$	$\int_0^t f(s)ds$	$F(t)$	$1-S(t)$	$1-\exp\left[-\int_0^t \lambda(s)ds\right]$
$S(t)$	$\int_t^\infty f(s)ds$	$1-F(t)$	$S(t)$	$\exp\left[-\int_0^t \lambda(s)ds\right]$
$\lambda(t)$	$f(t)/\int_t^\infty f(s)ds$	$\frac{d}{dt}F(t)/[1-F(t)]$	$-\frac{d}{dt}\log S(t)$	$\lambda(t)$

Table 2.1: Relationships between the PDF, $f(t)$, the CDF, $F(t)$, the survival function, $S(t)$, and the hazard function, $\lambda(t)$, assuming $f(t) = 0$ for $t < 0$.

2.3 Examples of reliability data

This section describes examples, and data sets, that illustrate the wide range of applications and characteristics of reliability data. These include pass/fail, failure count, failure age, and degradation data.

2.3.1 Bernoulli success/failure data

The simplest form of reliability data is "pass/fail" or Bernoulli trial data. This data arises from simple "pass/fail" testing. Table 2.2 contains outcomes from a set of Bernoulli trials. These data are the launch outcomes of new aerospace vehicles conducted by companies during the period between 1980 and 2002. A total of 11 launches occurred; 3 were successes and 8 were failures. Reliability in this case is the probability of a successful launch. The data is provided by Johnson et al. (2005). The data is also discussed in Hamada et al. (2008).

Vehicle	Outcome
Pegasus	Success
Percheron	Failure
AMROC	Failure
Conestoga	Failure
Ariane 1	Success
India SLV-3	Failure
India ASLV	Failure
India PSLV	Failure
Shavit	Success
Taepodong	Failure
Brazil VLS	Failure

Table 2.2: New launch vehicle outcomes (Johnson et al., 2005).

2.3.2 Failure count data

Bernoulli data can also be recorded as a function of time. Failure count data represent the number of failures that occur over a period of time. For example, Table 2.3 provides data on the number of pump failures, x_i, observed in t_i thousand operating hours for 10 different systems at the Farley 1 United States commercial nuclear power plant. The random variable is the number of pump failures, X_i, and the reliability is the probability that no pump failures occur in a given period of time. The data is provided by Gaver and O'Muircheartaigh (1987). The data is also discussed in Hamada et al. (2008).

System	x_i (failures)	t_i (thousand hours)
1	5	94.320
2	1	15.720
3	5	62.880
4	14	125.760
5	3	5.240
6	19	31.440
7	1	1.048
8	1	1.048
9	4	2.096
10	22	10.480

Table 2.3: Pump failure count data from the Farley 1 United States commercial nuclear power plant (number of pump failures, x_i, observed in t_i thousand operating hours) (Gaver and O'Muircheartaigh, 1987).

2.3.3 Lifetime or failure age data

Hard drive dataset

Table 2.4 presents an example of failure age data for hard drives from data backup company Backblaze. Backblaze is a company that offers cloud backup storage to protect against data loss. Since 2013, it has been collecting daily operational data on all of the hard drives operating at its facility. Some drives have been running since 2013 or before, while others were added at a later date. In other words, some drives have a history prior to data collection. Drives that failed prior to the start of data collection are not included in the dataset. Every quarter Backblaze makes its hard drive data publicly available through its website (Backblaze, 2022a).

Figure 2.1 provides a schematic representation of two possible life histories of the Backblaze hard drives and an example hard drive not included in the dataset. Drive 1 was in a failed

Drive	Start age (operating hours)	End age (operating hours)	Failure indicator
1	9369	15211	1
2	9314	21692	1
3	9524	21977	1
4	8853	17850	1
5	9541	10110	0
6	9647	21749	0
7	9	11550	0
8	9587	21530	0
9	9524	21978	0
10	9743	22671	0

Table 2.4: Failure age data for hard drives from data backup company Backblaze (Backblaze, 2022a).

state prior to the start of the study and was not included in the dataset. Drive 2 is right-censored (still in a working state) at the end of the study. Drive 3 failed prior to the end of the study. We define the failure age of a hard drive to be the age of the drive at failure. For example, the failure age for drive 3 is Y_3. We define calendar time to be the time since the study commenced. Under this terminology, the failure time for drive 3 is $\tilde{Y}_3$.

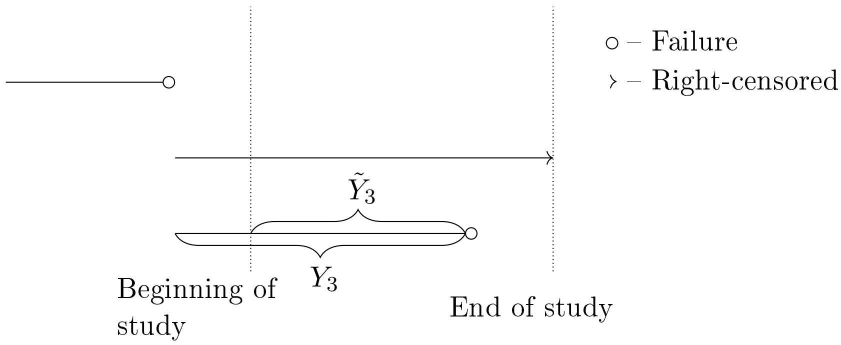

Figure 2.1: Two possible life histories and an example drive not included in the dataset.

Table 2.4 presents failure age data for ten example hard drives. Table 2.4 provides failure ages, for drives that are observed to fail before the end of the study, and right-censoring ages, for drives that are still in a working state at the end of the study. In addition, the dataset provides the age at the start of the study, and an indicator for each hard drive to indicate if the drive has failed (1) or not (0). From Table 2.4 we can see that the first four drives failed after working for 15211, 21692, 21977, and 17850 hours, respectively; and the final six drives were still in a working state after working for 10110, 21749, 11550, 21530, 21978, and 22671 hours, respectively. The random variable is the age of the drive at failure, and the reliability is the probability that the drive has not failed by a given age.

Warranty data

Drive	End age (operating hours)	Failure indicator
1	1521	1
2	2162	1
3	2077	1
4	1750	1
5	500	1
6	749	1
7	510	1
8	1530	1
9	978	1
10	671	1

Table 2.5: Warranty data with limited information for units that did not fail.

Table 2.5 illustrates an example warranty dataset. A particular item, after purchase, may be used regularly, sporadically, or not at all. The percentage of units put into regular use is unknown. During a particular production period, an incorrect component was installed in all of the units that were produced. When failures occur among the units, the units are returned to the manufacturer for repair or replacement under a long-term warranty program. The manufacturer learns about failures from this group of units only if the unit is put into service and if the unit fails before the analysis time. If a customer does not use the unit until failure (prior to the analysis time), the data will not be available to the manufacturer.

Figure 2.2 provides a schematic representation of three possible life histories and an example unit not included in the dataset. The first three units failed prior to the analysis time and were included in the dataset. The final item was not returned prior to analysis and hence the information about this item was not included in the dataset. The random variable is the age of the unit at failure, and the reliability is the probability that the unit has not failed by a given age.

2.3.4 Degradation data

In some applications it is useful to measure the degradation of an item rather than, or in addition to, its failure age. Degradation modelling is split into two categories: soft failure degradation modelling, where items fail when their degradation level reaches a fixed threshold; and hard failure degradation modelling, where items fail when their degradation level reaches a random threshold. The latter is used when degradation data and failure age data are both available. The former is used when failure age data is not available.

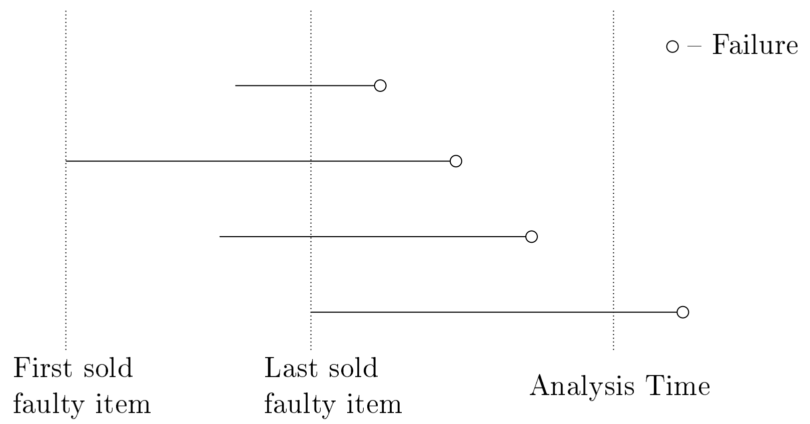

Figure 2.2: Three possible life histories and an example unit not included in the dataset.

2.3.4.1 Soft failures: fixed degradation level

For some items there is a gradual loss of performance over time. For example, decreasing light output from a fluorescent light bulb. We may define the light bulb to have failed when the light output is $\alpha\%$ of its initial output. This is an example of a soft failure. The light bulb is still in a working condition, but is not functioning at the required performance level. In this example, the degradation, $D(t)$, is the light output, as a percentage of the initial output, at age t.

A fixed value of $\mathcal{D}_f$ is used to denote the critical level for the degradation path. The failure age, T, is defined as the age when the degradation path crosses the fixed degradation level $\mathcal{D}_f$.

Table 2.6 shows degradation measurements of four organic coating specimens. The data shown in Table 2.6 is part of an outdoor weathering dataset from a study conducted by scientists at the National Institute of Standards and Technology (NIST). The data were collected in a study of the service life of organic coatings in outdoor environments. The degradation measurement of the organic coatings was recorded periodically (at intervals of several days) for each specimen using Fourier transform infrared spectroscopy (FTIR).

Generally, the degradation failure threshold is chosen to be the level of degradation at which the performance of the coating would not be acceptable (e.g., the level at which there would be customer-perceivable loss of gloss or colour). Hong et al. (2015) study the outdoor weathering dataset and choose the failure threshold to be $\mathcal{D}_f = -0.4$.

Figure 2.3 illustrates the degradation paths for the four specimens in Table 2.6. Using the failure threshold from Hong et al. (2015) we can see that specimens one, two, three, and four are considered failed after 72.08, 69.68, 70.52, and 70.19 days, respectively. The random variable is the age of the coating at failure, and the reliability is the probability that the coating has not failed by a given age.

Specimen Number	Time (days)					
	0	3	7	10	...	84
1	-0.008	-0.026	-0.044	-0.046	...	-0.431
2	-0.011	-0.032	-0.050	-0.052	...	-0.454
3	-0.002	-0.014	-0.034	-0.051	...	-0.459
4	-0.001	-0.014	-0.035	-0.054	...	-0.475

Table 2.6: Degradation data of organic coatings in outdoor environments obtained using Fourier transform infrared spectroscopy (FTIR) (Hong et al., 2015).

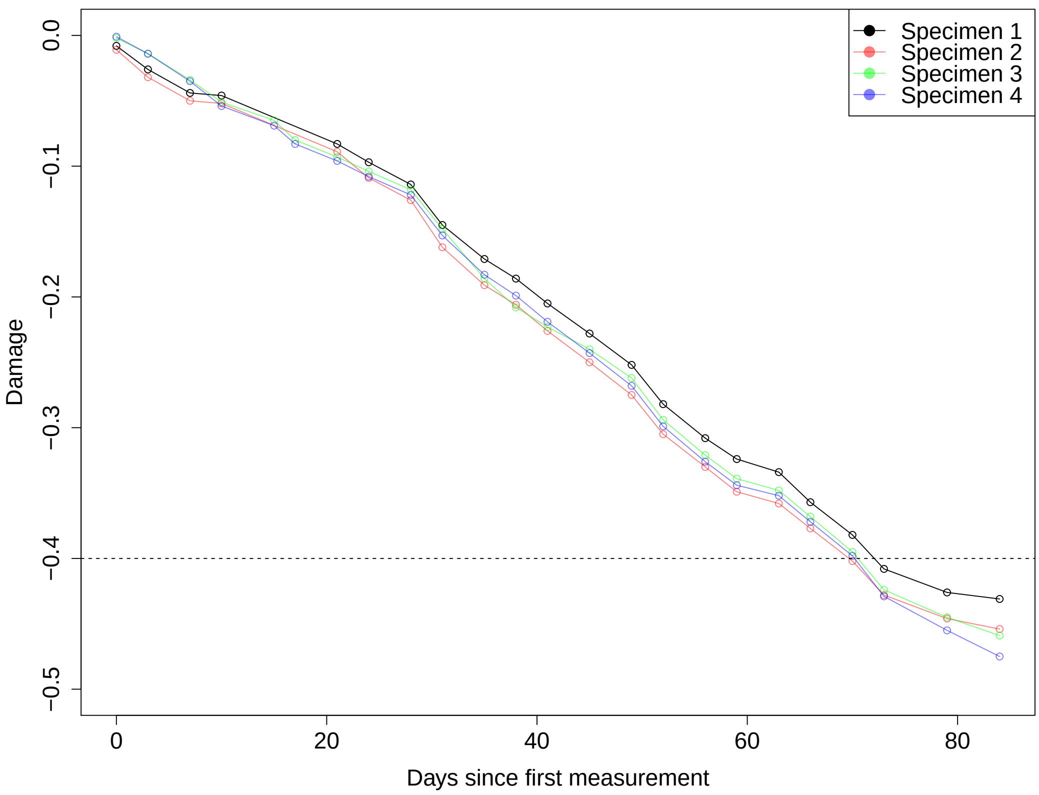

Figure 2.3: Plot of four representative degradation paths of organic coatings in outdoor environments.

2.3.4.2 Hard failures: random degradation level

For some products, the definition of failure is clear: the product stops working. These are called "hard failures". With hard failures, failure ages will not, in general, correspond with a fixed level of degradation. Instead, the level of degradation at which failure occurs will be random.

Hong and Meeker (2013b) model the failure ages of Product D2. Product D2 is similar to a high-end copying machine connected to the Internet and installed with a smart chip to record the number of pages that have been printed, as a function of time. Hong and Meeker (2013b) model the failure ages of Product D2 using the use-rate (cycles per week) as a dynamic covariate. In addition to the dynamic use-rate data, failure ages for failed units, and right-censored ages for units that did not fail, are also available.

Hong and Meeker (2013b) model the failure ages of Product D2 using the cumulative exposure or cumulative damage model. Given the entire covariate history, the cumulative exposure, $u(t)$, by age t is defined as

$$u(t) = u(t, \beta, \boldsymbol{x}(t)) = \int_0^t \exp[\beta x(s)]ds, \tag{2.3.1}$$

where β is a parameter that controls the rate at which exposure accrues as a function of the covariate values, $\boldsymbol{x}(t) = \{x(s) : 0 < s \leq t\}$ is the covariate history up to age t, and $x(s)$ is the covariate value at age s. Each unit accumulates an unobservable amount of the cumulative exposure, that depends on the dynamic covariate, $\boldsymbol{x}(t)$. The unit fails at age T when the amount of cumulative exposure reaches a random threshold U. That is, $U = u(T)$ and T is defined as the failure age of the unit. Thus, the relationship between cumulative exposure U and failure age T is

$$U = u(T) = \int_0^T \exp[\beta x(s)]ds. \tag{2.3.2}$$

The cumulative exposure threshold U has a CDF, $F(u; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of model parameters. Figure 2.4 illustrates the cumulative exposure model for three example covariate processes (i.e., three example use-rate processes). Figure 2.4 (left) depicts the cumulative exposure or the cumulative damage for three example covariate processes. The cumulative exposure can be thought of as a degradation path. The degradation path of a unit increases until the degradation reaches a random threshold U. Figure 2.4 (right) depicts the cumulative exposure CDF. From Figure 2.4 (right) we can see that the probability of failure increases as the cumulative exposure, $u(t)$, increases. Each level of exposure (or each level of degradation) has an associated probability of failure. This is in

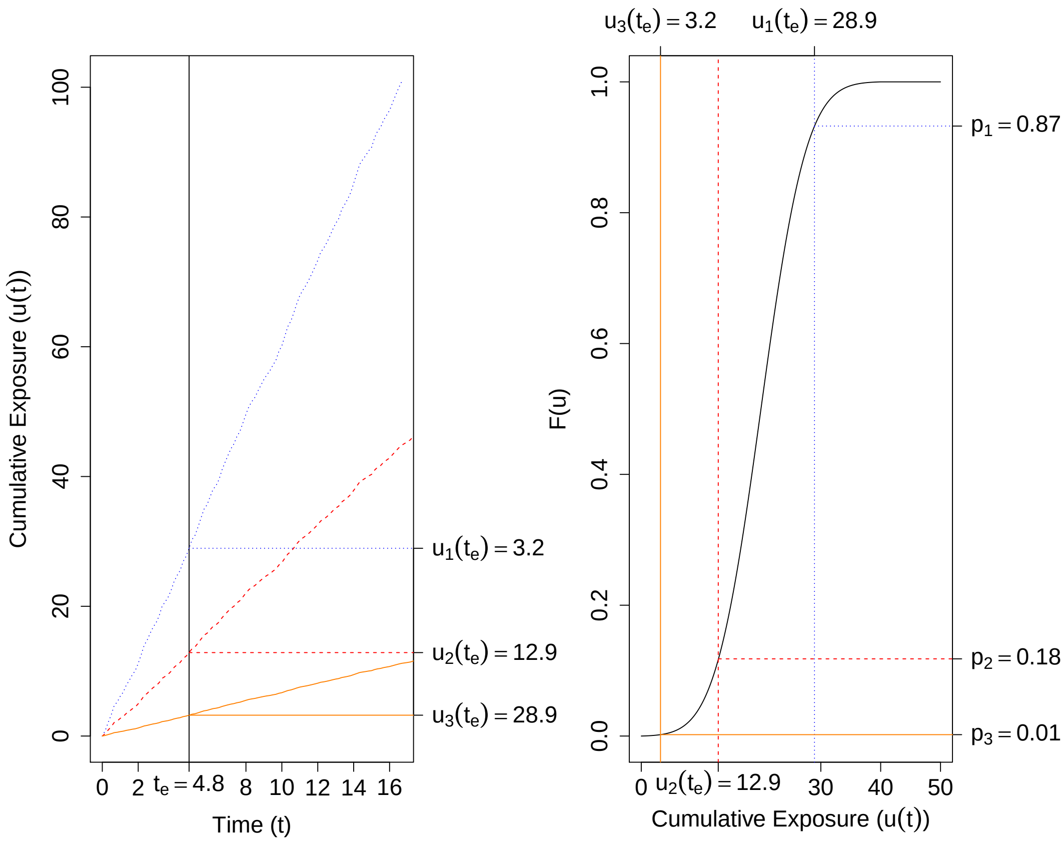

Figure 2.4: Illustration of the cumulative exposure model for three example covariate processes.

contrast to soft failures which use a fixed threshold, all units are assumed to fail when the degradation path reaches a fixed threshold, $\mathcal{D}_f$. We see from Figure 2.4, that the degradation level at failure varies from unit to unit.

2.4 Censoring

One common feature of reliability data is the presence of censoring. Failure age observations are censored when the exact failure age for a specific item is unknown. There are several types of censoring, including left, right, and interval censoring.

Left-censoring occurs when an item fails before the first inspection. For example, suppose than an experiment tests the failure age of a new hard drive. A set of 100 drives are inspected at 5 p.m. every day for two years to determine if the drive has failed. Suppose the drives are put in service at 5pm on the first day of the study. The failure age for any drive that fails before 5 p.m. on the second day is left-censored. Right-censoring occurs when an item has not failed by the last inspection. Consider our drive example. The failure age for any drive that has not failed after two years is right-censored.

Both left-censoring and right-censoring are special cases of interval-censoring. Interval-censoring occurs when an item's failure age is only known to be in an interval, (t_i, t_{i+1}). If an observation is left-censored at t, its failure age is in $(0, t)$. If an observation is right-censored at t, its failure age is in (t, ∞). In our drive example, the failure ages are interval-censored because they can only be determined to fail within a 24-hour interval.

We usually assume that the censoring and failure ages are independent. This is known as independent-censoring or noninformative-censoring. This implies that censoring ages of units provide no information about the failure age distribution. Using future events (or indicators of future events) to stop observing a unit could introduce bias.

Hard drives record many types of error over their lifespan. Modern drives can exhibit latent sector errors, reallocated sector counts, and many other read/write errors. Such errors weaken drives, but drives can still function as intended with such errors. The noninformative-censoring assumption would be violated, for example, if hard drives were removed from the study before actual failure, but in response to the drive recording drive errors. This type of censoring is informative, because the drive was removed after an event that is expected to increase the failure rate of the drive. Removing these drives would introduce bias to a standard reliability analysis.

2.5 Truncation

Another common feature of reliability data is the presence of truncation. It is important to distinguish between truncated observations and censored observations. Censoring occurs when there is a bound on an observation. Truncation, however, arises when even the existence of a potential observation would be unknown if its value were to lie in a certain range. Usually truncation occurs to the left of a specified point, τ^L, referred to as left-truncation, or to the right of a specified point, τ^R, referred to as right-truncation.

The hard drive dataset discussed in Section 2.3.3 is an example of left-truncated data. Drives that failed prior to the start of data collection are not included in the dataset. In other words, drives that failed before calendar time $\tau = 0$ are not observed. Consequently, the remaining observations are from a left-truncated distribution. Not incorporating the information provided by the left-truncated drives can introduce bias.

The warranty dataset discussed in Section 2.3.3 is an example of right-truncated data. Units that failed after the analysis time, τ^R, are not included in the dataset. In other words, units that failed after the analysis time are not observed. Consequently, the remaining observations are from a right-truncated distribution. Not incorporating the information provided by the right-truncated units can introduce bias.

Warranty data revisited

Suppose that all faulty units were sold to one buyer. For example, a data backup company buying hard drives in bulk. Moreover, suppose the manufacturer alerted the buyer of the faulty drives and hence the full batch of units were returned. In this scenario we observe the data presented in Table 2.7.

Drive	End age (operating hours)	Failure indicator
1	1521	1
2	2162	1
3	2077	1
4	1750	1
5	500	1
6	749	1
7	510	1
8	1530	1
9	978	1
10	671	1
11	500	0
12	1500	0
13	2100	0
14	1207	0
15	1785	0

Table 2.7: Warranty data with information for units that did not fail.

More specifically, we observe data on the failed units and the units that did not fail prior to analysis. This dataset is an example of right-censored data, since we have information about the units that did not fail. More specifically, we know the number of units that did not fail and we know a lower bound for the failure age for each unit that did not fail. In the warranty dataset discussed in Section 2.3.3, the units that did not fail prior to analysis were unknown.

2.6 Likelihood

Truncation and censoring are common in large populations of units. For example, left-truncation is common when units are required to be in a working state at the study start time. In this setting, sampling is biased towards longer unit failure ages since units are only included in the study if they are in a working state on entry. The analysis of truncated reliability data should apply statistical methods that account for this bias. Moreover, right-censoring is common in reliability studies when, due to advancements in technology, units are retired and replaced by newer technology, or because units simply will not have failed by the end of the study. Not accounting for censoring will also produce biased estimates. These biases can be accounted for through the likelihood function, for parametric methods for inferring the failure ages of units.

The full likelihood can be written as the joint probability of the observations as a function of the parameters of the chosen statistical model. Assuming n independent observations, the sample likelihood is

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}; \boldsymbol{X}) = \prod_{i=1}^{n} L_i(\boldsymbol{\theta}; x_i), \tag{2.6.1}$$

where $L_i(\boldsymbol{\theta}; x_i)$ is the likelihood contribution from observation i, x_i is observation i, $\boldsymbol{X} = (x_1, \ldots, x_n)$ is the vector of observations for all units, and $\boldsymbol{\theta}$ is the vector of model parameters to be estimated.

In this section, we provide the likelihood contributions for censored observations and the likelihood adjustments for truncated observations.

2.6.1 Likelihood contributions for censored observations

Figure 2.5 illustrates the intervals where the observations may lie for left-censored, interval-censored, and right-censored observations. The likelihood contribution for each of these cases, shown in Table 2.8, is the probability of failing in the corresponding interval.

Type of observation	Failure age	Contribution
Interval-censored	$t_{i-1} < T < t_i$	$F(t_i) - F(t_{i-1})$
Left-censored	$T \leq t_i$	$F(t_i)$
Right-censored	$T > t_i$	$1 - F(t_i)$
Uncensored	$T = t_i$	$f(t_i)$

Table 2.8: Likelihood contributions for an interval-censored, left-censored, right-censored, and uncensored observation.

Interval-censored observations

If a unit's failure age is known to have occurred between ages t_{i-1} and t_i, the likelihood contribution of the observation is

$$L_i(\boldsymbol{\theta}) = \int_{t_{i-1}}^{t_i} f(t)dt = F(t_i) - F(t_{i-1}). \tag{2.6.2}$$

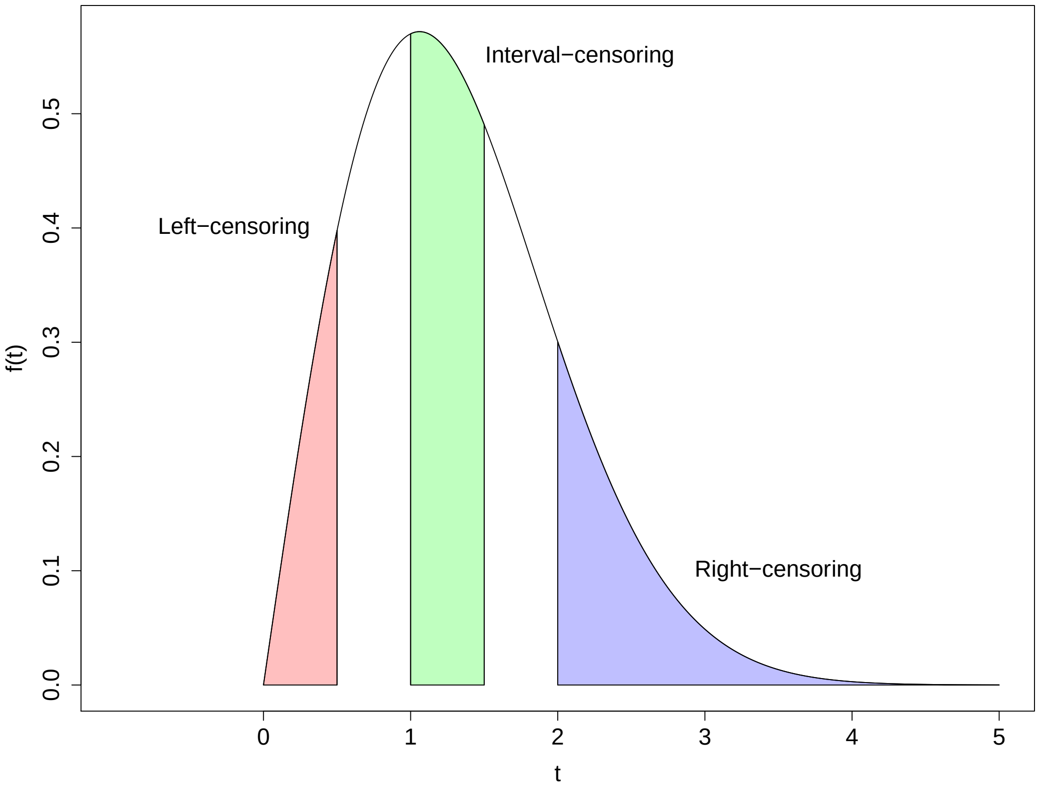

Figure 2.5: Likelihood contributions for different kinds of censoring.

Most data arising from observation of a time/age process can be thought of as having occurred in intervals similar to (t_{i-1}, t_i). Most modern products collect data using sensors or monitoring systems. Data may be collected every second, for example, and the failure ages of these products are interval-censored between (t_{i-1}, t_i) seconds, where unit i is observed to have failed by age t_i and was in a working state at age t_{i-1}. It is common to ignore interval-censoring if the failure age of a product is relatively much larger than the time between inspections. The following special cases warrant separate consideration.

Left-censored observations

Left-censored observations occur in life test applications when a unit has failed by the first inspection; all that is known is that the unit failed before the first inspection. If there is an upper bound t_i for the ith failure age, causing it to be left-censored, the likelihood contribution of the observation is

$$L_i(\boldsymbol{\theta}) = \int_0^{t_i} f(t)dt = F(t_i) - F(0) = F(t_i). \tag{2.6.3}$$

Right-censored observations

Right-censoring is common in reliability data analysis. It is uncommon for all units to have failed during a reliability study. Right-censored observations occur in life test applications when a unit has not failed by the final inspection; all that is known is that the unit will fail after the final inspection. If there is a lower bound t_i for the ith failure age, the failure age is somewhere in the interval (t_i, ∞). The likelihood contribution of the observation is

$$L_i(\boldsymbol{\theta}) = \int_{t_i}^{\infty} f(t)dt = F(\infty) - F(t_i) = 1 - F(t_i). \tag{2.6.4}$$

Uncensored observations

A failure age is considered uncensored if the failure is observed "exactly" at age t_i. The term exact is used loosely, since sensors and monitoring systems can only record data every t time units. Even if t is infinitesimally small, the failure ages will be interval-censored. However, as mentioned above, it is common to ignore interval-censoring if the failure age of a product is relatively much larger than the time between inspections. The likelihood contribution of an uncensored observation is

$$L_i(\boldsymbol{\theta}) = f(t_i). \tag{2.6.5}$$

2.6.2 Likelihood adjustments for truncated observations

In this section we provide the likelihood adjustments for left-truncated observations, right-truncated observations, and observations that are both left-truncated and right-truncated. We assume that we do not have any knowledge about the truncation distribution.

Likelihood with left-truncation

If a random variable T_i is truncated on the left at τ_i^L, then the likelihood of an observation in the interval $(t_{i-1}, t_i]$ is the conditional probability

$$L_i(\boldsymbol{\theta}) = \Pr(t_{i-1} < T_i \leq t_i \mid T_i > \tau_i^L) = \frac{F(t_i; \boldsymbol{\theta}) - F(t_{i-1}; \boldsymbol{\theta})}{1 - F(\tau_i^L; \boldsymbol{\theta})}, \ \tau_i^L < t_{i-1} < t_i. \quad (2.6.6)$$

For an observation reported as an exact failure at age t_i, the corresponding likelihood contribution is

$$L_i(\boldsymbol{\theta}) = \frac{f(t_i; \boldsymbol{\theta})}{1 - F(\tau_i^L; \boldsymbol{\theta})}, \ \tau_i^L < t_i. \quad (2.6.7)$$

It is possible to have censored observations when sampling from a left-truncated distribution. The recorded censored age(s) will exceed τ_i^L. To obtain $L_i(\boldsymbol{\theta})$ for a censored observation, we replace the numerator in Equation (2.6.7) by $F(t_i; \boldsymbol{\theta}) - F(\tau_i^L; \boldsymbol{\theta})$ for an observation that is left-censored at $t_i > \tau_i^L$ and by $1 - F(t_i; \boldsymbol{\theta})$ for an observation that is right-censored at $t_i > \tau_i^L$. Equation (2.6.6) provides the likelihood contribution for an observation that is interval-censored in (t_{i-1}, t_i).

Likelihood with right-truncation

If a random variable T_i is truncated on the right at τ_i^R, then the likelihood of an observation in the interval $(t_{i-1}, t_i]$ is the conditional probability

$$L_i(\boldsymbol{\theta}) = \Pr(t_{i-1} < T_i \leq t_i \mid T_i < \tau_i^R) = \frac{F(t_i; \boldsymbol{\theta}) - F(t_{i-1}; \boldsymbol{\theta})}{F(\tau_i^R; \boldsymbol{\theta})}, \ 0 < t_{i-1} < t_i < \tau_i^R. \quad (2.6.8)$$

For an observation reported as an exact failure at age t_i, the corresponding likelihood contribution is

$$L_i(\boldsymbol{\theta}) = \frac{f(t_i;\boldsymbol{\theta})}{F(\tau_i^R;\boldsymbol{\theta})},\ t_i < \tau_i^R. \tag{2.6.9}$$

As with left-truncation, it is possible to have censored observations when sampling from a right-truncated distribution. The recorded censored age(s) will not exceed τ_i^R. To obtain $L_i(\boldsymbol{\theta})$ for a censored observation, we replace the numerator in Equation (2.6.9) by $F(t_i;\boldsymbol{\theta})$ for an observation that is left-censored at $0 < t_i < \tau_i^R$ and by $F(\tau_i^R;\boldsymbol{\theta}) - F(t_i;\boldsymbol{\theta})$ for an observation that is right-censored at $t_i < \tau_i^R$. Equation (2.6.8) provides the likelihood contribution for an observation that is interval-censored in (t_{i-1}, t_i).

Likelihood with left-truncation and right-truncation

If a random variable T_i is truncated on the left at τ_i^L and on the right at τ_i^R, then the likelihood of an observation in the interval $(t_{i-1}, t_i]$ is the conditional probability

$$L_i(\boldsymbol{\theta}) = \Pr(t_{i-1} < T_i \leq t_i \mid \tau_i^L < T_i < \tau_i^R) = \frac{F(t_i;\boldsymbol{\theta}) - F(t_{i-1};\boldsymbol{\theta})}{F(\tau_i^R;\boldsymbol{\theta}) - F(\tau_i^L;\boldsymbol{\theta})},\ \tau_i^L < t_{i-1} < t_i < \tau_i^R. \tag{2.6.10}$$

For an observation reported as an exact failure at age t_i, the corresponding likelihood contribution is

$$L_i(\boldsymbol{\theta}) = \frac{f(t_i;\boldsymbol{\theta})}{F(\tau_i^R;\boldsymbol{\theta}) - F(\tau_i^L;\boldsymbol{\theta})},\ \tau_i^L < t_i < \tau_i^R. \tag{2.6.11}$$

As with left-truncation or right-truncation, it is possible to have censored observations when sampling from a distribution that is left-truncated and right-truncated. The recorded censored age(s) exceed τ_i^L and do not exceed τ_i^R. To obtain $L_i(\boldsymbol{\theta})$ for a censored observation, we replace the numerator in Equation (2.6.11) by $F(t_i;\boldsymbol{\theta}) - F(\tau_i^L;\boldsymbol{\theta})$ for an observation that is left-censored at $\tau_i^L < t_i < \tau_i^R$ and by $F(\tau_i^R;\boldsymbol{\theta}) - F(t_i;\boldsymbol{\theta})$ for an observation that is right-censored at $t_i < \tau_i^R$. Equation (2.6.10) provides the likelihood contribution for an observation that is interval-censored in (t_{i-1}, t_i).

2.7 Introduction to Bayesian inference and Bayesian computation

Statistical inference is the process of learning from data. The idea behind Bayesian inference is that we express our uncertainty about the values of unknown quantities by giving these quantities probability distributions. The probability distributions express our degree of belief in the quantity taking each possible subset of values.

Beliefs about quantities before we see the data are called prior beliefs. These are represented by prior distributions. When we observe data we use the information in the data to update our prior beliefs to posterior beliefs. These are represented by posterior distributions. Prior beliefs are updated to posterior beliefs using Bayes' theorem.

Suppose there is an unknown parameter θ and that our prior beliefs about θ are represented by a probability distribution with PDF $\pi(\theta)$. Suppose that we will observe a continuous quantity $\boldsymbol{Y}$ and that if we knew θ we could summarise our beliefs about $\boldsymbol{Y}$ through a conditional distribution of $\boldsymbol{Y}$ given θ with PDF $p(\boldsymbol{y} \mid \theta)$. Now suppose we observe $\boldsymbol{Y} = \boldsymbol{y}$. Then Bayes' theorem allows us to compute the PDF of the posterior distribution for θ through

$$\pi(\theta \mid \boldsymbol{y}) = \frac{\pi(\theta)p(\boldsymbol{y} \mid \theta)}{p(\boldsymbol{y})}, \tag{2.7.1}$$

where

$$p(\boldsymbol{y}) = \int_{\Theta} \pi(\theta)p(\boldsymbol{y} \mid \theta)d\theta, \tag{2.7.2}$$

is a normalising constant which ensures that $\int_{\Theta} \pi(\theta \mid \boldsymbol{y})d\theta = 1$, and Θ is the parameter space for θ. The posterior distribution of θ is the conditional distribution for θ given the observed data, $\boldsymbol{y}$. We call $p(\boldsymbol{y} \mid \theta)$, treated as a function of θ, the likelihood function. It is sometimes denoted $L(\theta; \boldsymbol{y})$ and contains the information from the data about θ. The normalising constant, $p(\boldsymbol{y})$, is sometimes called the marginal likelihood or the evidence.

In general, evaluation of the normalising constant, $p(\boldsymbol{y})$, is analytically intractable. In addition, rather than representing a univariate unknown quantity, θ is often a vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)^\top$ containing many unknowns, making numerical approximation of the integral computationally demanding. This was a major obstacle to Bayesian inference in complicated models until the 1990s when Markov chain Monte Carlo (MCMC) methods were adopted into widespread use. There are several software programs available that implement MCMC algorithms, such as BUGS (Lunn et al., 2000), JAGS (Plummer et al., 2003), and Stan (Carpenter et al., 2017). In this thesis we use Stan, which implements the Hamiltonian Monte Carlo algorithm, to evaluate posterior distributions for complex statistical models. Detailed guides on implementing statistical models in Stan are provided by the Stan development team (Stan, 2022). A conceptual overview of Hamiltonian Monte Carlo is provided by Betancourt (2017).

2.7.1 Posterior predictive simulation

Often the unknown quantities that we are interested in are future or hypothetical values of the data $\boldsymbol{Y}$ conditional on historical data $\tilde{\boldsymbol{y}}$. Averaging over the uncertainty in the other unknowns, $\boldsymbol{\theta}$, the posterior predictive distribution can be written as

$$p(\boldsymbol{y} \mid \tilde{\boldsymbol{y}}) = \int_{\boldsymbol{\Theta}} p(\boldsymbol{y} \mid \tilde{\boldsymbol{y}}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta} \mid \tilde{\boldsymbol{y}}) d\boldsymbol{\theta}. \tag{2.7.3}$$

Often, if we know $\boldsymbol{\theta}$ (typically the model parameters), then knowing about the historical data $\tilde{\boldsymbol{y}}$ provides no further information about the distribution of $\boldsymbol{Y}$. In this scenario, we say that $\boldsymbol{Y}$ and $\tilde{\boldsymbol{Y}}$ are conditionally independent given $\boldsymbol{\theta}$ and Equation (2.7.3) simplifies to

$$p(\boldsymbol{y} \mid \tilde{\boldsymbol{y}}) = \int_{\boldsymbol{\Theta}} p(\boldsymbol{y} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta} \mid \tilde{\boldsymbol{y}}) d\boldsymbol{\theta}. \tag{2.7.4}$$

We can sample from the posterior predictive distribution of $\boldsymbol{Y}$ using samples from the posterior distribution for $\boldsymbol{\theta}$, say $\boldsymbol{\theta}_j$, for $j = 1, \ldots, B$, by setting the parameters, $\boldsymbol{\theta}$, equal to $\boldsymbol{\theta}_j$ and sampling from $p(\boldsymbol{y} \mid \boldsymbol{\theta}_j)$, for $j = 1, \ldots, B$. This yields a sample from the posterior predictive distribution, $\boldsymbol{y}_j$, for $j = 1, \ldots, B$. The posterior predictive sample can be used to generate draws from the posterior predictive distribution of any summary of $\boldsymbol{Y}$ of interest, say $g(\boldsymbol{Y})$, by evaluating the function g at each sampled value to give $g(\boldsymbol{y}_j)$, for $j = 1, \ldots, B$. Gelman et al. (1995) provide a detailed coverage of posterior predictive simulation and other aspects of Bayesian data analysis.

In Part I we are interested in the posterior predictive failure age distribution. This can be written as

$$p(t \mid \boldsymbol{\mathcal{X}}) = \int_{\boldsymbol{\Theta}} p(t \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{X}}) d\boldsymbol{\theta}, \tag{2.7.5}$$

where $\boldsymbol{\mathcal{X}}$ is the failure age data and additional information collected by SMART (critical attributes), $\boldsymbol{\theta}$ is a vector of model parameters, and $\pi(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{X}})$ is the posterior distribution for $\boldsymbol{\theta}$ conditional on the observed data $\boldsymbol{\mathcal{X}}$.

2.8 Conclusions

In this chapter we introduced and described the relationships between the PDF, the CDF, the survival (or reliability) function, and the hazard function; four commonly used functions in the field of reliability. We provided examples of reliability data: including pass/fail, failure count, failure age, and degradation data. We introduced left, right, and interval-censoring and left and right-truncation. We illustrated how censoring and truncation can arise in practice. We showed how censored observations contribute to the likelihood function and described the likelihood adjustments for truncated observations. We concluded the chapter by introducing Bayesian inference and Bayesian computation.

Chapter 3

Multi State Models

In this chapter we extend the standard reliability methodology presented in Chapter 2. We introduce three multi-state models: the two-state model, the illness-death model, and the four-state multi-state model, referred to as the multi-state model. The two-state model is a simple multi-state model with two states and one transition between those states. The two-state model describes standard failure age data, introduced in Chapter 2, where units begin in a working state and eventually transition to a terminal (failed) state. The illness-death model is a multi-state model with three states and describes semi-competing risks data, an extension of standard failure age data, where units begin in a working state and are subject to a nonterminal event and a terminal event. The multi-state model is a multi-state model with four states and describes semi-competing risks data, where units begin in a working state and are subject to two nonterminal events and a terminal event. In this chapter we take the terminal event to be failure. We present the general form of left-truncated and right-censored data and derive the likelihood and the DRLs under each model. In Chapter 5, we model the failure ages and survival probabilities of hard drives using the two-state model, the illness-death model, and the multi-state model. The DRLs obtained under the multi-state model are used to examine the impact of critical attributes on hard drive failure ages and survival probabilities.

3.1 Introduction

A multi-state model is used to model a process where units transition from one state to the next. For instance, a standard survival or reliability analysis can be thought of as a simple multi-state model with two states (working and failed) and one transition between those two states. A diagram illustrating this process is shown in Figure 3.1. In these types of diagrams, each box is a state and each arrow is a possible transition. In Figure 3.1 units start in a working state (state 0) and eventually transition to a failed state (state 1).

Figure 3.3 depicts a classic semi-competing risks analysis with three states. Units begin in a working state (state 0) and are subject to a nonterminal event (state 1) and failure (state 2). Figure 3.5 depicts a semi-competing risks analysis with four states. Units begin in a working state (state 0) and are subject to two nonterminal events (states 1 and 2) and failure (state 3).

3.2 Two-state model

A standard survival or reliability analysis can be thought of as a two-state model, where units start in a working state (state 0) and eventually transition to a failed state (state 1). Let T denote the failure age. We assume a two-state model for each individual unit of the form shown in Figure 3.1. The two-state model is characterised by the transition hazard

$$\lambda(t \mid \boldsymbol{\theta}) = \lim_{\Delta \to 0} \frac{\Pr(T \in [t, t+\Delta) \mid T \geq t, \boldsymbol{\theta})}{\Delta}, \text{ for } t > 0, \tag{3.2.1}$$

where λ is the hazard rate (transition intensity) of the $0 \to 1$ transition, $\boldsymbol{\theta}$ is a vector of model parameters associated with λ, and states 0 and 1 are starting and failed states, respectively.

$$\boxed{\quad 0 \quad} \xrightarrow{\lambda(\cdot)} \boxed{\quad 1 \quad}$$

Figure 3.1: A two-state model.

3.2.1 Observed data under the two-state model

Let T_i denote the true failure age for unit i and let L_i and $C_i > L_i$ denote the left-truncation age and right-censoring age for unit i, respectively, which we assume are independent of T_i. Moreover, let $Y_i = \min(T_i, C_i)$ denote the observed failure age for unit i, with failure indicator $\delta_i = I\{T_i \leq C_i\}$, where $I(\cdot)$ is an indicator function, which is equal to 1 if $T_i \leq C_i$ and 0 otherwise. Data are left-truncated and satisfy the recruitment criterion $L_i < Y_i$; that is, units that failed before the study commenced are not included in the dataset. The observed data for the ith unit is $\mathcal{D}_i = \{l_i, y_i, \delta_i\}$. In words, for unit i, we observe the age at study entry, l_i; an indicator function, δ_i, which is equal to 1 if unit i failed and is equal to 0 if unit i is right-censored (has not failed); and the observed failure age, y_i. If unit i failed we observe the true failure age, and if unit i is right-censored, we observe the right-censoring age (a lower bound on the true failure age).

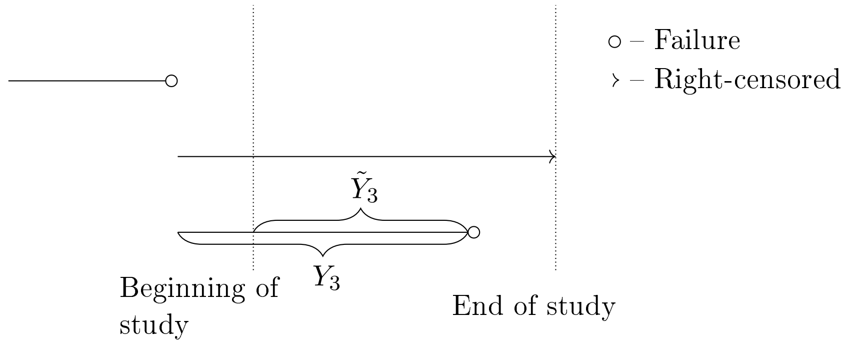

Figure 3.2: Two possible life histories under the two-state model and an example unit not included in the dataset.

Figure 3.2 provides a schematic representation of two possible life histories observed under the two-state model with left-truncation and right-censoring, and a description of a unit that is not included in the dataset since it did not satisfy the recruitment criterion, $L_i < Y_i$. We illustrate the difference between the failure age and the failure time using scenario 3 in Figure 3.2. The failure age of unit 3 is denoted Y_3 and the failure time is denoted $\tilde{Y}_3$. Unit 1 failed prior to the start of the study and was not included in the dataset. Unit 2 is right-censored at the end of the study. Unit 3 failed prior to the end of the study.

3.2.2 Likelihood under the two-state model

The observed data take the form $\mathcal{D}_n = \{l_i, y_i, \delta_i; i = 1, \ldots, n\}$. Assuming the ages of the units are independent, the censoring process is non-informative about T (Vakulenko-Lagun and Mandel, 2016), and that we do not have any knowledge about the truncation distribution (Vakulenko-Lagun and Mandel, 2016), the likelihood for the data, $\mathcal{D}_n$, can be written as

$$L(\boldsymbol{\theta}; \mathcal{D}_n) = \prod_{i=1}^{n} \frac{S(y_i \mid \boldsymbol{\theta}) \lambda^{\delta_i}(y_i \mid \boldsymbol{\theta})}{S(l_i \mid \boldsymbol{\theta})}, \tag{3.2.2}$$

as provided by Mittman et al. (2019), where $S(y_i \mid \boldsymbol{\theta}) = \exp\Big(- \int_0^{y_i} \lambda(u \mid \boldsymbol{\theta}) du \Big)$ is the probability that unit i does not transition from $0 \to 1$ by age y_i.

3.2.3 Posterior predictive survival distribution under the two-state model

Under the two-state model, the posterior predictive survival distribution is given by:

$$\Pr(T \geq \gamma + s \mid T > \gamma, \mathcal{D}_n) = \int \Pr(T \geq \gamma + s \mid T > \gamma, \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathcal{D}_n) d\boldsymbol{\theta}, \tag{3.2.3}$$

where

$$\Pr(T \geq \gamma + s \mid T > \gamma, \boldsymbol{\theta}) = S(\gamma + s \mid \gamma, \boldsymbol{\theta}). \tag{3.2.4}$$

In this thesis a Markov model is assumed for $\lambda(\cdot)$. Thus, $S(\gamma+s \mid \gamma, \boldsymbol{\theta}) = \exp\Big(-\int_{\gamma}^{\gamma+s} \lambda(u \mid \boldsymbol{\theta}) du\Big)$. Therefore, we obtain

$$\Pr(T \geq \gamma + s \mid T > \gamma, \boldsymbol{\theta}) = \frac{S(\gamma + s \mid \boldsymbol{\theta})}{S(\gamma \mid \boldsymbol{\theta})}, \tag{3.2.5}$$

In addition,

$$p(\boldsymbol{\theta} \mid \mathcal{D}_n) \tag{3.2.6}$$

is the posterior distribution of the two-state model parameters given the observed data, $\mathcal{D}_n$.

3.3 Illness-death model

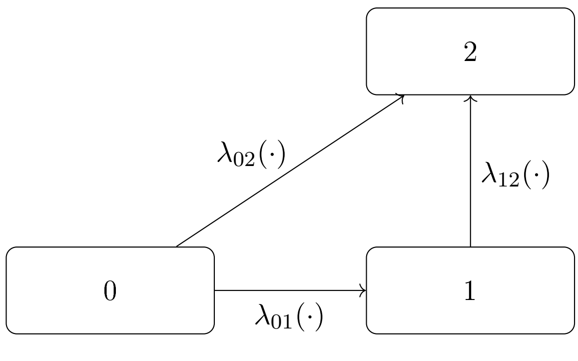

Figure 3.3: An illness-death model for semi-competing risks data.

This section focuses on modelling semi-competing risks data using the illness-death model. Let T_1 and T_2 denote the state 1 age and the failure age, respectively; where the stage 1 age is defined as the age a unit enters state 1, and the failure age is defined as the age a unit fails (or the age a unit enters state 2). We assume the convention of Xu et al. (2010) and Lee et al. (2021), by setting $T_1 = \infty$ for units which fail in the absence of the nonterminal event. We assume an illness-death model for each individual unit of the form shown in Figure 3.3. The illness-death model is characterised by the transition hazards:

$$\lambda_{01}(t_1 \mid \boldsymbol{\theta}_{01}) = \lim_{\Delta \to 0} \frac{\Pr(T_1 \in [t_1, t_1 + \Delta) \mid T_1 \geq t_1, T_2 \geq t_1, \boldsymbol{\theta}_{01})}{\Delta}, \text{ for } t_1 > 0 \tag{3.3.1}$$

$$\lambda_{02}(t_2 \mid \boldsymbol{\theta}_{02}) = \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 \geq t_2, T_2 \geq t_2, \boldsymbol{\theta}_{02})}{\Delta}, \text{ for } t_2 > 0 \tag{3.3.2}$$

$$\lambda_{12}(t_2 \mid T_1 = t_1, \boldsymbol{\theta}_{12}) = \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 = t_1, T_2 \geq t_2, \boldsymbol{\theta}_{12})}{\Delta}, \text{ for } 0 < t_1 < t_2, \tag{3.3.3}$$

where λ_{ij} is the hazard rate (transition intensity) of the $i \to j$ transition, $\boldsymbol{\theta}_{ij}$ is a vector of model parameters associated with λ_{ij}, and states 0, 1, and 2 correspond to the starting, state 1, and failed states, respectively. From Figures 3.1 and 3.3 we can see that the illness-death model is an extension of the two-state model. Removing state 1 from Figure 3.3 and relabeling state 2 to state 1 recovers the two-state model.

3.3.1 Observed data under the illness-death model

Let T_{i1} and T_{i2} denote the true state 1 age and failure age for unit i, respectively, and let L_i and $C_i > L_i$ denote the left-truncation age and right-censoring age for unit i, respectively, which we assume are independent of T_{i1} and T_{i2}. Moreover, let $Y_{i1} = \min(T_{i1}, T_{i2}, C_i)$ denote the observed state 1 age for unit i, with state 1 indicator $\delta_{i1} = I\{T_{i1} \leq \min(T_{i2}, C_i)\}$, and let $Y_{i2} = \min(T_{i2}, C_i)$ denote the observed failure age for unit i, with failure indicator $\delta_{i2} = I\{T_{i2} \leq C_i\}$.

Under the illness-death model there are multiple scenarios for left-truncation. For example, a study with a recruitment criterion that units must not have experienced the state 1 event prior to study entry or a study with a recruitment criterion that units must not have failed prior to study entry. In our application of interest, units must not have failed prior to study entry. Data are left-truncated and satisfy the recruitment criterion $L_i < Y_{i2}$; that is, units that failed before the study commenced are not included in the dataset. From herein, when we refer to left-truncation, under the illness-death model, we refer to the recruitment criterion $L_i < Y_{i2}$. The observed data for the ith unit is $\mathcal{D}_i = \{l_i, y_{i1}, \delta_{i1}, y_{i2}, \delta_{i2}\}$.

Figure 3.4 provides a schematic representation of the six possible life histories observed under the illness-death model with left-truncation and right-censoring, alongside two descriptions of units that are not included in the dataset since they do not satisfy the

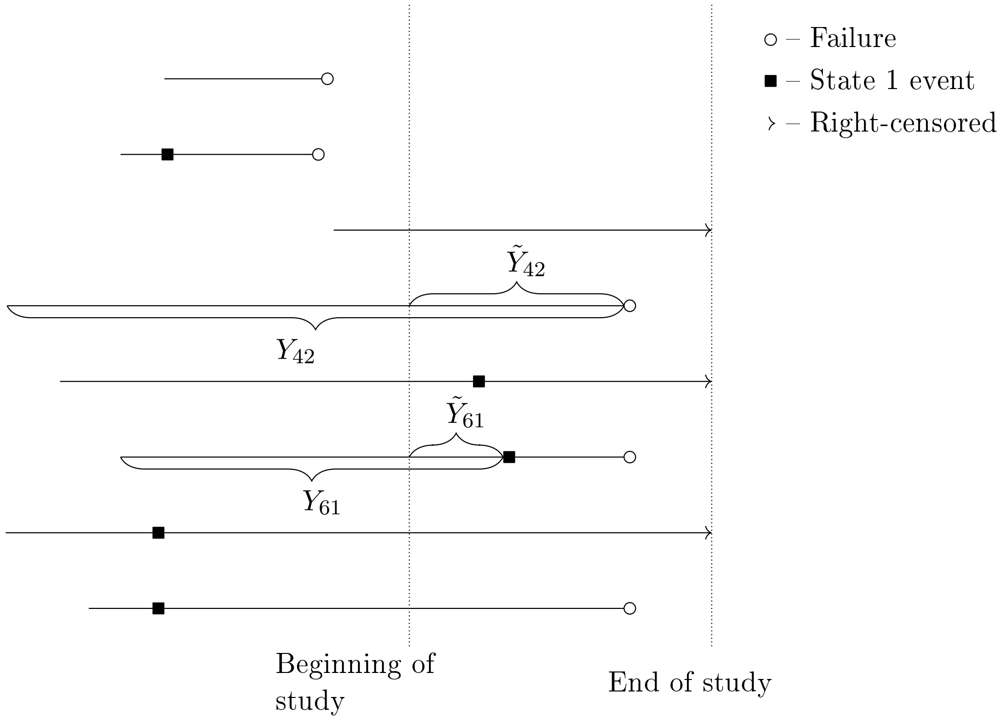

Figure 3.4: Six possible life histories under the illness-death model and two example units not included in the dataset.

recruitment criterion, $L_i < Y_{i2}$. We illustrate the difference between event ages and event times using scenarios 4 and 6 in Figure 3.4. The failure age of unit 4 is denoted Y_{42} and the failure time is denoted $\tilde{Y}_{42}$. The state 1 age of unit 6 is denoted Y_{61} and the state 1 time is denoted $\tilde{Y}_{61}$. Units 1 and 2 failed before the study began ($Y_{12} < L_1$, $Y_{22} < L_2$) and were not included in the dataset. Scenarios $3-8$ provide the six scenarios encountered under the illness-death model with left-truncation and right-censoring. Unit 3 was sampled in state 0 ($L_3 < Y_{31}, Y_{32}$) and remained in state 0 until the end of the study. Unit 3 is right-censored since it did not fail by the end of the study. Unit 4 was sampled in state 0 ($L_4 < Y_{41}, Y_{42}$) and remained in state 0 until failure ($T_{41} = \infty$). Unit 5 was sampled in state 0 ($L_5 < Y_{51}, Y_{52}$) and transitioned to state 1 prior to the end of the study ($Y_{51} < Y_{52}$). Unit 5 is right-censored since the study ended prior to the failure of this unit. Unit 6 was sampled in state 0 ($L_6 < Y_{61}, Y_{62}$) and transitioned to state 1 prior to failure ($Y_{61} < Y_{62}$). Units 7 and 8 were sampled in state 1 ($Y_{71} \leq L_7 < Y_{72}$, $Y_{81} \leq L_8 < Y_{82}$). Unit 7 is right-censored at the end of the study and unit 8 failed prior to the end of the study.

3.3.2 Likelihood under the illness-death model

In this case the observed data take the form $\mathcal{D}_n = \{l_i, y_{i1}, \delta_{i1}, y_{i2}, \delta_{i2}; i = 1, \ldots, n\}$. Assuming the ages of the units are independent, the censoring process is non-informative about T_1 and T_2, and that we do not have any knowledge about the truncation distribution, the likelihood for the data, $\mathcal{D}_n$, can be written as

$$L(\boldsymbol{\theta};\mathcal{D}_n) = \frac{L_0(\boldsymbol{\theta};\mathcal{D}_n) \times L_1(\boldsymbol{\theta};\mathcal{D}_n)}{L_{\text{Truncation}}(\boldsymbol{\theta};\mathcal{D}_n)}, \tag{3.3.4}$$

derived in Vakulenko-Lagun and Mandel (2016); where $\boldsymbol{\theta} = (\boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02}, \boldsymbol{\theta}_{12})$ is a vector of model parameters,

$$\begin{aligned} L_0(\boldsymbol{\theta};\mathcal{D}_n) = &\prod_{i:l_i<y_{i1},y_{i2}} \Big[S_{01}(y_{i2} \mid \boldsymbol{\theta}_{01}) S_{02}(y_{i2} \mid \boldsymbol{\theta}_{02}) \lambda_{02}^{\delta_{i2}}(y_{i2} \mid \boldsymbol{\theta}_{02}) \Big]^{1-\delta_{i1}} \\ &\times \Big[S_{01}(y_{i1} \mid \boldsymbol{\theta}_{01}) S_{02}(y_{i1} \mid \boldsymbol{\theta}_{02}) \lambda_{01}(y_{i1} \mid \boldsymbol{\theta}_{01}) S_{12}(y_{i2} \mid y_{i1}, \boldsymbol{\theta}_{12}) \lambda_{12}^{\delta_{i2}}(y_{i2} \mid y_{i1}, \boldsymbol{\theta}_{12}) \Big]^{\delta_{i1}}, \end{aligned} \tag{3.3.5}$$

is the contribution to the likelihood from units sampled in state 0,

$$\begin{aligned} L_1(\boldsymbol{\theta};\mathcal{D}_n) = \prod_{i:y_{i1}\le l_i<y_{i2}} \int_0^{l_i} &S_{01}(t \mid \boldsymbol{\theta}_{01}) S_{02}(t \mid \boldsymbol{\theta}_{02}) \lambda_{01}(t \mid \boldsymbol{\theta}_{01}) \\ &\times S_{12}(y_{i2} \mid t, \boldsymbol{\theta}_{12}) \lambda_{12}^{\delta_{i2}}(y_{i2} \mid t, \boldsymbol{\theta}_{12}) dt, \end{aligned} \tag{3.3.6}$$

is the contribution to the likelihood from units sampled in state 1, and

$$\begin{aligned} L_{\text{Truncation}}(\boldsymbol{\theta};\mathcal{D}_n) = \prod_i \Big\{ &S_{01}(l_i \mid \boldsymbol{\theta}_{01}) S_{02}(l_i \mid \boldsymbol{\theta}_{02}) \\ &+ \int_0^{l_i} S_{01}(t \mid \boldsymbol{\theta}_{01}) S_{02}(t \mid \boldsymbol{\theta}_{02}) \lambda_{01}(t \mid \boldsymbol{\theta}_{01}) S_{12}(l_i \mid t, \boldsymbol{\theta}_{12}) dt \Big\}, \end{aligned} \tag{3.3.7}$$

is the likelihood of survival up to the sampling age. In addition, $S_{0k}(t \mid \boldsymbol{\theta}_{0k}) = \exp\Big(- \int_0^t \lambda_{0k}(u \mid \boldsymbol{\theta}_{0k}) du \Big)$, for $k = 1, 2$, is the probability that unit i does not transition from $0 \to k$ by age t; and $S_{12}(t_2 \mid t_1, \boldsymbol{\theta}_{12}) = \exp\Big(- \int_{t_1}^{t_2} \lambda_{12}(u \mid \boldsymbol{\theta}_{12}) du \Big)$ is the probability that unit i does not transition from $1 \to 2$ by age t_2 given that unit i is in state 1 at age t_1.

The first term in Equation (3.3.5) is the likelihood contribution from a unit that is sampled in state 0 and does not transition to state 1 prior to the end of the study or failure (see scenarios 3 and 4 in Figure 3.4). In other words, this contribution is the likelihood of staying in state 0 up until age y_{i2}, for right-censored units, and then moving from state

0 to state 2, at age y_{i2}, for units that failed. The second term in Equation (3.3.5) is the likelihood contribution from a unit that is sampled in state 0 and transitions to state 1 prior to the end of the study or failure (see scenarios 5 and 6 in Figure 3.4). In other words, this contribution is the likelihood of staying in state 0 up until age y_{i1}, transitioning to state 1 at y_{i1}, and remaining in state 1 until age y_{i2}, for right-censored units, and moving from state 1 to state 2, at age y_{i2}, for units that failed.

Equation (3.3.6) is the likelihood contribution from units that are sampled in state 1 (see scenarios 7 and 8 in Figure 3.4). For these units, we know they transitioned to state 1 at age $y_{i1} \leq l_i < y_{i2}$ and remained in state 1 until the end of the study if right-censored, or until failure for units that failed. The integrand is identical to the second term in Equation (3.3.5), however, for units sampled in state 1, we need to integrate over all possible transition ages.

Equation (3.3.7) is the likelihood of survival up to the sampling age. We need to calculate the probability of survival up to the sampling age, l_i, for each unit. This probability is a sum of two terms: the probability of being sampled in state 0 (term 1), and the probability of being sampled in state 1 (term 2).

3.3.2.1 Posterior predictive survival distributions under the illness-death model

Let $z_i(\gamma)$ represent the state of drive i at age γ. Drive i is either in the healthy state, $\{0\}$, the critical state, $\{1\}$, or the failed state, $\{2\}$. Under the illness-death model, the posterior predictive survival distributions are given by:

$$\Pr(T_2 \geq \gamma+s \mid T_2 > \gamma, z(\gamma), \mathcal{D}_n) = \int \Pr(T_2 \geq \gamma+s \mid T_2 > \gamma, z(\gamma), \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathcal{D}_n) d\boldsymbol{\theta}, \quad (3.3.8)$$

for $z_i(\gamma) = 0, 1$, where

$$\begin{aligned} \Pr(T_2 \geq \gamma + s \mid T_2 > \gamma, z(\gamma) = 0, \boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02}, \boldsymbol{\theta}_{12}) = \frac{1}{S_{01}(\gamma \mid \boldsymbol{\theta}_{01}) S_{02}(\gamma \mid \boldsymbol{\theta}_{02})} \Bigg\{ \\ S_{01}(\gamma + s \mid \boldsymbol{\theta}_{01}) S_{02}(\gamma + s \mid \boldsymbol{\theta}_{02}) + \\ \int_{\gamma}^{\gamma+s} S_{01}(\nu \mid \boldsymbol{\theta}_{01}) S_{02}(\nu \mid \boldsymbol{\theta}_{02}) \lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) S_{12}(\gamma + s \mid \nu, \boldsymbol{\theta}_{12}) d\nu \Bigg\}, \end{aligned} \quad (3.3.9)$$

for drives in the healthy state at age γ, and

$$\Pr(T_2 \geq \gamma + s \mid T_2 > \gamma, z(\gamma) = 1, \boldsymbol{\theta}_{12}) = S_{12}(\gamma + s \mid \gamma, \boldsymbol{\theta}_{12}), \tag{3.3.10}$$

for drives in the critical state at age γ. In Equation (3.3.9), the first term represents the probability the drive remained in the healthy state from age γ to $\gamma + s$, and the second term is the probability the drive transitioned from the healthy state to the critical state at age $\nu \in (\gamma, \gamma + s)$ and then remained in the critical state from age ν to $\gamma + s$, integrated over ν.

In addition,

$$p(\boldsymbol{\theta} \mid \mathcal{D}_n) \tag{3.3.11}$$

is the posterior distribution of the illness-death model parameters given the observed data, $\mathcal{D}_n$.

3.4 Multi-state model

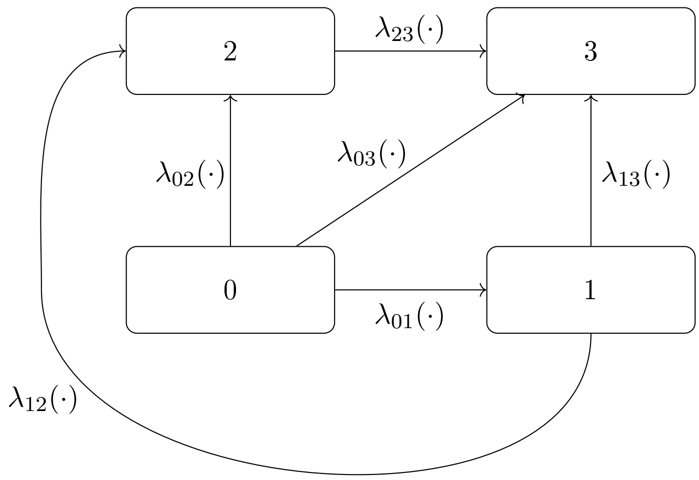

Figure 3.5: A multi-state model with four states.

Figure 3.5 depicts a multi-state model with four states: namely, the starting state (state 0), two nonterminal states (state 1 and state 2) and the failed state (state 3).

Let T_1, T_2, and T_3 denote state 1, state 2, and failure ages, respectively; where the state j age, for $j = 1, 2$, is defined as the age a unit enters state j, and the failure age is defined as the age a unit fails (or the age a unit enters state 3). We assume $T_1 = T_2 = \infty$ for units which fail in the absence of the nonterminal events. We assume a multi-state model for each individual unit of the form shown in Figure 3.5. Following the approach from Section 3.3, the multi-state model is characterised by the transition hazards:

$$\lambda_{01}(t_1 \mid \boldsymbol{\theta}_{01}) = \lim_{\Delta \to 0} \frac{\Pr(T_1 \in [t_1, t_1 + \Delta) \mid T_1 \geq t_1, T_2 \geq t_1, T_3 \geq t_1, \boldsymbol{\theta}_{01})}{\Delta}, \quad \text{for } t_1 > 0 \tag{3.4.1}$$

$$\lambda_{02}(t_2 \mid \boldsymbol{\theta}_{02}) = \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 \geq t_2, T_2 \geq t_2, T_3 \geq t_2, \boldsymbol{\theta}_{02})}{\Delta}, \quad \text{for } t_2 > 0 \tag{3.4.2}$$

$$\lambda_{03}(t_3 \mid \boldsymbol{\theta}_{03}) = \lim_{\Delta \to 0} \frac{\Pr(T_3 \in [t_3, t_3 + \Delta) \mid T_1 \geq t_3, T_2 \geq t_3, T_3 \geq t_3, \boldsymbol{\theta}_{03})}{\Delta}, \quad \text{for } t_3 > 0 \tag{3.4.3}$$

$$\lambda_{12}(t_2 \mid T_1 = t_1, \boldsymbol{\theta}_{12}) = \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 = t_1, T_2 \geq t_2, T_3 \geq t_2, \boldsymbol{\theta}_{12})}{\Delta}, \quad \text{for } 0 < t_1 < t_2 \tag{3.4.4}$$

$$\lambda_{13}(t_3 \mid T_1 = t_1, \boldsymbol{\theta}_{13}) = \lim_{\Delta \to 0} \frac{\Pr(T_3 \in [t_3, t_3 + \Delta) \mid T_1 = t_1, T_2 \geq t_3, T_3 \geq t_3, \boldsymbol{\theta}_{13})}{\Delta}, \quad \text{for } 0 < t_1 < t_3 \tag{3.4.5}$$

$$\lambda_{23}(t_3 \mid T_2 = t_2, \boldsymbol{\theta}_{23}) = \lim_{\Delta \to 0} \frac{\Pr(T_3 \in [t_3, t_3 + \Delta) \mid T_2 = t_2, T_3 \geq t_3, \boldsymbol{\theta}_{23})}{\Delta}, \quad \text{for } 0 < t_2 < t_3, \tag{3.4.6}$$

where λ_{ij} is the hazard rate (transition intensity) of the $i \to j$ transition, $\boldsymbol{\theta}_{ij}$ is a vector of model parameters associated with λ_{ij}, and states 0, 1, 2, and 3 correspond to the starting, state 1, state 2, and failed states, respectively. From Figures 3.3 and 3.5 we can see that the multi-state model is an extension of the illness-death model. Removing state 2 from Figure 3.5 and relabeling state 3 to state 2 recovers the illness-death model.

3.4.1 Observed data under the multi-state model

Let T_{i1}, T_{i2}, and T_{i3} denote the true state 1, state 2, and failure ages for unit i, respectively, and let L_i and $C_i > L_i$ denote the left-truncation age and right-censoring age for unit i, respectively, which we assume are independent of T_{i1}, T_{i2}, and T_{i3}. Moreover, let $Y_{i1} = \min(T_{i1}, T_{i2}, T_{i3}, C_i)$ denote the observed state 1 age for unit i, with state 1 indicator $\delta_{i1} = I\{T_{i1} \leq \min(T_{i2}, T_{i3}, C_i)\}$, let $Y_{i2} = \min(T_{i2}, T_{i3}, C_i)$ denote the observed state 2 age for unit i, with state 2 indicator $\delta_{i2} = I\{T_{i2} \leq \min(T_{i3}, C_i)\}$, and let $Y_{i3} = \min(T_{i3}, C_i)$ denote the observed failure age for unit i, with failure indicator $\delta_{i3} = I\{T_{i3} \leq C_i\}$.

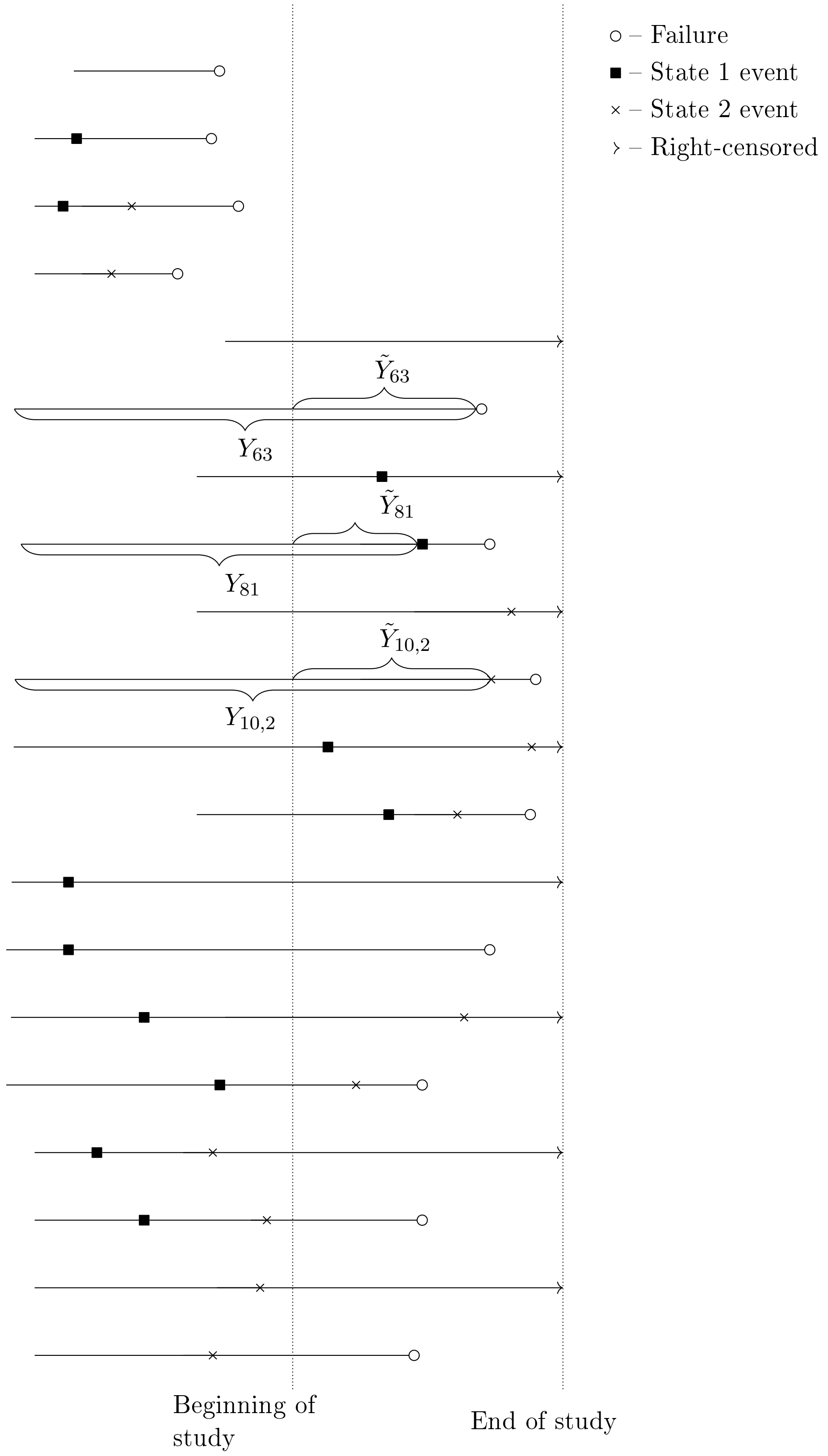

Figure 3.6: Sixteen possible life histories under the multi-state model and four example units not included in the dataset.

Under the multi-state model there are multiple scenarios for left-truncation. In our application of interest, units must not have failed prior to study entry. Data are left-truncated and satisfy the recruitment criterion $L_i < Y_{i3}$; that is, units that failed before the study commenced are not included in the dataset. From herein, when we refer to left-truncation, under the multi-state model, we refer to the recruitment criterion $L_i < Y_{i3}$. The observed data for the ith unit is $\mathcal{D}_i = \{l_i, y_{i1}, \delta_{i1}, y_{i2}, \delta_{i2}, y_{i3}, \delta_{i3}\}$.

Figure 3.6 provides a schematic representation of the sixteen possible life histories observed under the multi-state model with left-truncation and right-censoring, alongside four descriptions of units that are not included in the dataset since they do not satisfy the recruitment criterion, $L_i < Y_{i3}$. We illustrate the difference between event ages and event times using scenarios 6, 8 and 10 in Figure 3.6. The failure age of unit 6 is denoted Y_{63} and the failure time is denoted $\tilde{Y}_{63}$. The state 1 age of unit 8 is denoted Y_{81} and the state 1 time is denoted $\tilde{Y}_{81}$. The state 2 age of unit 10 is denoted $Y_{10,2}$ and the state 2 time is denoted $\tilde{Y}_{10,2}$. Units $1-4$ failed before the study began ($Y_{i3} < L_i$, for $i = 1, \ldots, 4$) and were not included in the dataset.

Scenarios $5-20$ provide the sixteen scenarios encountered under the multi-state model with left-truncation and right-censoring. Units $5-12$ were sampled in state 0 ($L_i < Y_{i1}, Y_{i2}, Y_{i3}$, for $i = 5, \ldots, 12$). Unit 5 remained in state 0 until the end of the study. This unit is right-censored since it did not fail prior to the end of the study. Unit 6 remained in state 0 until failure. Units 7, 9, and 11 are right-censored at the end of the study; unit 7 entered state 1 (but not state 2) prior to the end of the study; unit 9 entered state 2 (but not state 1) prior to the end of the study; and unit 11 entered state 1 and state 2 prior to the end of the study. Units 8, 10, and 12 failed prior to the end of the study; unit 8 entered state 1 (but not state 2) prior to failure; unit 10 entered state 2 (but not state 1) prior to failure; and unit 12 entered state 1 and state 2 prior to failure. Units $13-16$ were sampled in state 1 ($Y_{i1} \leq L_i < Y_{i2}, Y_{i3}$, for $i = 13, \ldots, 16$). Units 13 and 15 are right-censored at the end of the study; unit 13 remained in state 1 until the end of the study; and unit 15 entered state 2 prior to the end of the study. Units 14 and 16 failed prior to the end of the study; unit 14 remained in state 1 until failure and unit 16 entered state 2 prior to failure. Units 17 and 18 entered state 1 and state 2 prior to the start of the study; and hence were sampled in state 2 ($Y_{i1}, Y_{i2} \leq L_i < Y_{i3}$, for $i = 17, 18$). Unit 17 is right-censored at the end of the study and unit 18 failed prior to the end of the study. Units 19 and 20 entered state 2, but not state 1, prior to the start of the study; and hence were sampled in state 2 ($Y_{i2} \leq L_i < Y_{i3}$, $T_{i1} = \infty$, for $i = 19, 20$). Unit 19 is right-censored at the end of the study and unit 20 failed prior to the end of the study.

For units sampled in state 2, we do not know whether the unit transitioned from state 0 to state 1, and then to state 2 prior to the start of the study (see scenarios 17 and 18 in Figure 3.6), or if the unit transitioned from state 0 directly to state 2, without entering state 1, prior to the start of the study (see scenarios 19 and 20 in Figure 3.6). We only observe that these units are in state 2 upon study entry.

3.4.2 Likelihood under the multi-state model

In this case the observed data take the form $\mathcal{D}_n = \{l_i, y_{i1}, \delta_{i1}, y_{i2}, \delta_{i2}, y_{i3}, \delta_{i3}; i = 1, \ldots, n\}$. Assuming the ages of the units are independent, the censoring process is non-informative about T_1, T_2, and T_3, and that we do not have any knowledge about the truncation distribution, the likelihood for the data, $\mathcal{D}_n$, can be written as

$$L(\boldsymbol{\theta}; \mathcal{D}_n) = \frac{L_0(\boldsymbol{\theta}; \mathcal{D}_n) \times L_1(\boldsymbol{\theta}; \mathcal{D}_n) \times L_2(\boldsymbol{\theta}; \mathcal{D}_n)}{L_{\text{Truncation}}(\boldsymbol{\theta}; \mathcal{D}_n)}, \tag{3.4.7}$$

derived in Vakulenko-Lagun and Mandel (2016); where $\boldsymbol{\theta} = (\boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02}, \boldsymbol{\theta}_{03}, \boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{13}, \boldsymbol{\theta}_{23})$ is a vector of model parameters, and

$$\begin{aligned}
L_0(\boldsymbol{\theta}; \mathcal{D}_n) = &\prod_{i: l_i < y_{i1}, y_{i2}, y_{i3}} \\
&\Big[S_{01}(y_{i3} \mid \boldsymbol{\theta}_{01}) S_{02}(y_{i3} \mid \boldsymbol{\theta}_{02}) S_{03}(y_{i3} \mid \boldsymbol{\theta}_{03}) \lambda_{03}^{\delta_{i3}}(y_{i3} \mid \boldsymbol{\theta}_{03})\Big]^{(1-\delta_{i1})(1-\delta_{i2})} \times \\
&\Big[S_{01}(y_{i1} \mid \boldsymbol{\theta}_{01}) S_{02}(y_{i1} \mid \boldsymbol{\theta}_{02}) S_{03}(y_{i1} \mid \boldsymbol{\theta}_{03}) \lambda_{01}(y_{i1} \mid \boldsymbol{\theta}_{01}) \\
&\times S_{12}(y_{i3} \mid y_{i1}, \boldsymbol{\theta}_{12}) S_{13}(y_{i3} \mid y_{i1}, \boldsymbol{\theta}_{13}) \lambda_{13}^{\delta_{i3}}(y_{i3} \mid y_{i1}, \boldsymbol{\theta}_{13})\Big]^{\delta_{i1}(1-\delta_{i2})} \times \\
&\Big[S_{01}(y_{i2} \mid \boldsymbol{\theta}_{01}) S_{02}(y_{i2} \mid \boldsymbol{\theta}_{02}) S_{03}(y_{i2} \mid \boldsymbol{\theta}_{03}) \lambda_{02}(y_{i2} \mid \boldsymbol{\theta}_{02}) \\
&S_{23}(y_{i3} \mid y_{i2}, \boldsymbol{\theta}_{23}) \lambda_{23}^{\delta_{i3}}(y_{i3} \mid y_{i2}, \boldsymbol{\theta}_{23})\Big]^{(1-\delta_{i1})\delta_{i2}} \\
&\Big[S_{01}(y_{i1} \mid \boldsymbol{\theta}_{01}) S_{02}(y_{i1} \mid \boldsymbol{\theta}_{02}) S_{03}(y_{i1} \mid \boldsymbol{\theta}_{03}) \lambda_{01}(y_{i1} \mid \boldsymbol{\theta}_{01}) S_{12}(y_{i2} \mid y_{i1}, \boldsymbol{\theta}_{12}) \\
&S_{13}(y_{i2} \mid y_{i1}, \boldsymbol{\theta}_{13}) \lambda_{12}(y_{i2} \mid y_{i1}, \boldsymbol{\theta}_{12}) S_{23}(y_{i3} \mid y_{i2}, \boldsymbol{\theta}_{23}) \lambda_{23}^{\delta_{i3}}(y_{i3} \mid y_{i2}, \boldsymbol{\theta}_{23})\Big]^{\delta_{i1}\delta_{i2}},
\end{aligned} \tag{3.4.8}$$

is the contribution to the likelihood from units sampled in state 0,

$$
\begin{aligned}
L_1(\boldsymbol{\theta};\mathcal{D}_n) = & \prod_{i:y_{i1}\le l_i<y_{i2},y_{i3}} \int_0^{l_i} \Big[S_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{03}(\nu \mid \boldsymbol{\theta}_{03})\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) \\
& \times S_{12}(y_{i3} \mid \nu, \boldsymbol{\theta}_{12})S_{13}(y_{i3} \mid \nu, \boldsymbol{\theta}_{13})\lambda_{13}^{\delta_{i3}}(y_{i3} \mid \nu, \boldsymbol{\theta}_{13})\Big]^{\delta_{i1}(1-\delta_{i2})} \\
& \times \Big[S_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{03}(\nu \mid \boldsymbol{\theta}_{03})\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{12}(y_{i2} \mid \nu, \boldsymbol{\theta}_{12})S_{13}(y_{i2} \mid \nu, \boldsymbol{\theta}_{13}) \\
& \times \lambda_{12}(y_{i2} \mid \nu, \boldsymbol{\theta}_{12})S_{23}(y_{i3} \mid y_{i2}, \boldsymbol{\theta}_{23})\lambda_{23}^{\delta_{i3}}(y_{i3} \mid y_{i2}, \boldsymbol{\theta}_{23})\Big]^{\delta_{i1}\delta_{i2}} d\nu,
\end{aligned}
\tag{3.4.9}
$$

is the contribution to the likelihood from units sampled in state 1, and

$$
\begin{aligned}
L_2(\boldsymbol{\theta};\mathcal{D}_n) = & \prod_{i:y_{i2}\le l_i<y_{i3},t_{i1}=\infty} \\
& \int_0^{l_i} w(\nu \mid \boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02})\Big[S_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{03}(\nu \mid \boldsymbol{\theta}_{03})\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) \\
& \times \int_\nu^{l_i} S_{12}(\xi \mid \nu, \boldsymbol{\theta}_{12})S_{13}(\xi \mid \nu, \boldsymbol{\theta}_{13})\lambda_{12}(\xi \mid \nu, \boldsymbol{\theta}_{12})S_{23}(y_{i3} \mid \xi, \boldsymbol{\theta}_{23})\lambda_{23}^{\delta_{i3}}(y_{i3} \mid \xi, \boldsymbol{\theta}_{23})\Big] d\xi d\nu, \\
& + \prod_{i:y_{i1},y_{i2}\le l_i<y_{i3}} \int_0^{l_i} [1 - w(\nu \mid \boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02})]\Big[S_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{03}(\nu \mid \boldsymbol{\theta}_{03}) \\
& \times \lambda_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{23}(y_{i3} \mid \nu, \boldsymbol{\theta}_{23})\lambda_{23}^{\delta_{i3}}(y_{i3} \mid \nu, \boldsymbol{\theta}_{23})\Big] d\nu
\end{aligned}
\tag{3.4.10}
$$

where

$$
w(\nu \mid \boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02}) = \frac{\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01})}{\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) + \lambda_{02}(\nu \mid \boldsymbol{\theta}_{02})}
\tag{3.4.11}
$$

and

$$
1 - w(\nu \mid \boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02}) = 1 - \frac{\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01})}{\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) + \lambda_{02}(\nu \mid \boldsymbol{\theta}_{02})} = \frac{\lambda_{02}(\nu \mid \boldsymbol{\theta}_{02})}{\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) + \lambda_{02}(\nu \mid \boldsymbol{\theta}_{02})},
\tag{3.4.12}
$$

is the contribution to the likelihood from units sampled in state 2. In addition,

$$L_{\text{Truncation}}(\boldsymbol{\theta};\mathcal{D}_n) = \prod_i \Bigg\{ S_{01}(l_i \mid \boldsymbol{\theta}_{01})S_{02}(l_i \mid \boldsymbol{\theta}_{02})S_{03}(l_i \mid \boldsymbol{\theta}_{03})$$
$$+ \int_0^{l_i} S_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{03}(\nu \mid \boldsymbol{\theta}_{03})\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{12}(l_i \mid \nu,\boldsymbol{\theta}_{12})S_{13}(l_i \mid \nu,\boldsymbol{\theta}_{13})d\nu$$
$$+ \int_0^{l_i} S_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{03}(\nu \mid \boldsymbol{\theta}_{03})\lambda_{01}(\nu \mid \boldsymbol{\theta}_{01})$$
$$\times \int_\nu^{l_i} S_{12}(\xi \mid \nu,\boldsymbol{\theta}_{12})S_{13}(\xi \mid \nu,\boldsymbol{\theta}_{13})\lambda_{12}(\xi \mid \nu,\boldsymbol{\theta}_{12})S_{23}(l_i \mid \xi,\boldsymbol{\theta}_{23})d\xi d\nu$$
$$+ \int_0^{l_i} S_{01}(\nu \mid \boldsymbol{\theta}_{01})S_{02}(\nu \mid \boldsymbol{\theta}_{02})S_{03}(\nu \mid \boldsymbol{\theta}_{03})\lambda_{02}(\nu \mid \boldsymbol{\theta}_{01})S_{23}(l_i \mid \nu,\boldsymbol{\theta}_{12})d\nu \Bigg\}, \tag{3.4.13}$$

is the likelihood of survival up to the sampling age, $S_{0k}(\nu \mid \boldsymbol{\theta}_{0k}) = \exp\Big(-\int_0^\nu \lambda_{0k}(u \mid \boldsymbol{\theta}_{0k})du\Big)$, for $k = 1,2,3$, is the probability that unit i does not transition from $0 \to k$ by age ν; $S_{1k}(\xi \mid \nu,\boldsymbol{\theta}_{1k}) = \exp\Big(-\int_\nu^\xi \lambda_{1k}(u \mid \boldsymbol{\theta}_{1k})du\Big)$, for $k = 2,3$, is the probability that unit i does not transition from $1 \to k$ by age ξ given that unit i is in state 1 at age ν; and $S_{23}(\xi \mid \nu,\boldsymbol{\theta}_{23}) = \exp\Big(-\int_\nu^\xi \lambda_{23}(u \mid \boldsymbol{\theta}_{23})du\Big)$ is the probability that unit i does not transition from $2 \to 3$ by age ξ given that unit i is in state 2 at age ν.

Equation (3.4.8) is the likelihood contribution from units that are sampled in state 0. The first term in Equation (3.4.8) is the likelihood contribution from a unit that is sampled in state 0 and does not transition to any intermediate state prior to the end of the study, if the unit is right-censored, or prior to failure, if the unit failed (see scenarios 5 and 6 in Figure 3.6). The second term is the likelihood contribution from a unit that is sampled in state 0 and transitions to state 1 (but does not enter state 2) prior to the end of the study, if the unit is right-censored, or prior to failure, if the unit failed (see scenarios 7 and 8 in Figure 3.6). The third term is the likelihood contribution from a unit that is sampled in state 0 and transitions to state 2 (but does not enter state 1) prior to the end of the study, if the unit is right-censored, or prior to failure, if the unit failed (see scenarios 9 and 10 in Figure 3.6). The fourth term is the likelihood contribution from a unit that is sampled in state 0 and transitions to state 1 and then to state 2 prior to the end of the study, if the unit is right-censored, or prior to failure, if the unit failed (see scenarios 11 and 12 in Figure 3.6).

Equation (3.4.9) is the likelihood contribution from units that are sampled in state 1. The first term is the likelihood contribution from a unit that is sampled in state 1 and does not transition to state 2 prior to the end of the study, if the unit is right-censored, or prior to failure, if the unit failed (see scenarios 13 and 14 in Figure 3.6). The second term is the likelihood contribution from a unit that is sampled in state 1 and transitions to state 2 prior to the end of the study, if the unit is right-censored, or prior to failure, if the unit failed (see scenarios 15 and 16 in Figure 3.6).

Equation (3.4.10) is the likelihood contribution from units that are sampled in state 2. For units sampled in state 2, we do not know whether these units transitioned from state 0 to state 1, and then to state 2 prior to the start of the study (see scenarios 17 and 18 in Figure 3.6), or if these units transitioned from state 0 directly to state 2, without entering state 1, prior to the start of the study (see scenarios 19 and 20 in Figure 3.6). Therefore, the contribution to the likelihood for these units is a weighted average of both possibilities.

In other words, the first term in Equation (3.4.10) is the likelihood contribution of a unit that stays in state 0 until age ν, transitions to state 1 at age ν, remains in state 1 until age ξ, transitions to state 2 at age ξ, and remains in state 2 until age y_{i3}, if the unit is right-censored, or moves from state 2 to the failed state, at age y_{i3}, if the unit failed, weighted by the probability of the unit moving to state 1 at age ν conditional on the unit moving to either state 1 or state 2 at age ν, integrated over ξ and ν. The second term is the likelihood contribution of a unit that stays in state 0 until age ν, transitions to state 2 at age ν, and remains in state 2 until age y_{i3}, if the unit is right-censored, or moves from state 2 to the failed state, at age y_{i3}, if the unit failed, weighted by the probability of the unit moving to state 2 at age ν conditional on the unit moving to either state 1 or state 2 at age ν, integrated over ν.

Equation (3.4.13) is the likelihood of survival up to the sampling age. We need to calculate the probability of survival up to the sampling age, l_i, for each unit. This probability is a sum of four terms: the probability of being sampled in state 0 (term 1), the probability of being sampled in state 1 (term 2), the probability of being sampled in state 2 after entering state 1 (term 3), and the probability of being sampled in state 2 without entering state 1 (term 4).

3.4.3 Posterior predictive survival distributions under the multi-state model

Let $z_i(\gamma)$ represent the state of drive i at age γ. Drive i is either in the healthy state, $\{0\}$, the critical 1 state, $\{1\}$, the critical 2 state, $\{2\}$, or the failed state, $\{3\}$. Under the multi-state model, the posterior predictive survival distributions are given by:

$$\Pr(T_3 \geq \gamma + s \mid T_3 > \gamma, z(\gamma), \mathcal{D}_n) = \int \Pr(T_3 \geq \gamma + s \mid T_3 > \gamma, z(\gamma), \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathcal{D}_n) d\boldsymbol{\theta}, \tag{3.4.14}$$

for $z_i(\gamma) = 0, 1$, where

$$
\begin{aligned}
&\Pr(T_3 \geq \gamma + s \mid T_3 > \gamma, z(\gamma) = 0, \boldsymbol{\theta}_{01}, \boldsymbol{\theta}_{02}, \boldsymbol{\theta}_{03}, \boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{13}, \boldsymbol{\theta}_{23}) = \frac{1}{S_{01}(\gamma \mid \boldsymbol{\theta}_{01}) S_{02}(\gamma \mid \boldsymbol{\theta}_{02}) S_{03}(\gamma \mid \boldsymbol{\theta}_{03})} \Bigg\{ \\
&S_{01}(\gamma + s \mid \boldsymbol{\theta}_{01}) S_{02}(\gamma + s \mid \boldsymbol{\theta}_{02}) S_{03}(\gamma + s \mid \boldsymbol{\theta}_{03}) \\
&+ \int_{\gamma}^{\gamma+s} S_{01}(\nu \mid \boldsymbol{\theta}_{01}) S_{02}(\nu \mid \boldsymbol{\theta}_{02}) S_{03}(\nu \mid \boldsymbol{\theta}_{03}) \lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) S_{12}(\gamma + s \mid \nu, \boldsymbol{\theta}_{12}) S_{13}(\gamma + s \mid \nu, \boldsymbol{\theta}_{13}) d\nu \\
&+ \int_{\gamma}^{\gamma+s} S_{01}(\nu \mid \boldsymbol{\theta}_{01}) S_{02}(\nu \mid \boldsymbol{\theta}_{02}) S_{03}(\nu \mid \boldsymbol{\theta}_{03}) \lambda_{01}(\nu \mid \boldsymbol{\theta}_{01}) \\
&\times \int_{\nu}^{\gamma+s} S_{12}(\xi \mid \nu, \boldsymbol{\theta}_{12}) S_{13}(\xi \mid \nu, \boldsymbol{\theta}_{13}) \lambda_{12}(\xi \mid \boldsymbol{\theta}_{12}) S_{23}(\gamma + s \mid \xi, \boldsymbol{\theta}_{23}) d\xi d\nu \\
&+ \int_{\gamma}^{\gamma+s} S_{01}(\nu \mid \boldsymbol{\theta}_{01}) S_{02}(\nu \mid \boldsymbol{\theta}_{02}) S_{03}(\nu \mid \boldsymbol{\theta}_{03}) \lambda_{02}(\nu \mid \boldsymbol{\theta}_{02}) S_{23}(\gamma + s \mid \nu, \boldsymbol{\theta}_{23}) d\nu \Bigg\},
\end{aligned} \tag{3.4.15}
$$

for drives in the healthy state at age γ,

$$
\begin{aligned}
&\Pr(T_3 \geq \gamma + s \mid T_3 > \gamma, z(\gamma) = 1, \boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{13}, \boldsymbol{\theta}_{23}) = \frac{1}{S_{12}(\gamma \mid \boldsymbol{\theta}_{12}) S_{13}(\gamma \mid \boldsymbol{\theta}_{13})} \Bigg\{ \\
&S_{12}(\gamma + s \mid \boldsymbol{\theta}_{12}) S_{13}(\gamma + s \mid \boldsymbol{\theta}_{13}) + \int_{\gamma}^{\gamma+s} S_{12}(\nu \mid \boldsymbol{\theta}_{12}) S_{13}(\nu \mid \boldsymbol{\theta}_{13}) \lambda_{12}(\nu \mid \boldsymbol{\theta}_{12}) S_{23}(\gamma + s \mid \nu, \boldsymbol{\theta}_{23}) d\nu \Bigg\},
\end{aligned} \tag{3.4.16}
$$

for drives in the critical 1 state at age γ, and

$$
\Pr(T_3 \geq \gamma + s \mid T_3 > \gamma, z(\gamma) = 2, \boldsymbol{\theta}_{23}) = S_{23}(\gamma + s \mid \gamma, \boldsymbol{\theta}_{23}), \tag{3.4.17}
$$

for drives in the critical 2 state at age γ. In Equation (3.4.15), the first term represents the probability the drive remained in the healthy state from age γ to $\gamma + s$; the second term is the probability the drive transitioned from the healthy state to the critical 1 state at age $\nu \in (\gamma, \gamma + s)$ and remained in the critical 1 state from age ν to $\gamma + s$, integrated over ν; the third term is the probability the drive transitioned from the healthy state to the critical 1 state at age $\nu \in (\gamma, \gamma + s)$, then transitioned from the critical 1 state to the critical 2 state at age $\xi \in (\nu, \gamma + s)$ and remained in the critical 2 state from age ξ to $\gamma + s$, integrated over ξ and ν; and the fourth term is the probability the drive transitioned from the healthy state to the critical 2 state at age $\nu \in (\gamma, \gamma + s)$ and remained in the critical 2 state from age ν to $\gamma + s$, integrated over ν.

In Equation (3.4.16), the first term represents the probability the drive remained in the critical 1 state from age γ to $\gamma + s$, and the second term is the probability the drive transitioned from the critical 1 state to the critical 2 state at age $\nu \in (\gamma, \gamma + s)$ and remained in the critical 2 state from age ν to $\gamma + s$, integrated over ν.

In addition,

$$p(\boldsymbol{\theta} \mid \mathcal{D}_n) \tag{3.4.18}$$

is the posterior distribution of the multi-state model parameters given the observed data, $\mathcal{D}_n$.

3.5 CONCLUSIONS

In this chapter we extended the standard reliability methodology presented in Chapter 2. We presented the general form of left-truncated and right-censored data under the two-state model. The two-state model is a simple multi-state model with two states and one transition between those states. The two-state model describes standard failure age data, introduced in Chapter 2, where units start in a working state and eventually transition to a terminal (failed) state. We then derived the likelihood and the DRLs under the two-state model. Next, we introduced the illness-death model, an extension of the two-state model. The illness-death model is a multi-state model with three states and describes semi-competing risks data with two competing risks, an extension of standard failure age data, where units begin in a working state and are subject to a nonterminal event and a terminal event (failure). We presented the general form of left-truncated and right-censored data and derived the likelihood and the DRLs under the illness-death model. Finally, we introduced a four-state multi-state model, an extension of the illness-death model. The multi-state model describes semi-competing risks data with three competing risks, where units begin in a working state and are subject to two nonterminal events and a terminal event (failure). We presented the general form of left-truncated and right-censored data and derived the likelihood and the DRLs under the multi-state model. In Chapter 5, we model the failure ages and survival probabilities of hard drives using the two-state model, the illness-death model, and the multi-state model. The DRLs obtained under the multi-state model are used to examine the impact of critical attributes on hard drive failure ages and survival probabilities.

Chapter 4

Measuring predictive performance

The assessment of the predictive performance of time-to-event models has received a lot of attention in the statistical literature. In general, the developed methodology has focused on calibration, i.e., how well the model predicts the observed data (Schemper and Henderson, 2000; Gerds and Schumacher, 2006), or discrimination, i.e., how well the model can discriminate between units that had the event of interest and units that did not (Harrell Jr et al., 1996; Proust-Lima and Taylor, 2009). In this chapter we present discrimination and calibration measures, in the presence of right-censoring, to assess model performance, where the event of interest is the terminal event (failure). In Chapter 5, we use the calibration and discrimination measures presented in this chapter to compare how well the two-state model, the illness-death model, and the multi-state model, presented in Chapter 3, can predict hard drive failure ages, and how well each model can discriminate between drives that fail and drives that do not fail, within a forecast horizon of interest. Our motivating example concerns a large dataset of hard drives, from data backup company Backblaze (Backblaze, 2022a), that is subject to left-truncation and right-censoring.

4.1 Introduction

To assess the performance of dynamic predictions, we use the time-dependent area under the receiver operating characteristic curve (AUC) (Rizopoulos et al., 2017), as a discriminative measure, to compare how well the two-state model, the illness-death model, and the multi-state model can discriminate between drives that fail and drives that do not fail; and the expected prediction error (PE) (Henderson et al., 2002; Rizopoulos et al., 2017), as a measure of model calibration, to compare how well each model can predict hard drive failure ages.

In this chapter, we first describe the AUC and the PE under the illness-death model. Then, we provide adaptions for the two-state model and the multi-state model. We assume an illness-death model for each individual unit of the form shown in Figure 4.1.

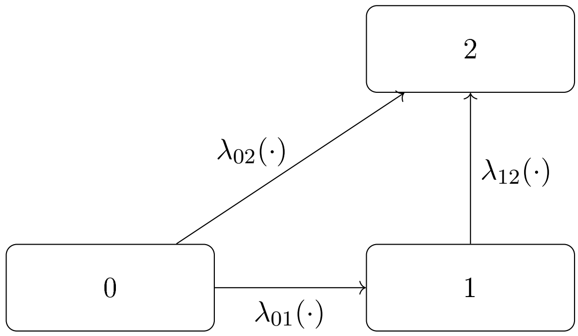

Figure 4.1: An illness-death model for semi-competing risks data.

Define the survival probability

$$\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) = \begin{cases} \Pr(T_{i2} \geq \tau_i + s \mid T_{i2} > \tau_i, z(\tau_i) = 0, \mathcal{D}_n), \\ \text{if unit } i \text{ is in the healthy state at age } \tau_i, \\ \Pr(T_{i2} \geq \tau_i + s \mid T_{i2} > \tau_i, z(\tau_i) = 1, \mathcal{D}_n), \\ \text{if unit } i \text{ is in the critical state at age } \tau_i, \end{cases} \tag{4.1.1}$$

where τ_i is the age of unit i at time τ, $\tau \geq 0$ is the time elapsed since the beginning of the study, $z(\tau_i)$ is the state of unit i at age τ_i, $s > 0$ is the forecast horizon of interest, T_{i2} is the true failure age of unit i, $\mathcal{D}_n$ is the observed data used to obtain the posterior distribution of the illness-death model parameters, and the relevant expressions are given in Equations (3.3.9) and (3.3.10), respectively. The time elapsed since the beginning of the study, τ, will be referred to as the calendar time from herein. The quantity $\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$ is the probability that unit i does not fail by age $\tau_i + s$ conditional on being in state $z(\tau_i)$ at age τ_i. In other words, the quantity $\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$ is the probability that unit i does not fail by calendar time $\tau + s$ conditional on being in state $z(\tau_i)$ at time τ.

4.2 Time-dependent area under the receiver operating characteristic curve (AUC)

The AUC is a measure of discrimination of a model. At calendar time τ, unit i is defined as a case if $\tilde{T}_{i2} \in (\tau, \tau + s]$ and a control if $\tilde{T}_{i2} > \tau + s$, where $\tilde{T}_{i2}$ is the true failure time of unit i. In other words, unit i is a case if it fails within the forecast horizon $(\tau, \tau + s]$ and a control if it survives the forecast horizon. The AUC measures our model's ability to distinguish between a case and a control. Following Rizopoulos et al. (2017), consider at calendar time τ a pair of randomly chosen units, (i, j), for which unit i is a case and unit j is a control. The AUC is given by

$$\begin{aligned}\text{AUC}(\tau, s) = \Pr\big[&\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n) \\ &\mid \{\tilde{T}_{i2} \in (\tau, \tau + s]\} \cap \{\tilde{T}_{j2} > \tau + s\}\big],\end{aligned} \tag{4.2.1}$$

where $\tilde{T}_{i2}$ and $\tilde{T}_{j2}$ are the failure times for units i and j, respectively, and T_{i2} and T_{j2} are the failure ages for units i and j, respectively.

If unit i fails within the forecast horizon and unit j is in a working state after the forecast horizon, we would expect the assumed model to provide a higher probability of surviving the forecast horizon to the unit that survives (unit j) compared to the unit that fails (unit i).

If there are no right-censored failure ages within the forecast horizon, the AUC can be estimated by

$$\begin{aligned}\hat{\text{AUC}}(\tau, s) = \frac{1}{n_{\text{fail}}(\tau) n_{\text{surv}}(\tau)} \sum_{i=1}^{n_{\text{fail}}(\tau)} \sum_{j=1}^{n_{\text{surv}}(\tau)} \mathbb{E}_{\rho_i(\cdot), \rho_j(\cdot)} \Big[I\big\{&\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \\ &\rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n)\big\}\Big],\end{aligned} \tag{4.2.2}$$

where $n_{\text{fail}}(\tau)$ is the number of units that fail within the forecast horizon $(\tau, \tau + s]$ and $n_{\text{surv}}(\tau)$ is the number of units that survive the forecast horizon and the expectation is taken with respect to the posterior predictive distributions of $\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$ and $\rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n)$. If $\hat{\text{AUC}}(\tau, s) = 1$, then our model always gives a higher probability of surviving to units that survive compared to units that fail.

Since many unit failure ages are right-censored within the forecast horizon, estimation of the AUC is based on counting the concordant pairs of units by appropriately distinguishing between the comparable and the "partially-comparable" (due to censoring) pairs of units at calendar time τ. More specifically, the pairs of units that are comparable are given by the set

$$\Omega_{ij}^{(1)}(\tau, s) = \big[\{\tilde{Y}_{i2} \in (\tau, \tau + s]\} \cap \{\delta_{i2} = 1\}\big] \cap \{\tilde{Y}_{j2} > \tau + s\}, \tag{4.2.3}$$

for $i, j = 1, \ldots, n(\tau)$, with $i \neq j$, $n(\tau)$ is the number of units in a working state at calendar time τ, and $\tilde{Y}_{i2}$ and $\tilde{Y}_{j2}$ are the observed failure times for units i and j, respectively. In words, $\Omega_{ij}^{(1)}(\tau, s)$, for $i, j = 1, \ldots, n(\tau)$, with $i \neq j$, is the set of all pairs of units, such that unit i is observed to fail within the forecast horizon (a case), and unit j is observed to survive the forecast horizon (a control). The remaining pairs of units which, due to censoring, cannot be directly compared, are given by the following sets:

$$\begin{aligned}
\Omega_{ij}^{(2)}(\tau, s) &= \big[\{\tilde{Y}_{i2} \in (\tau, \tau + s]\} \cap \{\delta_{i2} = 0\}\big] \cap \{\tilde{Y}_{j2} > \tau + s\}, \\
\Omega_{ij}^{(3)}(\tau, s) &= \big[\{\tilde{Y}_{i2} \in (\tau, \tau + s]\} \cap \{\delta_{i2} = 1\}\big] \cap \big[\{\tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s\} \cap \{\delta_{j2} = 0\}\big], \\
\Omega_{ij}^{(4)}(\tau, s) &= \big[\{\tilde{Y}_{i2} \in (\tau, \tau + s]\} \cap \{\delta_{i2} = 0\}\big] \cap \big[\{\tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s\} \cap \{\delta_{j2} = 0\}\big].
\end{aligned} \tag{4.2.4}$$

In order for a pair of units (i, j) to be directly comparable, unit i must be observed to fail within the forecast horizon and unit j must be observed to survive the forecast horizon; we need a case and a control.

The observed failure time of unit i in $\Omega_{ij}^{(2)}(\tau, s)$ is right-censored. Since the true failure time of unit i is not observed, we do not know if unit i fails within the forecast horizon or not. Thus, units in $\Omega_{ij}^{(2)}(\tau, s)$ are not directly comparable since we cannot directly identify a case and a control. Similarly, the observed failure time for unit j in $\Omega_{ij}^{(3)}(\tau, s)$ is right-censored such that $\tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$. Therefore, we do not know if unit j survives the forecast horizon. Thus, units in $\Omega_{ij}^{(3)}(\tau, s)$ are not directly comparable. Finally, the observed failure times of both units in $\Omega_{ij}^{(4)}(\tau, s)$ are right-censored, such that $\tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$. Thus, units in $\Omega_{ij}^{(4)}(\tau, s)$ are not directly comparable.

The partially-comparable units contribute to the overall AUC after being appropriately weighted with the probability of being comparable

$$\begin{aligned}
\nu_i^{(2)}(\tau, s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n) &= 1 - \rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n), \\
\nu_j^{(3)}(\tau, s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n) &= \rho_j(\tau_j + s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n), \\
\nu_{ij}^{(4)}(\tau, s \mid Y_{i2}, z(Y_{i2}), Y_{j2}, z(Y_{j2}), \mathcal{D}_n) &= \big(1 - \rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)\big) \\
&\quad \times \rho_j(\tau_j + s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n),
\end{aligned} \tag{4.2.5}$$

where Y_{i2} and Y_{j2} are the observed failure ages for units i and j, respectively. In words, $\nu_i^{(2)}(\tau, s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)$ is the probability unit i fails within the forecast horizon, conditional on surviving until age Y_{i2}. In other words, $\nu_i^{(2)}(\tau, s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)$ is the probability that a pair of units in $\Omega_{ij}^{(2)}(\tau, s)$ is comparable. Similarly, $\nu_j^{(3)}(\tau, s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n)$ is the probability that unit j survives the forecast horizon, conditional on surviving until age Y_{j2}. In other words, $\nu_j^{(3)}(\tau, s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n)$ is the probability that a pair of units in $\Omega_{ij}^{(3)}(\tau, s)$ is comparable. Finally, $\nu_{ij}^{(4)}(\tau, s \mid Y_{i2}, z(Y_{i2}), Y_{j2}, z(Y_{j2}), \mathcal{D}_n)$ is the probability that unit i fails within the forecast horizon, conditional on surviving until age Y_{i2}, and unit j survives the forecast horizon, conditional on surviving until age Y_{j2}. In other words, $\nu_{ij}^{(4)}(\tau, s \mid Y_{i2}, z(Y_{i2}), Y_{j2}, z(Y_{j2}), \mathcal{D}_n)$ is the probability that a pair of units in $\Omega_{ij}^{(4)}(\tau, s)$ is comparable.

The time-dependent area under the receiver operating characteristic curve at time τ, as proposed by Rizopoulos et al. (2017), is given by

$$\hat{\text{AUC}}(\tau, s) = \frac{\kappa_1(\tau, s) + \kappa_2(\tau, s) + \kappa_3(\tau, s) + \kappa_4(\tau, s)}{\zeta_1(\tau, s) + \zeta_2(\tau, s) + \zeta_3(\tau, s) + \zeta_4(\tau, s)}, \tag{4.2.6}$$

where

$$\begin{aligned}
\kappa_1(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} \mathbb{E}_{\rho_i(\cdot), \rho_j(\cdot)} \Big[I\big\{\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n)\big\}\Big] \\
&\quad \times I\{\Omega_{ij}^{(1)}(\tau, s)\}, \\
\kappa_2(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} \mathbb{E}_{\rho_i(\cdot), \rho_j(\cdot)} \Big[I\big\{\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n)\big\}\Big] \\
&\quad \times I\{\Omega_{ij}^{(2)}(\tau, s)\}\nu_i^{(2)}, \\
\kappa_3(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} \mathbb{E}_{\rho_i(\cdot), \rho_j(\cdot)} \Big[I\big\{\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n)\big\}\Big] \\
&\quad \times I\{\Omega_{ij}^{(3)}(\tau, s)\}\nu_j^{(3)}, \\
\kappa_4(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} \mathbb{E}_{\rho_i(\cdot), \rho_j(\cdot)} \Big[I\big\{\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n)\big\}\Big] \\
&\quad \times I\{\Omega_{ij}^{(4)}(\tau, s)\}\nu_{ij}^{(4)},
\end{aligned} \tag{4.2.7}$$

and

$$
\begin{aligned}
\zeta_1(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} I\{\Omega_{ij}^{(1)}(\tau, s)\}, \\
\zeta_2(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} I\{\Omega_{ij}^{(2)}(\tau, s)\}\nu_i^{(2)}, \\
\zeta_3(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} I\{\Omega_{ij}^{(3)}(\tau, s)\}\nu_j^{(3)}, \\
\zeta_4(\tau, s) &= \sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} I\{\Omega_{ij}^{(4)}(\tau, s)\}\nu_{ij}^{(4)}.
\end{aligned}
\tag{4.2.8}
$$

In addition, I_A is an indicator function, which is equal to 1 if event A occurs and 0 otherwise.

For all pairs of units (i, j) in the set $\Omega_{ij}^{(1)}(\tau, s)$, $\kappa_1(\tau, s)$ is the number of times the illness-death model gives the control a higher probability of surviving the forecast horizon compared to the case. Moreover, $\zeta_1(\tau, s)$ is the number of pairs of units in the set $\Omega_{ij}^{(1)}(\tau, s)$. If all pairs of units were comparable, i.e., if all pairs of units were contained in the set $\Omega_{ij}^{(1)}(\tau, s)$, then Equation (4.2.6) reduces to

$$
\hat{\text{AUC}}(\tau, s) = \frac{\kappa_1(\tau, s)}{\sum_{i=1}^{n(\tau)} \sum_{j=1, j\neq i}^{n(\tau)} I\{\Omega_{ij}^{(1)}(\tau, s)\}},
\tag{4.2.9}
$$

which is equivalent to Equation (4.2.2).

Similar definitions can be made for $\kappa_l(\tau, s)$ and $\zeta_l(\tau, s)$, for $l = 2, 3, 4$, but each term in $\kappa_l(\tau, s)$ and $\zeta_l(\tau, s)$, for $l = 2, 3, 4$, is weighted by the probability of the pair of units, (i, j), being comparable, i.e., the probability that unit i fails within the forecast horizon and unit j survives the forecast horizon. $\hat{\text{AUC}}(\tau, s)$, given by Equation (4.2.6), accounts for right-censored observations in the forecast horizon, and is interpreted as the probability the illness-death model gives the control a higher probability of surviving the forecast horizon compared to the case.

Illustration of the AUC in the presence of right-censoring

There are four sets to consider, in the presence of right-censoring, when calculating the AUC. The AUC assesses whether a model can distinguish between units that fail and units that survive a forecast horizon of interest. For a pair of units to be directly comparable, unit i needs to fail within the forecast horizon and unit j needs to survive the forecast horizon. We need a case and a control.

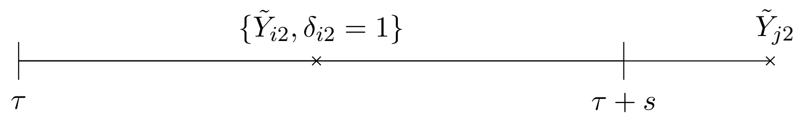

Figure 4.2: A pair of comparable units in $\Omega_{ij}^{(1)}(\tau, s)$. Unit i failed within the forecast horizon and unit j survived the forecast horizon.

Figure 4.2 depicts an example pair of comparable units in the set $\Omega_{ij}^{(1)}(\tau, s)$. Pairs of units in the set $\Omega_{ij}^{(1)}(\tau, s)$ are directly comparable under the definition of the AUC.

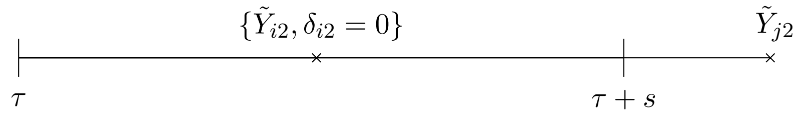

Figure 4.3: A pair of "partially comparable" units in $\Omega_{ij}^{(2)}(\tau, s)$. Unit i is right-censored in the forecast horizon and unit j survived the forecast horizon.

Figure 4.3 depicts an example pair of "partially comparable" units in the set $\Omega_{ij}^{(2)}(\tau, s)$. In this scenario, unit j survives the forecast horizon, but unit i is right-censored within the forecast horizon at $\tau < \tilde{Y}_{i2} \leq \tau + s$. If we observed unit i until failure, we would know if unit i failed within the forecast horizon or if unit i survived the forecast horizon. In the former case, the pair (i, j) would be comparable under the definition of the AUC, whereas in the latter case, the pair (i, j) would not be comparable. The contribution of this pair of units is weighted by the probability that the pair is comparable, which is the probability that unit i fails within the forecast horizon conditional on surviving until age Y_{i2}, $1 - \rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)$.

In the limit that $1 - \rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)$ tends to 0, the pair (i, j) does not contribute to the AUC; in this limit, the probability of unit i failing within the forecast horizon is zero, and hence the pair of units is not comparable. In the limit that $1 - \rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)$ tends to 1, the pair (i, j) contributes to the AUC analogously to a directly comparable pair of units; in this limit, the probability of unit i failing within the forecast horizon is one, and hence the pair of units is directly comparable.

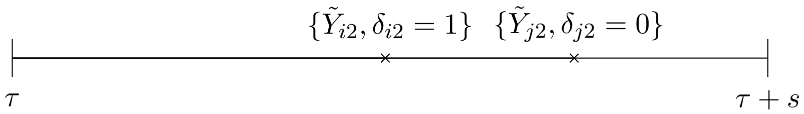

Figure 4.4: A pair of "partially comparable" units in $\Omega_{ij}^{(3)}(\tau, s)$. Unit i failed within the forecast horizon and unit j is right-censored in the forecast horizon, such that $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$.

Figure 4.4 depicts an example pair of "partially comparable" units in the set $\Omega_{ij}^{(3)}(\tau, s)$. In this scenario, unit i fails within the forecast horizon, but unit j is right-censored within the forecast horizon at $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$. The contribution of this pair of units is weighted by the probability that the pair is comparable, which is the probability that unit j survives the forecast horizon conditional on surviving until age Y_{j2}, $\rho_j(\tau_j + s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n)$.

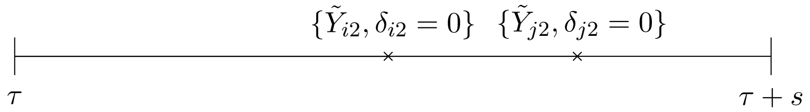

Figure 4.5: A pair of "partially comparable" units in $\Omega_{ij}^{(4)}(\tau, s)$. Units i and j are both right-censored in the forecast horizon, such that $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$.

Figure 4.5 depicts an example pair of "partially comparable" units in the set $\Omega_{ij}^{(4)}(\tau, s)$. In this scenario, units i and j are both right-censored within the forecast horizon, such that $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$. The contribution of this pair of units is weighted by the probability that the pair is comparable, which is the probability that unit i fails within the forecast horizon, conditional on surviving until age Y_{i2}, and unit j survives the forecast horizon, conditional on surviving until age Y_{j2}, $\big(1 - \rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)\big) \times \rho_j(\tau_j + s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n)$.

We note that in the set $\Omega_{ij}^{(3)}(\tau, s)$ we require $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$. In other words, the right-censoring time for unit j must be larger than the failure time for unit i. Consider the example depicted in Figure 4.6, where unit i fails within the forecast horizon and unit j is right-censored, such that $\tau < \tilde{Y}_{j2} < \tilde{Y}_{i2} \leq \tau + s$. One could ask, should this pair be considered and the contribution for this pair weighted by the probability that unit j survives the forecast horizon conditional on surviving until age Y_{j2}, $\rho_j(\tau_j + s \mid Y_{j2}, z(Y_{j2}), \mathcal{D}_n)$? After all, if unit j survives the forecast horizon, then we have a comparable pair (one unit failing within the forecast horizon and one unit surviving the forecast horizon).

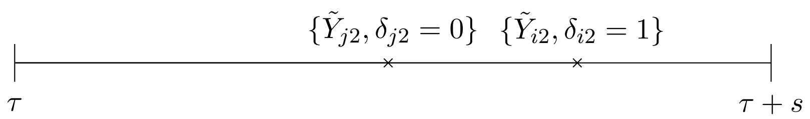

Figure 4.6: A forecast horizon with a pair of units (i, j). Unit i failed within the forecast horizon and unit j is right-censored in the forecast horizon, such that $\tau < \tilde{Y}_{j2} < \tilde{Y}_{i2} \leq \tau + s$.

The restriction that the right-censoring time for unit j must be larger than the failure time for unit i is in place to prevent pairs of units from switching the "failed" unit to the "survived" unit over time. The AUC is defined for all time points τ. To assess the performance of a model over time, the AUC should be calculated at times $\tau = 0, \tau_1, \tau_2, \ldots, \tau_n$. Now, suppose we ignore this restriction and consider the example depicted in Figure 4.7.

Figure 4.7 depicts a possible scenario when we remove the restriction that $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$. The top plot in Figure 4.7 depicts the forecast horizon at calendar time $\tau_1 < \tau_2$ and the bottom plot in Figure 4.7 depicts the forecast horizon at calendar time τ_2. From the point of view of τ_1, unit j survives the forecast horizon and unit i is right-censored within the forecast horizon, hence this pair is comparable if unit i fails within the forecast horizon, conditional on surviving until age Y_{i2}. Hence, we are trying to discriminate between unit j, a unit that survived the horizon, and unit i, a unit that will potentially fail within the forecast horizon. From the point of view of τ_2, unit j fails within the forecast horizon and unit i is right-censored within the forecast horizon, hence this pair is comparable if unit i survives the forecast horizon, conditional on surviving until age Y_{i2}. Hence, we are trying to discriminate between unit j, a unit that failed and unit i, a unit that will potentially survive the forecast horizon.

We see that removing the restriction that $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$ results in the switching of definitions of "survived" and "failed" units over time for the same pair of units. At time τ_1, the AUC is "rewarded" by the contribution $1 - \rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)$ (assuming $\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n)$). At time τ_2, the AUC is "rewarded" by the contribution $\rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n)$ (assuming $\rho_j(\tau_j + s \mid \tau_j, z(\tau_j), \mathcal{D}_n) < \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$). In other words, at τ_1 the AUC is larger if $\rho_i(\tau_i+s \mid \tau_i, z(\tau_i), \mathcal{D}_n) < \rho_j(\tau_j+s \mid \tau_j, z(\tau_j), \mathcal{D}_n)$ and at τ_2 the AUC is larger if $\rho_i(\tau_i+s \mid \tau_i, z(\tau_i), \mathcal{D}_n) > \rho_j(\tau_j+s \mid \tau_j, z(\tau_j), \mathcal{D}_n)$.

Consequently, we require that $\tau < \tilde{Y}_{i2} < \tilde{Y}_{j2} \leq \tau + s$ to prevent this unintuitive rewarding of models. Similar arguments hold for the restriction for the set $\Omega^{(4)}_{ij}(\tau, s)$.

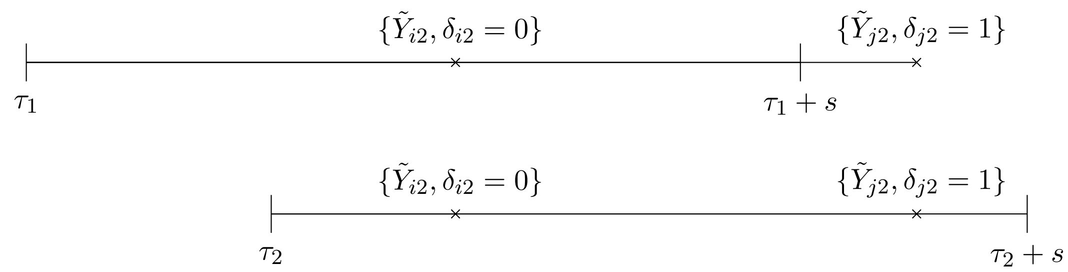

Figure 4.7: Two consecutive forecast horizons with a pair of units (i, j).

4.3 Time-dependent expected predicted error (PE)

The expected predicted error in predicting future failures can be used to assess the accuracy of dynamic predictions Rizopoulos et al. (2017). As for the AUC we focus our interest in predicting failures that occur by calendar time $\tau + s > \tau$ given the information available up to calendar time τ. The expected predicted error for a working unit at time τ is given by

$$\mathrm{PE}(\tau + s \mid \tau) = \mathbb{E}_{\rho_i(\cdot)}\Big[\mathbb{E}_{T_{i2}}\left[L\{N_i(\tau + s) - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)\}\right]\Big], \qquad (4.3.1)$$

where $N_i(\tau + s) = I(T_{i2} > \tau_i + s) = I(\tilde{T}_{i2} > \tau + s)$ denotes the survival status of unit i at calendar time $\tau + s$, where $N_i(\tau + s) = 1$ if unit i survives the forecast horizon and $N_i(\tau + s) = 0$ if unit i fails within the forecast horizon; $L(\cdot)$ denotes a loss function, such as the absolute or square loss; the inner expectation is taken with respect to the posterior predictive distribution of T_{i2} and the outer expectation is taken with respect to the posterior predictive distribution of $\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$.

The PE is a calibration measure that measures how well a model predicts failures. The smaller the value of $N_i(\tau + s) - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$, the smaller the PE, or the more accurate the model is. That is, for units that survive, the closer $\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$ is to 1 (i.e., the higher the probability of survival), the smaller $N_i(\tau + s) - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$ will be. We want our model to give a high probability of surviving the forecast horizon of interest to units that survive. In contrast, we want our model to give a low probability of surviving the forecast horizon to units that fail within the forecast horizon. For units that fail, the closer $\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$ is to 0 (i.e., the lower the survival probability), the smaller $N_i(\tau + s) - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)$ will be. An accurate model will give high probabilities of surviving to units that survive and low probabilities of surviving to units that fail.

An estimate of $\text{PE}(\tau + s \mid \tau)$ that accounts for censoring has been proposed by Henderson et al. (2002) and is given by

$$
\begin{aligned}
&\hat{\text{PE}}(\tau + s \mid \tau) = \frac{1}{n(\tau)} \sum_{i=1}^{n(\tau)} \hat{\text{PE}}_i(\tau + s \mid \tau) = \\
&\quad \frac{1}{n(\tau)} \sum_{i=1}^{n(\tau)} I\{\tilde{Y}_{i2} > \tau + s\} \, \mathbb{E}_{\rho_i(\cdot)} \Big[L(1 - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)) \Big] \\
&\quad + \delta_i I\{\tilde{Y}_{i2} < \tau + s\} \, \mathbb{E}_{\rho_i(\cdot)} \Big[L(0 - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)) \Big] \\
&\quad + (1 - \delta_i) I\{\tilde{Y}_{i2} < \tau + s\} \Big\{ \mathbb{E}_{\rho_i(\cdot)} \Big[\rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n) \Big] \\
&\quad \times \mathbb{E}_{\rho_i(\cdot)} \Big[L(1 - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)) \Big] \\
&\quad + \Big(1 - \mathbb{E}_{\rho_i(\cdot)} \Big[\rho_i(\tau_i + s \mid Y_{i2}, z(Y_{i2}), \mathcal{D}_n) \Big] \Big) \mathbb{E}_{\rho_i(\cdot)} \Big[L(0 - \rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n)) \Big] \Big\},
\end{aligned} \tag{4.3.2}
$$

where $n(\tau)$ denotes the number of units that are in a working state at calendar time τ. The first term in Equation (4.3.2) is the contribution from a unit that is in a working state after time $\tau + s$, in other words, the contribution from a unit that survives the forecast horizon; the second term is the contribution from a unit that fails within the forecast horizon; the third and fourth terms are the contributions from a unit that is censored in the forecast horizon $(\tau, \tau + s)$. Using the information up to time τ, $\mathrm{PE}(\tau + s \mid \tau)$ measures the predictive accuracy at calendar time τ.

4.4 Adaptions for the two-state model and the multi-state model

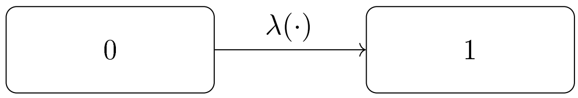

Figure 4.8: A two-state model.

In this chapter we have described the AUC and the PE under the illness-death model shown in Figure 4.1. The definitions of the AUC and the PE can be adapted for the two-state model shown in Figure 4.8 and the multi-state model shown in Figure 4.9.

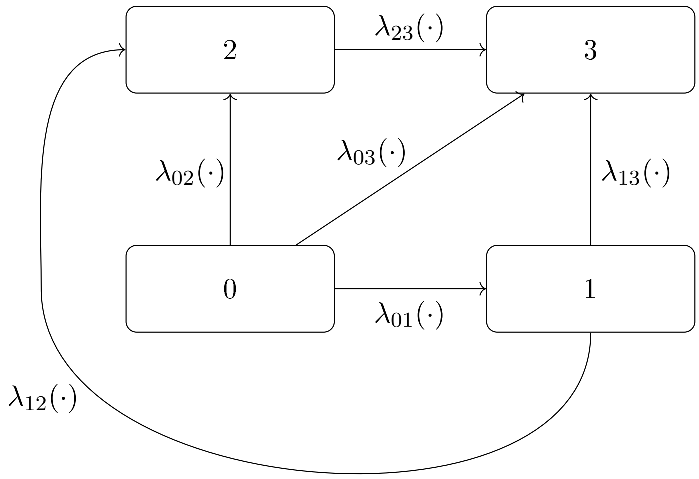

Figure 4.9: A multi-state model with four states.

Under the two-state model we define

$$\rho_i(\tau_i + s \mid \tau_i, \mathcal{D}_n) = \Pr(T_i \geq \tau_i + s \mid T_i > \tau_i, \mathcal{D}_n), \tag{4.4.1}$$

where τ_i is the age of unit i at time τ, T_i is the true failure age of unit i, $\mathcal{D}_n$ is the observed data used to obtain the posterior distribution of the two-state model parameters, and the relevant expression is given in Equation (3.2.3). Under the two-state model, in the definitions of the AUC and the PE, given in Sections 4.2 and 4.3, respectively, we replace the $\{i2\}$ and $\{j2\}$ subscripts with $\{i\}$ and $\{j\}$ subscripts, respectively. For example, under the two-state model, we replace $\tilde{Y}_{i2}$ and $\tilde{Y}_{j2}$ with $\tilde{Y}_i$ and $\tilde{Y}_j$, respectively.

Similarly, under the multi-state model, we define

$$\rho_i(\tau_i + s \mid \tau_i, z(\tau_i), \mathcal{D}_n) = \begin{cases} \Pr(T_{i3} \geq \tau_i + s \mid T_{i3} > \tau_i, z(\tau_i) = 0, \mathcal{D}_n), \\ \text{if unit } i \text{ is in the healthy state at age } \tau_i, \\ \Pr(T_{i3} \geq \tau_i + s \mid T_{i3} > \tau_i, z(\tau_i) = 1, \mathcal{D}_n), \\ \text{if unit } i \text{ is in the critical 1 state at age } \tau_i, \\ \Pr(T_{i3} \geq \tau_i + s \mid T_{i3} > \tau_i, z(\tau_i) = 2, \mathcal{D}_n), \\ \text{if unit } i \text{ is in the critical 2 state at age } \tau_i, \end{cases} \quad (4.4.2)$$

where τ_i is the age of unit i at time τ, $z(\tau_i)$ is the state of unit i at age τ_i, T_{i3} is the true failure age of unit i, $\mathcal{D}_n$ is the observed data used to obtain the posterior distribution of the multi-state model parameters, and the relevant expressions are given in Equations (3.4.15) - (3.4.17), respectively. Under the multi-state model, in the definitions of the AUC and the PE, given in Sections 4.2 and 4.3, respectively, we replace the $\{i2\}$ and $\{j2\}$ subscripts with $\{i3\}$ and $\{j3\}$ subscripts, respectively. For example, under the multi-state model, we replace $\tilde{Y}_{i2}$ and $\tilde{Y}_{j2}$ with $\tilde{Y}_{i3}$ and $\tilde{Y}_{j3}$, respectively.

4.5 Conclusions

In this chapter we presented discrimination and calibration measures, in the presence of right-censoring, to assess model performance, where the event of interest is the terminal event (failure). We introduced the time-dependent area under the receiver operating characteristic curve (AUC) to assess the discriminative ability of a statistical model over time. We described the comparable and "partially comparable" (due to censoring) pairs of units and described how the standard AUC measure is adjusted in the presence of right-censoring. We then introduced the expected predicted error (PE) to assess the accuracy of predicting future failures. In Chapter 5, we use the calibration and discrimination measures presented in this chapter to compare how well the two-state model, the illness-death model, and the multi-state model, presented in Chapter 3, can predict hard drive failure ages, and how well each model can discriminate between drives that fail and drives that do not fail, within a forecast horizon of interest. Our motivating example concerns a large dataset of hard drives, from data backup company Backblaze (Backblaze, 2022a), that is subject to left-truncation and right-censoring.

Chapter 5

Multi-state models for left-truncated and right-censored data

In this chapter, we provide a novel way to model the failure ages of hard drives using data collected by SMART. The proposed models enable us to specifically identify the impact of critical attributes on hard drive survival probabilities and failure ages.

We apply the three multi-state models described in Chapter 3 to a large dataset of hard drives, from data backup company Backblaze, that is subject to left-truncation and right-censoring. We propose transient states, named the critical states, for hard drives using data collected by SMART and model the resulting semi-competing risks data using multi-state models. The proposed multi-state models provide a coherent and novel way to model failure ages of hard drives and allow us to statistically examine the impact of critical attributes on hard drive failure ages. We illustrate how multi-state models can be used to obtain the DRLs for hard drives using the current state of a drive, and compare our results to previous work by Mittman et al. (2019) using the methods presented in Chapter 4.

5.1 Multi-state models

In this section, we define critical attributes and critical states for hard drives using data collected by SMART. We first remind the reader of the Backblaze hard drive dataset.

5.1.1 Backblaze hard drive data

Backblaze is a company that offers cloud backup storage to protect against data loss. Since 2013, it has been collecting daily operational data, using SMART, on all of the hard drives operating at its facility. Every quarter Backblaze makes its hard drive data publicly available through its website (Backblaze, 2022a).

As of the first quarter of 2022, Backblaze was collecting and reporting data on 150 different drive models. Some drive models have been running since 2013 or before, while others were added at a later date. The number of drives and the number of failed drives vary by drive model; some models have no recorded failures.

Each day SMART takes a snapshot of each operational drive at Backblaze. This snapshot includes basic drive information, along with the SMART attributes reported by that drive. Consequently, for each drive, we have a time series for each recorded attribute. Each day, for each operational drive at Backblaze, SMART records: the date, the drive serial number, the drive model, the capacity of the drive in bytes, an indicator denoting if the drive failed that day, and multiple SMART attributes.

Over time, the reporting technology is upgraded, and as a result, the number of recorded attributes changes over time. For example, between 2013 and 2014, SMART provided 80 columns of data per day, alongside basic drive information, for each hard drive. These columns correspond to the raw and normalised values of 40 different SMART attributes. From 2018 onwards, SMART provided 124 columns of data per day, alongside basic drive information, for each drive (corresponding to the raw and normalised values of 62 different SMART attributes).

The Backblaze hard drive failure ages are left-truncated and right-censored. When an observation is left-truncated, it would not have been observed if it had occurred prior to a particular time. Many Backblaze hard drives have a history prior to data collection, and hard drives that were in a failed state when data collection commenced are not included in the dataset. Hence, the ages of Backblaze hard drives at failure are left-truncated. In addition, it is rare that all drives in a study are observed until failure. If a drive has not failed when the study ends, it is considered right-censored. Right-censoring puts a lower bound on the failure age. Left-truncation and right-censoring must both be incorporated to avoid biased estimates.

In this chapter, we extend the methods proposed by Mittman et al. (2019) to incorporate the attributes collected by SMART. We compare our results to the results presented in Section 4.2 of Mittman et al. (2019) using the methods presented in Chapter 4. More specifically, to illustrate the multi-state models proposed in this chapter, we present an analysis of drive model 14. Drive model 14 was a hard drive used by Backblaze up to (and including) the last quarter of 2015. Between 2013 and 2016, Backblaze deployed 4707 model 14 drives.

5.1.2 The critical states

5.1.2.1 Attribute selection

Backblaze uses five SMART attributes as a means of helping determine if a drive is going to fail (Rincón et al., 2017; Backblaze, 2022b). Namely, SMART 5, the reallocated sectors count; SMART 187, the reported uncorrectable errors; SMART 188, command timeout; SMART 197, the current pending sector count; and SMART 198, the uncorrectable sector count. When the raw value for at least one of these five attributes is greater than zero, Backblaze has a reason to investigate (Rincón et al., 2017; Backblaze, 2022b).

Rincón et al. (2017) investigated three different machine learning models using Backblaze data. They performed a statistical analysis to reduce the number of SMART attributes to consider and to eliminate irrelevant variables (variables without any relationship to drive failure). Their trend test identified six attributes to be used in the machine learning models. Namely, SMART 196, the reallocation event count, and the five attributes used by Backblaze.

Ma et al. (2015) designed RAIDSHIELD, consisting of PLATE and ARMOR. PLATE monitors individual drive health by tracking the number of reallocated sector counts (SMART 5) and proactively detecting unstable drives. ARMOR utilises joint failure probabilities to quantify and predict how likely a RAID group (multiple hard drives grouped together to decrease the risk of data loss) is to face multiple simultaneous drive failures. The joint failure probabilities depend on the number of reallocated sector counts (SMART 5). Their results show that the accumulation of reallocated sectors is correlated with a higher probability of failure.

Following from Ma et al. (2015), Rincón et al. (2017), and Backblaze (2022b) we consider SMART attributes 5, 187, 188, 197, and 198. From herein these SMART attributes are referred to as critical attributes. We note that SMART 196 is missing for all hard drives in our dataset.

5.1.2.2 Model for critical attributes

A parametric model for each critical attribute process is needed for the purpose of prediction. Figure 5.1 (left) depicts the longitudinal profile of the reallocated sector count (SMART 5) for five example hard drives. Figure 5.1 (right) depicts the longitudinal

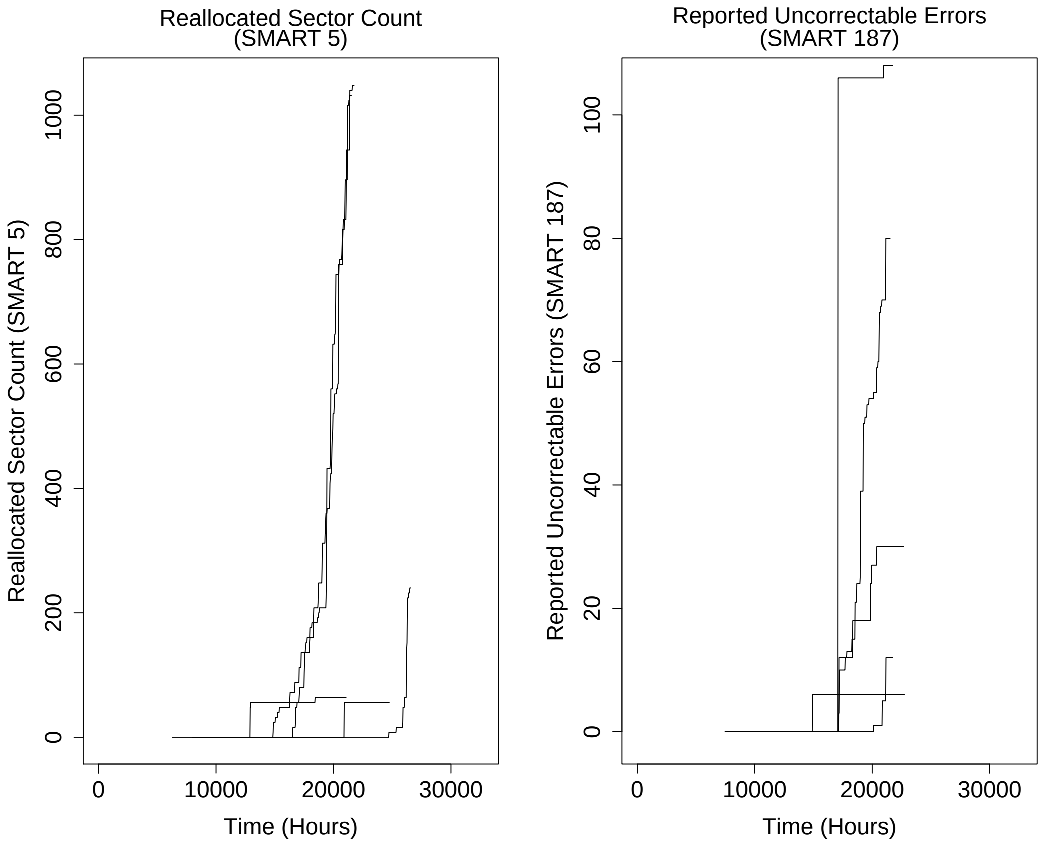

Figure 5.1: Figure (left) depicts the longitudinal profile of the reallocated sector count (SMART 5) for five example hard drives. Figure (right) depicts the longitudinal profile of the reported uncorrectable errors (SMART 187) for five example hard drives.

profile of the reported uncorrectable errors (SMART 187) for five example hard drives. From Figure 5.1, we can see that there is no clear trend in the reallocated sector count over time, nor in the reported uncorrectable errors over time (other than both being non-decreasing).

Table 5.1 provides summary statistics for the five critical attributes for model 14 hard drives. More specifically, Table 5.1 provides, for each critical attribute, quantiles of the time between jumps (nonzero increases), in hours (for hard drives with at least two nonzero values); quantiles of the size of the first nonzero jump (for hard drives with at least one nonzero value); and quantiles of the number of nonzero jumps (for hard drives with at least one nonzero value). The quantiles were obtained using the empirical distributions for each critical attribute.

For drive model 14, the central 99% empirical interval for the time between jumps in reallocated sector count, for hard drives that experience at least two jumps, is $(23, 1360)$ hours with a median of 24 hours; the central 99% empirical interval for the size of the first jump in reallocated sector count, for drives with at least one nonzero jump, is $(8, 34200)$

Critical Attribute	**Time between jumps (hours)**			**Size of jump**			**Number of nonzero jumps**			**Proportion of drives with at least**	
	$\alpha_{0.005}$	$\alpha_{0.5}$	$\alpha_{0.995}$	$\alpha_{0.005}$	$\alpha_{0.5}$	$\alpha_{0.995}$	$\alpha_{0.005}$	$\alpha_{0.5}$	$\alpha_{0.995}$	**1 jump**	**2 jumps**
SMART 5	23.0	24.0	1360	8.00	72.0	34200	1.00	22.0	176	0.253	0.239
SMART 187	23.0	119.0	4310	1.00	6.00	432	1.00	2.00	39.7	0.368	0.236
SMART 188	23.0	720	11800	1.00	4.72×10^{10}	2.10×10^{11}	1.00	2.00	18.3	0.840	0.557
SMART 197	22.8	121	8890	8.00	8.00	65500	1.00	2.00	24.5	0.362	0.216
SMART 198	23.0	120	3930	8.00	8.00	65500	1.00	2.00	22.5	0.275	0.159

Table 5.1: Empirical distribution quantiles for the time between jumps, in hours, the size of each jump, and the number of nonzero jumps, for each critical attribute; alongside the proportion of drives with at least one and at least two jumps.

with a median of 72; and the central 99% empirical interval for the number of nonzero jumps in reallocated sector count, for drives that experienced at least one jump, is $(1, 176)$ with a median of 22 jumps.

Many hard drives do not experience a nonzero jump in reallocated sector count during their lifespan. More specifically, 75% of model 14 hard drives in our dataset do not experience a nonzero jump in reallocated sector count. For hard drives that experience a nonzero jump in reallocated sector count: some hard drives do not exhibit further jumps, some drives have frequent sporadic jumps thereafter, some of the subsequent jumps could be within days or thousands of hours later, and the size of the jumps vary drastically in size. Similar conclusions can be drawn from the other critical attributes. The erratic nature of these poorly understood processes makes it difficult to predict their values over time (with a "reasonable" amount of certainty that the predictions could be useful in a predictive model).

We utilise the probabilities of changes in critical attributes and the age of a hard drive to predict the probability of drive failure over time. We treat changes in critical attributes as drives entering critical states. In this setting, the data collected by SMART can be considered semi-competing risks data. Under this definition of the critical states, we do not need to forecast the process for any critical attribute. Instead, we must forecast the probability of entering the critical states. It is difficult to predict the value of the critical attributes over time, but it is more manageable to obtain the probability of entering the critical states. In Sections 5.2.5 and 5.2.6, we model the semi-competing risks data more formally using the illness-death model and the multi-state model.

5.2 Application to the Backblaze hard drives

In this section, we present an analysis of drive model 14. This drive model is chosen to illustrate how multi-state models can be used to obtain survival probabilities and DRLs for hard drives using the age and state of a drive. We illustrate how the survival probabilities and DRLs allow us to define the impact of critical attributes on hard drive survival probabilities and failure ages. In addition, we assess the performance of the two-state model, the illness-death model, and the multi-state model through a simulation study using the discrimination (AUC) and calibration (PE) measures presented in Chapter 4. First, we introduce a useful alternative parameterisation of the Weibull distribution and propose the generalised limited failure population model.

5.2.1 Weibull distribution and reparameterisation

The Weibull CDF is

$$\Pr(T \leq t \mid \alpha, \beta) = F(t \mid \alpha, \beta) = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right],\ t > 0, \tag{5.2.1}$$

where $\beta > 0$ is the Weibull shape parameter and $\alpha > 0$ is the scale parameter. Because $\log(T)$ has a smallest extreme value distribution, a member of the location-scale family of distributions (Meeker et al., 2022), the Weibull CDF can also be written as

$$\Pr(T \leq t \mid \mu, \sigma) = F(t \mid \mu, \sigma) = \Phi_{\text{SEV}}\left[\frac{\log(t) - \mu}{\sigma}\right],\ t > 0, \tag{5.2.2}$$

where $\Phi_{\text{SEV}}(z) = 1 - \exp[-\exp(z)]$ is the standard smallest extreme value distribution CDF and $\mu = \log(\alpha)$ and $\sigma = 1/\beta$ are, respectively, location and scale parameters for the distribution of $\log(T)$.

Following Mittman et al. (2019), we use an alternative parameterisation where the usual scale parameter α is replaced by the p quantile $t_p = \alpha[-\log(1-p)]^{\sigma}$ (which is also a scale parameter). The p quantile, t_p, is defined as $\Pr(T \leq t_p) = p$. Replacing α in Equation (5.2.1) with $t_p/[-\log(1-p)]^{\sigma}$ and β with $1/\sigma$ gives

$$F(t \mid t_p, \sigma) = 1 - \exp\left[\log(1-p)\left(\frac{t}{t_p}\right)^{1/\sigma}\right],\ t > 0. \tag{5.2.3}$$

The Weibull PDF is

$$f(t \mid t_p, \sigma) = -\frac{\log(1-p)}{\sigma t_p}\left(\frac{t}{t_p}\right)^{1/\sigma - 1} \exp\left[\log(1-p)\left(\frac{t}{t_p}\right)^{1/\sigma}\right],\ t > 0. \tag{5.2.4}$$

The Weibull hazard function is

$$\lambda(t \mid t_p, \sigma) = \frac{f(t \mid t_p, \sigma)}{1 - F(t \mid t_p, \sigma)} = -\frac{\log(1-p)}{\sigma t_p}\left(\frac{t}{t_p}\right)^{1/\sigma - 1},\ t > 0. \tag{5.2.5}$$

There are good reasons for using this parameterisation.

1. Especially with a highly-reliable product, it will be easier to elicit prior information about a quantile (t_p) in the lower tail of the distribution than it will be to elicit prior information about α (approximately the 0.63 quantile). In addition, there is generally available information about the shape parameter, σ, for a given failure mechanism. For example, if the failure is due to a wearout mechanism, then it is known that $\sigma < 1$ (Mittman et al., 2019).

2. Because of heavy censoring in reliability data, the parameters μ and σ will generally be highly correlated and the specification of independent marginal prior distributions would be inappropriate (Mittman et al., 2019). However, t_p and σ, for some appropriately chosen value of p, will be approximately independent, allowing the easier elicitation and specification of independent marginal prior distributions. For example, if a dataset has 10% of units entering a particular state (e.g., the failure state), then choosing $t_{0.05}$ would work well (Mittman et al., 2019).

3. Bayesian MCMC estimation will be better behaved due to reduced correlation between t_p and σ (relative to α and σ or μ and σ) (Mittman et al., 2019).

5.2.2 Generalised limited failure population model

Let F_1 and F_2 be the CDFs of Weibull distributions with parameters (t_{p_1}, σ_1) and (t_{p_2}, σ_2), respectively. The GLFP model of Chan and Meeker (1999) is defined as follows. Let $T \sim \text{GLFP}(\pi, t_{p_1}, \sigma_1, t_{p_2}, \sigma_2)$. Then

$$\Pr(T \le t \mid \boldsymbol{\theta}) = F(t \mid \boldsymbol{\theta}) = 1 - (1 - \pi F_1(t \mid \boldsymbol{\theta}_1))(1 - F_2(t \mid \boldsymbol{\theta}_2)),\ t > 0, 0 < \pi < 1, \tag{5.2.6}$$

where $\boldsymbol{\theta} = (\pi, t_{p_1}, \sigma_1, t_{p_2}, \sigma_2)$, $\boldsymbol{\theta}_1 = (t_{p_1}, \sigma_1)$, and $\boldsymbol{\theta}_2 = (t_{p_2}, \sigma_2)$.

The GLFP model can be understood as a mixture model with a binary latent variable, $\zeta_i \sim \text{Bernoulli}(\pi)$. ζ_i is an indicator for whether drive i is defective or not (i.e., susceptible to an early failure), and π is the probability that drive i is defective. Here $F_1(t)$ is the CDF for the early failures, and $F_2(t)$ is the CDF for the wear-out failures. Expressed conditional on ζ_i,

$$\begin{aligned}\Pr(T \le t \mid \zeta_i = 1, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) &= 1 - (1 - F_1(t \mid \boldsymbol{\theta}_1))(1 - F_2(t \mid \boldsymbol{\theta}_2)),\\ \Pr(T \le t \mid \zeta_i = 0, \boldsymbol{\theta}_2) &= F_2(t \mid \boldsymbol{\theta}_2).\end{aligned} \tag{5.2.7}$$

The parameter π represents the proportion of drives susceptible to early failure, and hence susceptible to both failure modes.

The GLFP PDF is

$$f(t \mid \boldsymbol{\theta}) = \pi f_1(t \mid \boldsymbol{\theta}_1)(1 - F_2(t \mid \boldsymbol{\theta}_2)) + f_2(t \mid \boldsymbol{\theta}_2)(1 - \pi F_1(t \mid \boldsymbol{\theta}_1)). \tag{5.2.8}$$

Furthermore, the GLFP hazard function at t is given by

$$\lambda(t \mid \boldsymbol{\theta}) = \frac{f(t \mid \boldsymbol{\theta})}{1 - F(t \mid \boldsymbol{\theta})}, \tag{5.2.9}$$

where $f(t)$ is given by Equation (5.2.8) and $F(t)$ is given by Equation (5.2.6).

Hard drives, like other engineered products, have a relatively high rate of early failure due to manufacturing defects. After this “burn-in” period, failure rates stabilise, once the majority of defective units have failed. Finally, after prolonged use, rates of failure increase due to wear-out. The GLFP model is used to capture both failure modes of hard drives.

5.2.3 Drive model 14

Backblaze deployed 4707 model 14 drives. Three of these drives had only one observation and were removed from the dataset. This leaves 4704 hard drives. By the end of 2015, 1707 of these drives had failed. Consequently, 2997 drives are right-censored.

Figure 5.3 depicts the illness-death model used to model hard drive failure ages, with the inclusion of critical attributes through the critical state; where the critical event is defined as at least one critical attribute being nonzero. The critical age is the age a unit enters the critical state. We assume an illness-death model for each individual unit of the form shown in Figure 5.3. Under the illness-death model, 4341 (92%) hard drives entered the critical state (see Figure 5.3) at some point in their lifetime. Although not shown, this model performed poorly. From this model we observed that as hard drives got older, the probability of failure within a forecast horizon of interest was higher for drives in the healthy state compared to drives in the critical state. This indicated that we misspecified the definition of the critical state. Upon further investigation, we found that 3950 (84%) hard drives experienced at least one nonzero jump in command timeout (SMART 188). Consequently, we removed SMART 188 from the set of critical attributes.

From herein, SMART 5, SMART 187, SMART 197, and SMART 198 are considered critical attributes. Under the illness-death model, with this set of critical attributes, 2428 (52%) hard drives entered the critical state, 283 drives transitioned from the healthy state to the failed state, and 1424 drives transitioned from the critical state to the failed state.

Figure 5.4 depicts the multi-state model used to model hard drive failure ages, with the inclusion of critical attributes through the critical 1 and critical 2 states; where the critical 1 and critical 2 events are defined as one critical attribute being nonzero, and at least two critical attributes being nonzero, respectively. The critical 1 age is the age a unit enters the critical 1 state. The critical 2 age is the age a unit enters the critical 2 state. Under the multi-state model, 2009 drives transitioned from the healthy state to the critical 1 state (see Figure 5.4), 1316 drives transitioned from the critical 1 state to the critical 2 state (see Figure 5.4), and 419 drives transitioned from the healthy state to the critical 2 state (without transitioning to the critical 1 state). In total 1707 drives failed; 283 drives transitioned from the healthy state to the failed state, 288 hard drives transitioned from the critical 1 state to the failed state, and 1136 drives transitioned from the critical 2 state to the failed state.

We model the failure ages of hard drives using the two-state model (depicted by Figure 5.2), the illness-death model (depicted by Figure 5.3), and the multi-state model (depicted by Figure 5.4). We specify the two-state model, the illness-death model, and the multi-state model more formally in the next sections.

5.2.4 Two-state model specifications

In this section we present the two-state model proposed by Mittman et al. (2019). Figure 5.2 depicts the two-state model used to model hard drive failure ages without the inclusion of critical attributes. We assume a two-state model for each individual unit of the form shown in Figure 5.2. The two-state model is characterised by the transition hazard:

$$\lambda(t \mid \boldsymbol{\theta}) = \lim_{\Delta \to 0} \frac{\Pr(T \in [t, t+\Delta) \mid T \geq t, \boldsymbol{\theta})}{\Delta}, \text{ for } t > 0, \tag{5.2.10}$$

where λ is the hazard rate (transition intensity) of the $0 \to 1$ transition, $\boldsymbol{\theta}$ is a vector of model parameters associated with λ, T is the failure age, and states 0 and 1 are the healthy and failed states, respectively.

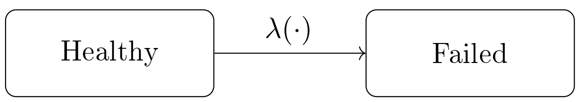

Figure 5.2: A two-state model without the inclusion of critical attributes.

A GLFP hazard is used to describe the transition rate for hard drives in the healthy state; i.e., a GLFP hazard is used to describe the $0 \to 1$ transition, where state 0 is the healthy state and state 1 is the failed state. This model describes the early failure mode and the wear-out failure mode of hard drives. More specifically,

$$T \sim \text{GLFP}(\pi, t_{p_1}, \sigma_1, t_{p_2}, \sigma_2), \tag{5.2.11}$$

where T is the failure age of a hard drive. The transition hazard is obtained using Equations (5.2.3), (5.2.4), (5.2.6), (5.2.8), and (5.2.9), and is given by

$$\begin{aligned}\lambda(t \mid \boldsymbol{\theta}) = & \frac{-t^{1/\sigma_1 - 1}\pi \log(1-p_1)\sigma_1^{-1} t_{p_1}^{-1/\sigma_1} \exp\left[\left(\frac{t}{t_{p_1}}\right)^{1/\sigma_1} \log(1-p_1)\right]}{1 - \pi\left(1 - \exp\left[\left(\frac{t}{t_{p_1}}\right)^{1/\sigma_1} \log(1-p_1)\right]\right)} \\ & - \frac{\log(1-p_2)}{\sigma_2 t_{p_2}}\left(\frac{t}{t_{p_2}}\right)^{1/\sigma_2 - 1},\end{aligned} \tag{5.2.12}$$

where $\boldsymbol{\theta} = (\pi, t_{p_1}, \sigma_1, t_{p_2}, \sigma_2)$. To infer the parameters of the two-state model, we use a Bayesian approach, selecting proper, but generally diffuse, prior distributions to improve the identification of the model parameters. Following Mittman et al. (2019) we use the 0.50 quantile for the early failure mode ($p_1 = 0.5$) and the 0.2 quantile for the wear-out failure mode ($p_2 = 0.2$).

The prior distributions used in our analysis are

$$\begin{aligned}\sigma_1 &\overset{ind.}{\sim} \text{LogNormal}(0,1), \\ \sigma_2 &\overset{ind.}{\sim} \text{LogNormal}(0,1)\ \text{Tr}(0,1), \\ t_{p_1} &\overset{ind.}{\sim} \text{LogNormal}(7,1), \\ t_{p_2} &\overset{ind.}{\sim} \text{LogNormal}(10,1), \\ \pi &\overset{ind.}{\sim} \text{LogitNormal}(-3,2),\end{aligned} \tag{5.2.13}$$

where $p_1 = 0.5$ and $p_2 = 0.2$. We truncate the distribution of σ_2 at 1 (indicated by $\text{Tr}(0,1)$), restricting the wear-out failure mode to have an increasing hazard function. The two-state model is fit using RStan, the R interface to Stan (Carpenter et al., 2017; Stan, 2022). We run one chain with 4000 iterations (2000 warm-up samples and 2000 post warm-up samples).

5.2.5 Illness-death model specifications

The illness-death model is characterised by the transition hazards:

$$\lambda_{01}(t_1 \mid \boldsymbol{\theta}_{01}) = \lim_{\Delta \to 0} \frac{\Pr(T_1 \in [t_1, t_1 + \Delta) \mid T_1 \geq t_1, T_2 \geq t_1, \boldsymbol{\theta}_{01})}{\Delta}, \text{ for } t_1 > 0 \tag{5.2.14}$$

$$\lambda_{02}(t_2 \mid \boldsymbol{\theta}_{02}) = \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 \geq t_2, T_2 \geq t_2, \boldsymbol{\theta}_{02})}{\Delta}, \text{ for } t_2 > 0 \tag{5.2.15}$$

$$\lambda_{12}(t_2 \mid T_1 = t_1, \boldsymbol{\theta}_{12}) = \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 = t_1, T_2 \geq t_2, \boldsymbol{\theta}_{12})}{\Delta}, \text{ for } 0 < t_1 < t_2, \tag{5.2.16}$$

where λ_{ij} is the hazard rate (transition intensity) of the $i \to j$ transition, $\boldsymbol{\theta}_{ij}$ is a vector of model parameters associated with λ_{ij}, T_1 and T_2 denote the critical and failure ages, respectively, and states 0, 1, and 2 correspond to the healthy, critical, and failed states, respectively.

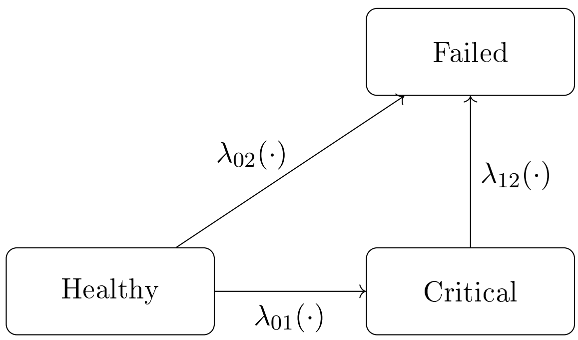

Figure 5.3: An illness-death model where the critical state is defined as at least one critical attribute being nonzero.

A GLFP hazard is used to describe the $0 \to 2$ transition and Weibull hazards are used to describe the $0 \to 1$ and $1 \to 2$ transitions, where state 0 is the healthy state, state 1 is the critical state, and state 2 is the failed state (see Figure 5.3). The hard drives in our dataset enter the critical state after the "early failure" phase and hence hard drives in the critical state are not expected to suffer from the early failure mode. More specifically, the hazard rate for the $0 \to 1$ transition is given by

$$\lambda_{01}(t \mid \boldsymbol{\theta}_{01}) = -\frac{\log(1 - p_{01})}{\sigma_{01} t_{p_{01}}} \left(\frac{t}{t_{p_{01}}} \right)^{1/\sigma_{01} - 1}, \tag{5.2.17}$$

where $\boldsymbol{\theta}_{01} = (t_{p_{01}}, \sigma_{01})$ and $p_{01} = 0.5$, the hazard rate for the $0 \to 2$ transition is given by

$$\lambda_{02}(t \mid \boldsymbol{\theta}_{02}) = \frac{-t^{1/\sigma_{1,02}-1}\pi \log(1-p_{1,02})\sigma_{1,02}^{-1} t_{p_{1,02}}^{-1/\sigma_{1,02}} \exp\left[\left(\frac{t}{t_{p_{1,02}}}\right)^{1/\sigma_{1,02}} \log(1-p_{1,02})\right]}{1-\pi\left(1-\exp\left[\left(\frac{t}{t_{p_{1,02}}}\right)^{1/\sigma_{1,02}} \log(1-p_{1,02})\right]\right)} - \frac{\log(1-p_{2,02})}{\sigma_{2,02} t_{p_{2,02}}}\left(\frac{t}{t_{p_{2,02}}}\right)^{1/\sigma_{2,02}-1}, \tag{5.2.18}$$

where $\boldsymbol{\theta}_{02} = (\pi, t_{p_{1,02}}, \sigma_{1,02}, t_{p_{2,02}}, \sigma_{2,02})$, $p_{1,02} = 0.5$, and $p_{2,02} = 0.2$, and the hazard rate for the $1 \to 2$ transition is given by

$$\lambda_{12}(t \mid \boldsymbol{\theta}_{12}) = -\frac{\log(1-p_{12})}{\sigma_{12} t_{p_{12}}}\left(\frac{t}{t_{p_{12}}}\right)^{1/\sigma_{12}-1}, \tag{5.2.19}$$

where $\boldsymbol{\theta}_{12} = (t_{p_{12}}, \sigma_{12})$ and $p_{12} = 0.1$.

To infer the parameters of the illness-death model, we use a Bayesian approach, selecting proper, but generally diffuse, prior distributions to improve the identification of the model parameters. The prior distributions for the healthy to failed transition parameters, $\boldsymbol{\theta}_{02}$, are defined in Section 5.2.4. The remaining prior distributions used in our analysis are

$$\begin{aligned} t_{p_{01}} &\overset{ind.}{\sim} \text{LogNormal}(10, 1), \\ t_{p_{12}} &\overset{ind.}{\sim} \text{LogNormal}(9, 1), \\ \sigma_{01}, \sigma_{12} &\overset{ind.}{\sim} \text{LogNormal}(0, 1)\ \text{Tr}(0, 1). \end{aligned} \tag{5.2.20}$$

We truncate the distributions of $\sigma_{01}, \sigma_{2,02}$ and σ_{12} at 1, restricting the associated failure modes to have increasing hazard functions. The illness-death model is fit using RStan. We run one chain with 4000 iterations (2000 warm-up samples and 2000 post warm-up samples).

The likelihood for the illness-death model is given by Equations (3.3.4) - (3.3.7). The definite integrals in the likelihood in Equations (3.3.6) and (3.3.7) are evaluated using the composite Simpson's rule with $M = 100$ equal subdivisions. More specifically, for the integral of the general function $f(x \mid \boldsymbol{\theta})$ over the interval $[0, l_i]$, the evaluation takes the form

$$\int_0^{l_i} f(x \mid \boldsymbol{\theta}) dx = \frac{h}{3}\left\{ f(x_0 \mid \boldsymbol{\theta}) + 4\sum_{j=1}^{M/2} f(x_{2j-1} \mid \boldsymbol{\theta}) + 2\sum_{j=1}^{M/2-1} f(x_{2j} \mid \boldsymbol{\theta}) + f(x_M \mid \boldsymbol{\theta}) \right\}, \tag{5.2.21}$$

where $x_j = jh$, for $j = 0, 1, \ldots, M$, with $h = l_i/M$; in particular, $x_0 = 0$ and $x_M = l_i$.

5.2.6 Multi-state model specifications

The multi-state model is characterised by the transition hazards:

$$\begin{aligned}
\lambda_{01}(t_1 \mid \boldsymbol{\theta}_{01}) &= \lim_{\Delta \to 0} \frac{\Pr(T_1 \in [t_1, t_1 + \Delta) \mid T_1 \geq t_1, T_2 \geq t_1, T_3 \geq t_1, \boldsymbol{\theta}_{01})}{\Delta}, && (5.2.22)\\
&\text{for } t_1 > 0\\
\lambda_{02}(t_2 \mid \boldsymbol{\theta}_{02}) &= \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 \geq t_2, T_2 \geq t_2, T_3 \geq t_2, \boldsymbol{\theta}_{02})}{\Delta}, && (5.2.23)\\
&\text{for } t_2 > 0\\
\lambda_{03}(t_3 \mid \boldsymbol{\theta}_{03}) &= \lim_{\Delta \to 0} \frac{\Pr(T_3 \in [t_3, t_3 + \Delta) \mid T_1 \geq t_3, T_2 \geq t_3, T_3 \geq t_3, \boldsymbol{\theta}_{03})}{\Delta}, && (5.2.24)\\
&\text{for } t_3 > 0\\
\lambda_{12}(t_2 \mid T_1 = t_1, \boldsymbol{\theta}_{12}) &= \lim_{\Delta \to 0} \frac{\Pr(T_2 \in [t_2, t_2 + \Delta) \mid T_1 = t_1, T_2 \geq t_2, T_3 \geq t_2, \boldsymbol{\theta}_{12})}{\Delta}, && (5.2.25)\\
&\text{for } 0 < t_1 < t_2\\
\lambda_{13}(t_3 \mid T_1 = t_1, \boldsymbol{\theta}_{13}) &= \lim_{\Delta \to 0} \frac{\Pr(T_3 \in [t_3, t_3 + \Delta) \mid T_1 = t_1, T_2 \geq t_3, T_3 \geq t_3, \boldsymbol{\theta}_{13})}{\Delta}, && (5.2.26)\\
&\text{for } 0 < t_1 < t_3\\
\lambda_{23}(t_3 \mid T_2 = t_2, \boldsymbol{\theta}_{23}) &= \lim_{\Delta \to 0} \frac{\Pr(T_3 \in [t_3, t_3 + \Delta) \mid T_2 = t_2, T_3 \geq t_3, \boldsymbol{\theta}_{23})}{\Delta}, && (5.2.27)\\
&\text{for } 0 < t_2 < t_3,
\end{aligned}$$

where λ_{ij} is the hazard rate (transition intensity) of the $i \to j$ transition, $\boldsymbol{\theta}_{ij}$ is a vector of model parameters associated with λ_{ij}, T_1, T_2, and T_3 are critical 1, critical 2, and failure ages, respectively, and states 0, 1, 2, and 3 correspond to the healthy, critical 1, critical 2, and failed states, respectively.

A GLFP hazard is used to describe the $0 \to 3$ transition, where state 0 is the healthy state and state 3 is the failed state, and Weibull hazards are used to describe all other transitions. More specifically, the hazard rate for the $i \to j$ transition is given by

$$\lambda_{ij}(t \mid \boldsymbol{\theta}_{ij}) = -\frac{\log(1 - p_{ij})}{\sigma_{ij} t_{p_{ij}}} \left(\frac{t}{t_{p_{ij}}} \right)^{1/\sigma_{ij} - 1}, \tag{5.2.28}$$

where $\boldsymbol{\theta}_{ij} = (t_{p_{ij}}, \sigma_{ij})$, for $i = 0$ and $j = 1, 2$, for $i = 1$ and $j = 2, 3$, and for $i = 2$ and $j = 3$, and $p_{01} = 0.5$, $p_{02} = 0.05$, $p_{12} = 0.3$, $p_{13} = 0.1$, and $p_{23} = 0.25$. In addition, the hazard rate for the $0 \to 3$ transition is given by

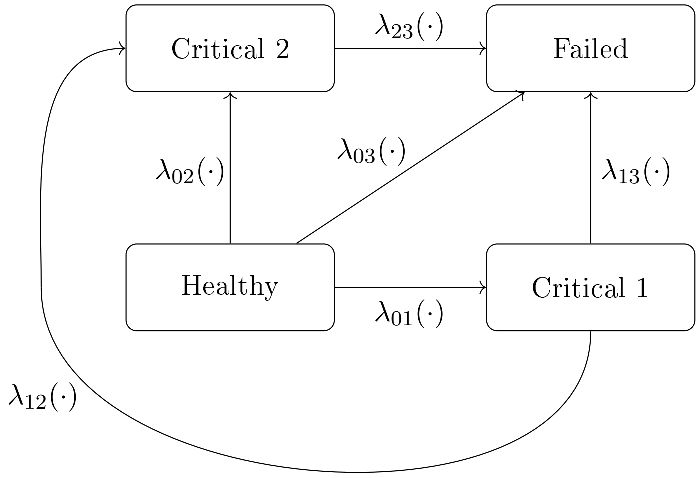

Figure 5.4: A multi-state model where the critical 1 state is defined as one critical attribute being nonzero and the critical 2 state is defined as at least two critical attributes being nonzero.

$$\lambda_{03}(t \mid \boldsymbol{\theta}_{03}) = \frac{-t^{1/\sigma_{1,03}-1}\pi\log(1-p_{1,03})\sigma_{1,03}^{-1}t_{p_{1,03}}^{-1/\sigma_{1,03}}\exp\left[\left(\frac{t}{t_{p_{1,03}}}\right)^{1/\sigma_{1,03}}\log(1-p_{1,03})\right]}{1-\pi\left(1-\exp\left[\left(\frac{t}{t_{p_{1,03}}}\right)^{1/\sigma_{1,03}}\log(1-p_{1,03})\right]\right)} - \frac{\log(1-p_{2,03})}{\sigma_{2,03}t_{p_{2,03}}}\left(\frac{t}{t_{p_{2,03}}}\right)^{1/\sigma_{2,03}-1}, \tag{5.2.29}$$

where $\boldsymbol{\theta}_{03} = (\pi, t_{p_{1,03}}, \sigma_{1,03}, t_{p_{2,03}}, \sigma_{2,03})$, $p_{1,03} = 0.5$, and $p_{2,03} = 0.2$.

To infer the parameters of the multi-state model, we use a Bayesian approach, selecting proper, but generally diffuse, prior distributions to improve the identification of the model parameters. The prior distributions for the healthy to failed transition parameters, $\boldsymbol{\theta}_{03}$, are defined in Section 5.2.4; and the prior distributions for the healthy to critical 1 transition parameters, $\boldsymbol{\theta}_{01}$, and the critical 1 to failed transition parameters, $\boldsymbol{\theta}_{13}$, are defined in Section 5.2.5. The remaining prior distributions used in our analysis are

$$\begin{aligned} t_{p_{02}} &\overset{ind.}{\sim} \text{LogNormal}(9, 1), \\ t_{p_{12}} &\overset{ind.}{\sim} \text{LogNormal}(9.5, 1), \\ t_{p_{23}} &\overset{ind.}{\sim} \text{LogNormal}(9.25, 1), \\ \sigma_{02}, \sigma_{12}, \sigma_{23} &\overset{ind.}{\sim} \text{LogNormal}(0, 1)\ \text{Tr}(0, 1). \end{aligned} \tag{5.2.30}$$

We truncate the distributions of $\sigma_{01}, \sigma_{02}, \sigma_{2,03}, \sigma_{12}, \sigma_{13}$ and σ_{23} at 1, restricting the associated failure modes to have increasing hazard functions. The multi-state model is fit using RStan. The multi-state model is fit using RStan. We run one chain with 4000 iterations (2000 warm-up samples and 2000 post warm-up samples).

5.2.7 Results

5.2.7.1 Survival probabilities and remaining useful life prediction

Figure 5.5 illustrates dynamic predictions of conditional survival probabilities over time for hard drives in the healthy state and hard drives in the critical states, obtained under the multi-state model using the parameter posterior distributions summarised in Table A.3 in Appendix A.1, conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours. The multi-state model allows us to coherently examine the impact of critical attributes on the survival probability of hard drives. The posterior predictive survival distributions can be used to compare the probabilities of failure, within a forecast horizon of interest, of drives in the healthy state to drives in the critical states and to compare the probabilities of failure of drives in the critical 1 state to drives in the critical 2 state. This allows us to concretely define the impact of a single critical attribute, and the impact of multiple critical attributes, on the survival probability of hard drives; which in turn allows us to examine the impact of a single critical attribute and the impact of multiple critical attributes on the RUL distributions of hard drives.

In addition, Figure 5.5 illustrates dynamic predictions of conditional survival probabilities over time obtained under the two-state model, using the parameter posterior distributions summarised in Table A.1 in Appendix A.1, conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours. From Figure 5.5, it appears that the survival probabilities, conditional on surviving until age τ_i, for $\tau_i = 5000, 10000, 15000, 20000$, obtained under the two-state model are a weighted mixture of the three survival curves (corresponding to the survival curves for drives in the healthy, critical 1 and critical 2 states) obtained under the multi-state model, conditional on surviving until age τ_i, for $\tau_i = 5000, 10000, 15000, 20000$. The two-state model may underestimate the survival probability of drives in the healthy state and overestimate the survival probability of drives in the critical states.

Table 5.2 provides the posterior median and central posterior 95% prediction intervals of the probability of surviving an 84 day (2016 hour) forecast horizon under the two-state model and the multi-state model, for drives in the healthy state and drives in the critical states, conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours; 84 days is approximately 3 months (or a quarter of a year) and this is how often Backblaze releases new data. The quantities obtained from the multi-state model posterior predictive survival distributions allow us to explicitly identify the impact of a single critical attribute, and the impact of multiple critical attributes, on hard drive survival probabilities. For example, from Table 5.2, we can see that the median posterior probabilities of surviving an 84 day (2016 hour) horizon for a drive of age 20,000 hours are 0.9410, 0.7658, and 0.5866 for drives with no critical attributes, one critical attribute and multiple critical attributes, respectively. In addition, under the two-state model, the median posterior probability of surviving an 84 day (2016 hour) horizon for a drive of age 20,000 hours is 0.8179 regardless of how many critical attributes the drive has acquired.

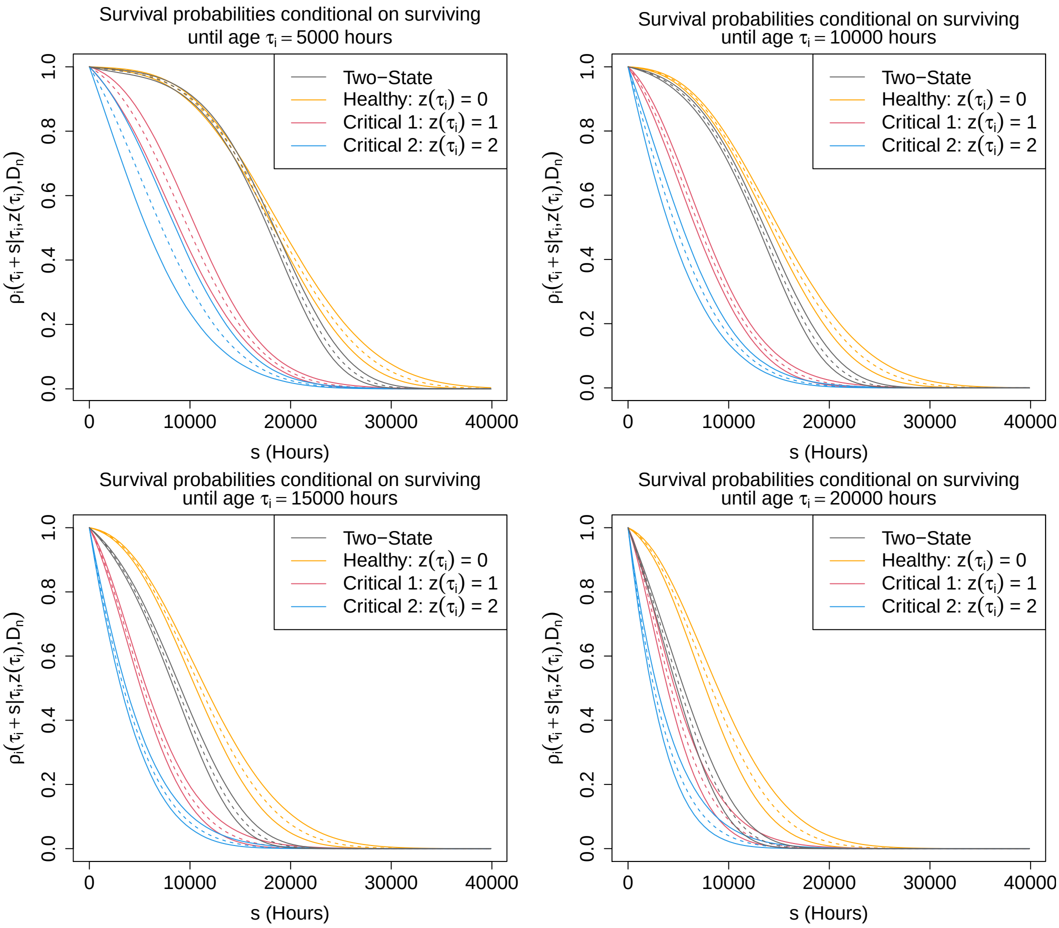

Figure 5.5: Posterior medians (dashed lines) and 95% central prediction intervals (solid lines) of the survival probability over time under the two-state model (grey) and the multi-state model, for hard drives in the healthy state (orange), hard drives in the critical 1 state (red), and hard drives in the critical 2 state (blue), conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours.

Figure 5.6 illustrates the dynamic RUL predictions obtained under the multi-state model (for drives in the healthy state, the critical 1 state, and the critical 2 state), using the parameter posterior distributions summarised in Table A.3 in Appendix A.1. Figure 5.6 provides the RUL predictions, conditional on surviving until age

Age (τ_i)	Two-State median and 95% PI	Healthy median and 95% PI	Critical 1 median and 95% PI	Critical 2 median and 95% PI
5000	0.9910 (0.9866, 0.9941)	0.9952, (0.9925, 0.9969)	0.9497, (0.9235, 0.9677)	0.8793, (0.8160, 0.9223)
10000	0.9815 (0.9787, 0.9839)	0.9910, (0.9892, 0.9924)	0.8992, (0.8818, 0.9153)	0.7748, (0.7298, 0.8130)
15000	0.9308 (0.9256, 0.9355)	0.9745, (0.9716, 0.9771)	0.8432, (0.8297, 0.8561)	0.6768, (0.6536, 0.6993)
20000	0.8179 (0.8050, 0.8309)	0.9408, (0.9323, 0.9485)	0.7842, (0.7581, 0.8060)	0.5874, (0.5543, 0.6230)

Table 5.2: Posterior median and central posterior 95% prediction intervals of the probability of surviving an 84 day (2016 hour) forecast horizon under the two-state model and the multi-state model, for hard drives in the healthy state, drives in the critical 1 state, and drives in the critical 2 state, conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours.

$\tau_i = 5000, 10000, 15000, 20000$ hours. As depicted in Figure 5.6, the RUL, conditional on surviving until age τ_i, for drives in the critical 1 state is lower than the RUL for drives in the healthy state; and the RUL, conditional on surviving until age τ_i, for drives in the critical 2 state is lower than the RUL for drives in the critical 1 state. This illustrates that drives with multiple critical attributes are more prone to failure than drives with only one critical attribute and drives with one critical attribute are more prone to failure than drives without any critical attributes.

In addition, Figure 5.6 illustrates the dynamic RUL predictions obtained under the two-state model, using the parameter posterior distributions summarized in Table A.1 in Appendix A.1. From Figure 5.6, it appears that the RUL distribution, conditional on surviving until age τ_i, for $\tau_i = 5000, 10000, 15000, 20000$, obtained under the two-state model is a weighted mixture of the three RUL distributions (corresponding to the RUL distributions for drives in the healthy, critical 1 and critical 2 states) obtained under the multi-state model, conditional on surviving until age τ_i, for $\tau_i = 5000, 10000, 15000, 20000$. The two-state model appears to underestimate the RUL of drives in the healthy state and overestimate the RUL of drives in the critical states. We will assess this more formally in the next section.

The survival probabilities were obtained by sampling from the appropriate posterior predictive survival distributions under each model; and the RUL distributions were obtained by inverse transform sampling from the relevant conditional survival probabilities.

5.2.7.2 Model assessment

In this section, we investigate the performance of the two-state model, the illness-death model and the multi-state model in a simulation study, using the AUC (discrimination) and the PE (calibration) described in Chapter 4; the square loss function is used to obtain the PE. We obtain the AUC and the PE under all models every 672 hours (28 days), assuming three time intervals for prediction of 672 hours (28 days), 1344 hours (56 days), and 2016 hours (84 days), i.e., $s = 672, 1344, 2016$. More specifically, we obtain the AUC and the PE under all models at the beginning of the study, at calendar time $\tau = 0$, and then at calendar time $\tau = 672, 1344, \ldots$, with $s = 672, 1344, 2016$. We perform Monte Carlo cross-validation, splitting the data into training (60%) and validation (40%) data. For each split, we fit the two-state model, the illness-death model and the multi-state model to the training data to obtain the joint posterior distribution of the model parameters under each model. The joint posterior distribution of the model parameters is obtained once under each model using all of the training data. We then obtain the AUC and the PE at each time point and for each forecast horizon, s. We run 1000 Monte Carlo simulations.

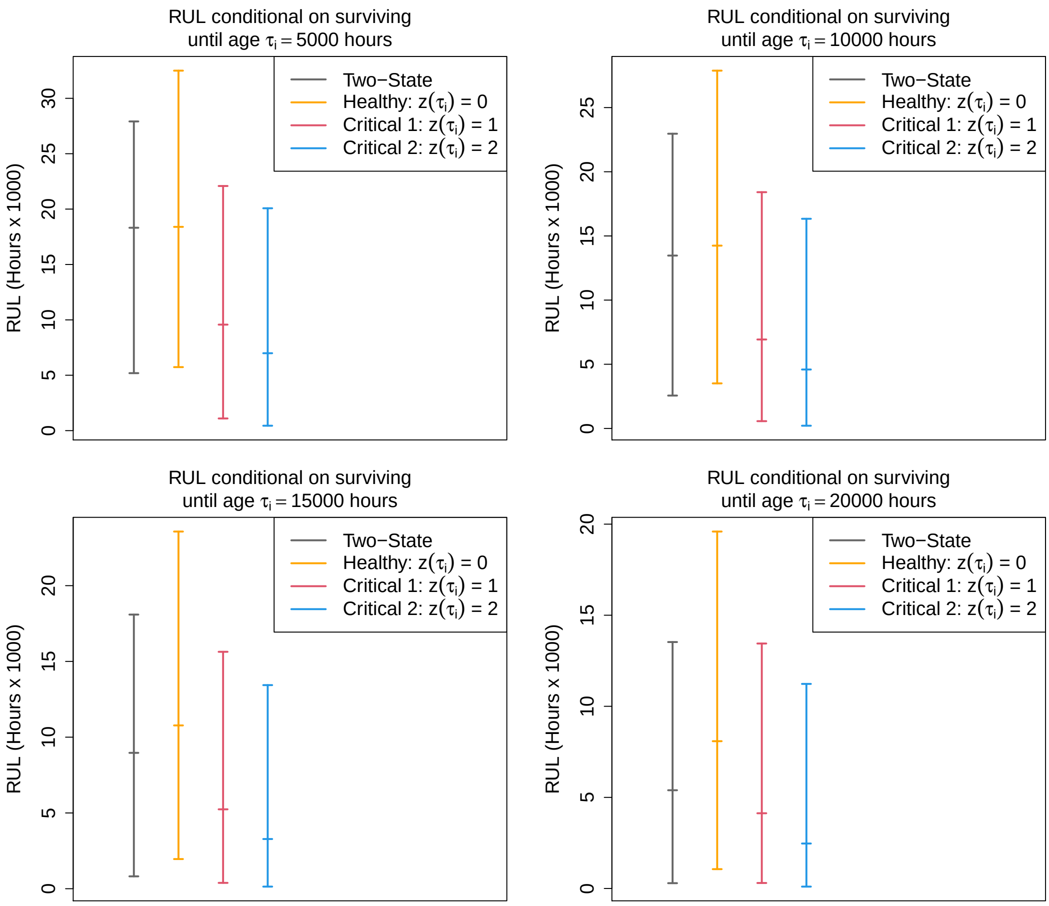

Figure 5.6: Posterior medians and 95% central prediction intervals of the posterior predictive RUL distributions for hard drives under the two-state model (grey) and the multi-state model, for hard drives in the healthy state (orange), hard drives in the critical 1 state (red), and hard drives in the critical 2 state (blue), conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours. The scales of the y-axes vary between plots.

Table A.1, Table A.2 and Table A.3 in Appendix A.1 provide summaries of the parameter posterior distributions for the two-state model, the illness-death model and the multi-state model, respectively, for one Monte Carlo cross-validation.

In Figure 5.7, and Figures A.1 and A.2 in Appendix A.1, we present the results of the simulation study, for $s = 672, 1344, 2016$. Figure 5.7 presents the results comparing the multi-state model to the two-state model; Figure A.1 presents the results comparing the illness-death model to the two-state model; and Figure A.2 presents the results comparing the multi-state model to the illness-death model. The calendar times shown were selected based on when the most failures occur in the data (Papageorgiou et al., 2019).

The boxplots in the top, middle and bottom left panels of Figure 5.7 represent the difference in the AUC between the multi-state model and the two-state model, at multiple time points, for $s = 672, 1344, 2016$, respectively. The boxplots in the top, middle and bottom right

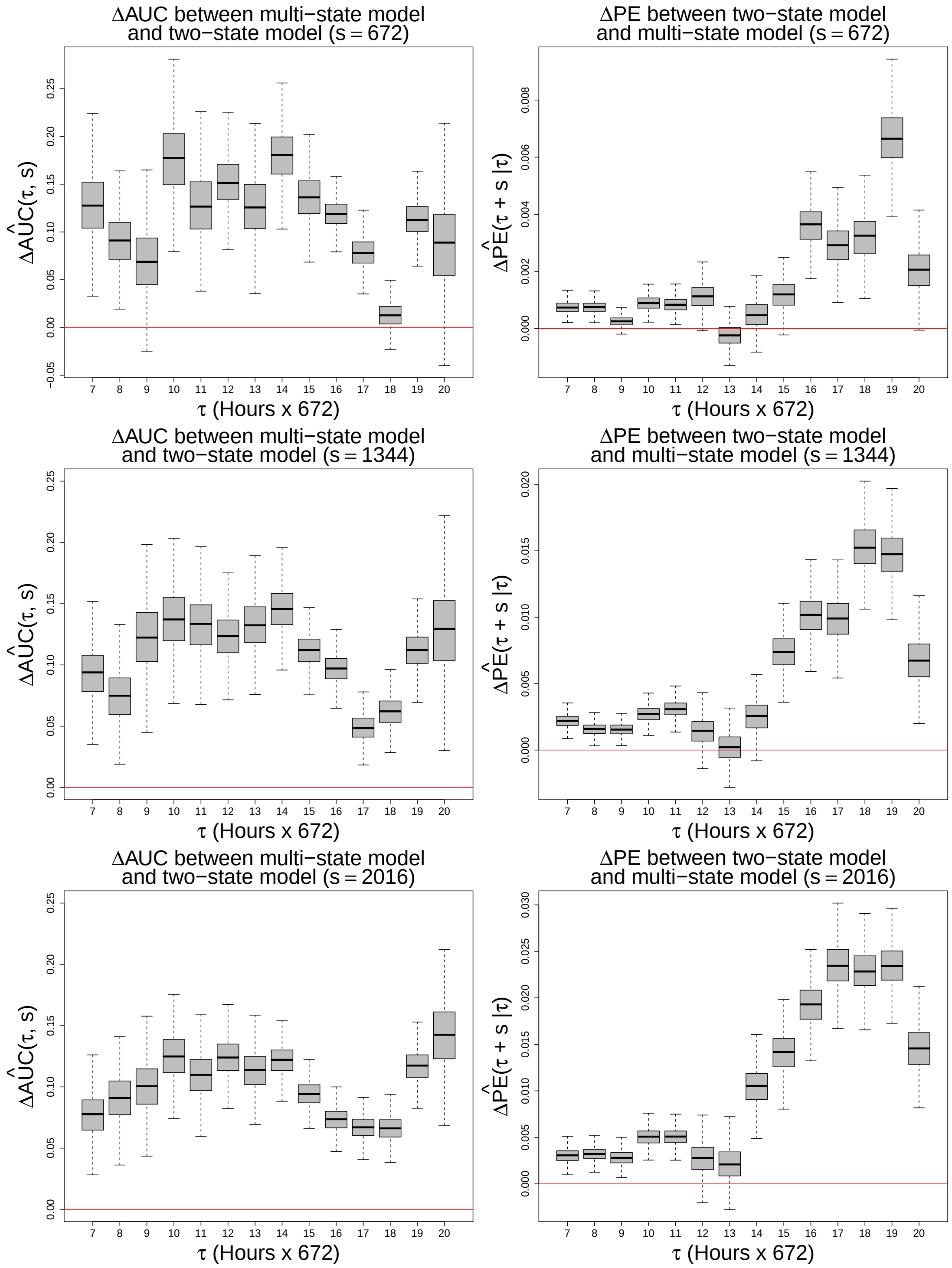

Figure 5.7: Top, middle and bottom left panels show the difference in the AUC, at multiple time points, between the multi-state model and the two-state model, for $s = 672, 1344$ and 2016, respectively. Top, middle and bottom right panels show the difference in the PE, at multiple time points, between the two-state model and the multi-state model, for $s = 672, 1344$ and 2016, respectively. The scales of the y-axes vary between plots.

panels of Figure 5.7 represent the difference in the PE between the two-state model and the multi-state model. The structure is the same in Figures A.1 and A.2, but these figures compare the illness-death model to the two-state model and the multi-state model to the illness-death model, respectively.

From the left panels in Figures 5.7 and A.1, we can see that the AUC is larger for the multi-state model and the illness-death model, compared to the two-state model, at all time points, indicating that the multi-state model and the illness-death model are better at discriminating between hard drives than the two-state model. This makes sense since the two-state model does not take into account the effect of obtaining critical attributes. Under the two-state model, the probability of surviving the forecast horizon will always be higher for a "younger" drive. The results shown in the left panels of Figures 5.7 and A.1 indicate that the multi-state model and the illness-death model are able to identify younger drives that are more likely to fail within the relevant time interval, due to these drives being in a critical states.

Moreover, from the right panels in Figures 5.7 and A.1, we can see that the PE is larger for the two-state model, compared to the multi-state model and the illness-death model, at all time points (except for prediction 13 for $s = 672$, where the PE for the two-state model is smaller, compared to the multi-state model and the illness-death model; and prediction 9 for $s = 672$ and prediction 13 for $s = 1344$, where the PE for the two-state model is approximately the same as the PE under the illness-death model). This suggests that the multi-state model and the illness-death model are more accurate than the two-state model.

Furthermore, from the left panels in Figure A.2, we can see that the AUC is larger for the multi-state model, compared to the illness-death model, at all time points (except for prediction 18 for $s = 672, 2016$, where the AUC is slightly larger for the illness-death model compared to the multi-state model). This indicates that the multi-state model is better at discriminating between hard drives compared to the illness-death model. This makes sense, since the illness-death model does not differentiate between drives with one critical attribute and drives with multiple critical attributes. Under the illness-death model, the probability of surviving the forecast horizon, for a fixed age, is identical for drives with one critical attribute and drives with multiple critical attributes. However, according to the multi-state model (see Figure 5.5), drives in the critical 2 state have a lower survival probability at all time points compared to drives in the critical 1 state, suggesting drives with multiple critical attributes are more prone to failure than drives with a single critical attribute. The results shown in the left panels of Figure A.2 indicate that the multi-state model is able to identify drives that are more likely to fail within the relevant time interval due to these drives having multiple critical attributes rather than a single critical attribute.

Finally, from the right panels in Figure A.2, we can see that the PE is larger for the illness-death model, compared to the multi-state model, at all time points. This suggests that the multi-state model is more accurate than the illness-death model.

We note that the differences in the AUC and the PE between the multi-state model and the illness-death model are not as large as the differences between the multi-state model and the two-state model or between the illness-death model and the two-state model (see the scales in Figures 5.7 and A.1 compared to Figure A.2). This suggests that more complex models, for example, a multi-state model with five states (with three critical states), may not be superior to the four-state multi-state model presented in this chapter. In addition, this model would have more parameters and be more challenging to train (due to fewer drives transitioning between each state).

5.3 Conclusions

In this chapter, we proposed a coherent and novel way to use data collected by SMART to obtain survival probabilities and DRLs for hard drives; and we examined the impact of a single critical attribute and the impact of multiple critical attributes on hard drive failure ages. We showed how to use posterior predictive survival distributions and posterior predictive RUL distributions (see Figure 5.5, Table 5.2 and Figure 5.6) to concretely examine the impact of critical attributes on hard drive survival probabilities and failure ages.

Following from Ma et al. (2015), Rincón et al. (2017), and Backblaze (2022b) we reduced the number of SMART attributes to a reduced set of "critical" attributes. A parametric model for the critical attributes is needed for the purpose of prediction. However, the evolution of each critical attribute is poorly understood, and it is challenging to predict the values of the critical attributes over time. To overcome this problem, we proposed critical states for hard drives using the critical attributes. Under our definition of the critical states, we do not need to forecast the process for any critical attribute. Instead, we must forecast the probability of entering the critical states. It is challenging to predict the value of critical attributes over time, but it is more manageable to obtain the probability of entering the critical states. This provided a framework for incorporating the erratic critical attributes.

We modelled the semi-competing risks data using the illness-death model and the multi-state model, and examined the impact of critical attributes on hard drive survival probabilities and failure ages using the multi-state model. The proposed models provided a coherent and novel way to model the failure ages of hard drives while incorporating the attributes provided by SMART.

We illustrated the multi-state modelling approach using a dataset of hard drives that is left-truncated and right-censored. We proposed transition-specific hazard functions for each event. We used the GLFP model to model the probability of failure for drives in the healthy state. The GLFP model described the early failure mode due to manufacturing defects and captured the wear-out failures observed after the initial “burn-in” period. Weibull hazards were used to describe the transitions between all other states.

We investigated the impact of critical attributes on hard drive failure ages using the multi-state model. The multi-state model suggested that the RUL for drives was impacted by a single critical attribute and impacted further by multiple critical attributes. The RUL, conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours, for drives in the critical 1 state was lower than the RUL for drives in the healthy state; and the RUL, conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours, for drives in the critical 2 state was lower than the RUL for drives in the critical 1 state. The multi-state model suggested that drives with multiple critical attributes are more prone to failure than drives with only one critical attribute and drives with one critical attribute are more prone to failure than drives without any critical attributes. We obtained the posterior predictive survival probabilities over time, conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$ hours, for drives with no critical attributes (drives in the healthy state), for drives with one critical attribute (drives in the critical 1 state) and for drives with multiple critical attributes (drives in the critical 2 state). The posterior predictive survival curves allow us to concretely define the impact of a single critical attribute and the impact of multiple critical attributes on the survival probabilities of a hard drive, which in turn allows us to examine the impact of a single critical attribute and the impact of multiple critical attributes on the RUL of a hard drive.

We assessed the performance of the multi-state model using discrimination (AUC) and calibration (PE) measures, comparing predictions obtained from the multi-state model to the illness-death model and the two-state model. We performed Monte Carlo cross-validation, splitting the data into training (60%) and validation (40%) data. For each split, we fitted the two-state model, the illness-death model and the multi-state model to the training data. We obtained the AUC and the PE every four weeks (672 hours), assuming relevant time intervals of $s = 672, 1344, 2016$ hours for prediction. We found that the multi-state model and the illness-death model outperformed the two-state model. Furthermore, we found that the multi-state model outperformed the illness-death model. The results illustrated the importance of incorporating data collected by SMART, and the multi-state model provided a framework to do this.

The differences in the AUC and the PE between the multi-state model and the illness-death model were not as large as the differences between the multi-state model and the two-state model or between the illness-death model and the two-state model. This suggested that more complex models, for example, a multi-state model with five states, may not be superior to the four-state multi-state model presented in this chapter. Furthermore, we found that command timeout, SMART 188, did not appear to be a critical attribute.

Part II

Chapter 6

Introduction

In this chapter we introduce the reliability and maintenance concepts required for the subsequent chapters and we outline the contributions in Part II of the thesis. We begin by introducing system reliability for multi-component systems. Next, define economic and stochastic dependence in the context of maintenance. Finally, we introduce degradation processes and degradation thresholds.

6.1 Multi-Component Systems

In Chapter 5 we obtained hard drive reliability (or survival) functions. This chapter provided an example of obtaining individual drive (component) reliability. Individual component reliability is the foundation of reliability assessment. The reliability of a group of hard drives, called a RAID group, is a function of the reliability of each hard drive in the RAID group. In this context, the RAID group is a multi-component system. In this section we introduce system reliability for multi-component systems. We will explore ways of representing multi-component systems using reliability block diagrams and provide examples of system reliability for systems with different structural properties.

Series systems

In reliability analysis, we often model systems graphically. This provides a visual representation of the components and how they are configured to form a system. One of the most commonly used system representations in reliability analysis is the reliability block diagram.

Figure 6.1 provides a reliability block diagram of a series system with n potentially non-identical components. A series system is a system that functions if and only if all of its n components are functioning.

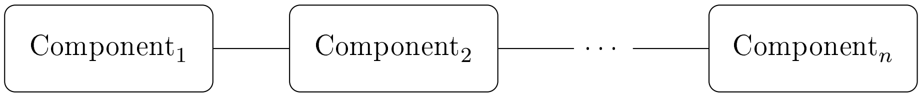

Figure 6.1: A series system with n (potentially non-identical) components.

The system reliability at time t of a series system, with n components, is

$$R_{\text{sys}}(t) = \prod_{i=1}^{n} R_i(t), \tag{6.1.1}$$

where $R_i(t)$ is the reliability of component i at time t.

Parallel systems

Figure 6.2 provides a reliability block diagram of a parallel system with n potentially non-identical components. A parallel system is a system that functions if at least one of its n components is functioning.

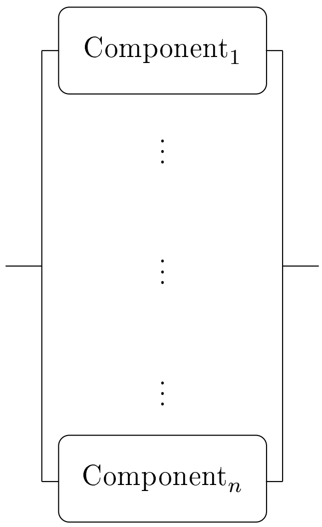

Figure 6.2: A parallel system with n (potentially non-identical) components.

The system reliability at time t of a parallel system, with n components, is

$$R_{\text{sys}}(t) = 1 - \prod_{i=1}^{n} (1 - R_i(t)), \tag{6.1.2}$$

where $R_i(t)$ is the reliability of component i at time t.

SERIES-PARALLEL SYSTEMS

Figure 6.3 provides a reliability block diagram of a series-parallel system with m subsystems. Subsystem i has n_i potentially non-identical components. A series-parallel system is a system that functions if and only if all of its m subsystems are functioning and each subsystem is a system that functions if at least one of its n_i components is functioning.

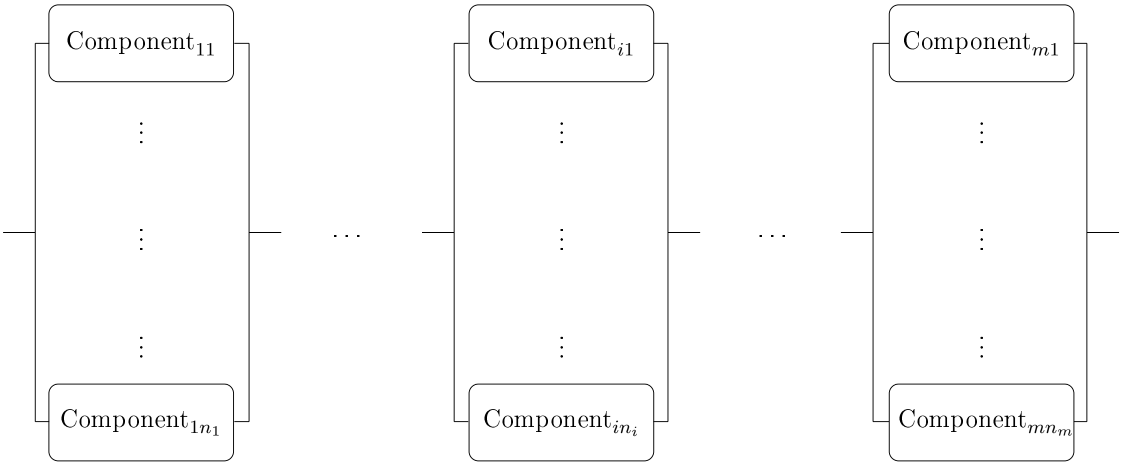

Figure 6.3: A series-parallel system with m subsystems. Subsystem i has n_i potentially non-identical components.

The system reliability at time t of the series-parallel system shown in Figure 6.3 is

$$R_{\text{sys}}(t) = \prod_{i=1}^{m} \left\{ 1 - \prod_{j=1}^{n_i} (1 - R_{ij}(t)) \right\}, \tag{6.1.3}$$

where $R_{ij}(t)$ is the reliability of component$_{ij}$ at time t.

6.2 ECONOMIC AND STOCHASTIC DEPENDENCE

Economic dependence exists when the cost of maintaining or inspecting multiple units simultaneously is different from the sum of the costs of maintaining or inspecting these units separately. Economic dependence is usually incorporated as a fixed set-up cost. For example, suppose component$_1$ in Figure 6.1 is in a failed state and all other components are in a functioning state. Hence the system is in a failed state. Maintenance is required to repair or replace component$_1$ in order for the system to function. A maintenance crew could replace each component upon failure. This would result in the most frequent maintenance, since a maintenance crew is required at each failure to replace one and only one component (assuming no simultaneous failures, for illustration).

It may be beneficial, especially if the cost of hiring a maintenance crew is large, to replace some or all of the functioning components whilst the maintenance crew replaces component$_1$ (the only failed component). This grouping of maintenance activities can reduce the number of maintenance call-outs, the number of times the maintenance crew has to set-up equipment (such as scaffolding, etc), the number of times the system is in an idle state, etc, and can result in lower long-term maintenance costs. On the other hand, there are some downsides to replacing some or all of the functioning components whilst replacing component$_1$. Replacing components that are still in a functioning state wastes RUL of working components, and replaces working components before it is necessary to do so.

Stochastic dependence arises when the state of a component influences the deterioration processes or lifetime distributions of other components, or when components are subjected to common-cause failures. In Part II of this thesis, we study failure-based load sharing systems. Failure-based load sharing systems are subject to stochastic dependence since component failures increase the workload on at least one of the remaining components, and consequently the deterioration rates, or failure rates, of these components also increase.

We assume workload enters each subsystem and is split between components within the same subsystem (see Figure 6.3). Under failure-based load sharing, the total load within subsystem i (see Figure 6.3) is shared among all functioning units within the subsystem and thus the load faced by at least one component in the subsystem changes upon failure of another component within the same subsystem. If a component fails, the system keeps operating but at least one of the remaining components, within the subsystem of the failed component, need to work harder to realise the same output level. The failure of a component thus increases the load of at least one of the working components in the same subsystem. Hence some of the working components will deteriorate faster. The system shown in Figure 6.3 is in a failed state when all components within any subsystem are in a failed state.

Stochastic dependence provides incentive to replace components upon failure, to prevent the failure rates of working components from increasing. On the other hand, economic dependence provides the incentive to group maintenance as late as possible (we assume there is no system failure penalty and that components are replaced instantly so there is no system idle time when replacing all components at once), to reduce maintenance frequency, and to reduce wasting resources; for example, as a result of replacing working components. For example, consider the system shown in Figure 6.2; to minimise the effects of economic dependence we aim to replace all components at the same time, at the failure of the system (i.e., when all components have failed). This will result in the least frequent maintenance and no wasted resources as a result of replacing working components. If we replace all components prior to system failure, this will results in replacing at least one working component; and hence some wasted resources.

Replacing all components at system failure results in no loss, due to economic dependence, and infrequent maintenance call-outs. However, this results in the maximum loss, due to stochastic dependence. By leaving components in failed states (in order to replace all components together at system failure), at least one of the working components degrade more quickly (after each component failure), since they must work harder to realise the same output. Replacing components upon failure will result in no loss due to stochastic dependence, but increases the loss due to economic dependence, since this results in the most frequent maintenance.

The optimal maintenance decisions will depend on the condition of components, the loss due to economic dependence and the loss due to stochastic dependence.

6.3 Degradation Processes

Consider Figure 6.3. We assume each subsystem receives workload at time t, $X_i(t;\boldsymbol{\theta}_{X_i})$, for $i = 1, \ldots, m$, where $\boldsymbol{\theta}_{X_i}$ is a vector of parameters. For example, $\boldsymbol{\theta}_{X_i}$ may be the shape and rate parameters of a gamma distribution. This workload is split between all working components in subsystem i, for $i = 1, \ldots, m$. Component$_{ij}$, for $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, receives workload $X_{ij}(t)$, such that

$$\sum_{j=1}^{n_i} X_{ij}(t;\boldsymbol{\theta}_{X_i}) = X_i(t;\boldsymbol{\theta}_{X_i}). \tag{6.3.1}$$

Let $\boldsymbol{Y}_i(t) = (Y_{i1}(t), \ldots, Y_{in_i}(t))$ be the failure indicator vector for subsystem i, where $Y_{ij}(t) = 1$ if Component$_{ij}$ is in a working state at time t and $Y_{ij}(t) = 0$ if Component$_{ij}$ is in a failed state at time t, for $j = 1, \ldots, n_i$. The number of functioning components in subsystem i at time t is denoted $\Pi_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t)$. If the workload in subsystem i, $X_i(t;\boldsymbol{\theta}_{X_i})$, is split evenly between all working components, we have

$$X_{ij}(t;\boldsymbol{\theta}_{X_i}) = \frac{X_i(t;\boldsymbol{\theta}_{X_i})}{\Pi_i(t)}, \tag{6.3.2}$$

for $j = 1, \ldots, n_i$. The workload received at time t, $X_{ij}(t)$, causes Component$_{ij}$ to degrade. More specifically, the increase in degradation of Component$_{ij}$ due to $X_{ij}(t)$ is

$$\Delta D_{ij}(t) = \Delta D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, \boldsymbol{\theta}_{X_i}) = \Delta D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, X_{ij}(t; \boldsymbol{\theta}_{X_i})), \tag{6.3.3}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, where $\boldsymbol{\theta}_{D_{ij}}$ is a vector of parameters that characterises how much the workload allocated to Component$_{ij}$ at time t affects the degradation level of Component$_{ij}$. When all components within subsystem i are functioning, Component$_{ij}$, $j = 1, \ldots, n_i$, degrades at the baseline rate (the "no load sharing" rate). When a component within subsystem i fails, at least one component within subsystem i receives additional workload. Hence, at least one component in subsystem i degrades more quickly than the baseline rate.

The workload vector is denoted by $\boldsymbol{X}_{ij}(a, b; \boldsymbol{\theta}_{X_i}) = \{X_{ij}(t; \boldsymbol{\theta}_{X_i}) : a < t \leq b\}$, and is the workload of Component$_{ij}$ from time a to time b. The total degradation of Component$_{ij}$ at time t is

$$D_{ij}(t) = D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, \boldsymbol{\theta}_{X_i}) = D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, \boldsymbol{X}_{ij}(t^*_{ij}, t; \boldsymbol{\theta}_{X_i})) = D_{ij}(t^*_{ij}) + \int_{t^*_{ij}}^{t} \Delta D_{ij}(s) ds, \tag{6.3.4}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, where t^*_{ij} is the time of the most recent replacement of Component$_{ij}$.

6.3.1 Degradation Thresholds

Typically papers define component failure as the time the component degradation level, $D_{ij}(t)$, reaches a fixed failure threshold, L_{ij}. However, in practice, many components will fail before, or after, the degradation reaches L_{ij}. In this paper the probability of failure (the reliability) will increase (decrease) as the degradation level increases. Components will not be assumed to fail as soon as a fixed threshold, L_{ij}, is reached. Every observed degradation level, $D_{ij}(t)$, will come with an associated probability of failure, rather than having a probability of zero of failure for $D_{ij}(t) < L_{ij}$ and a probability of one of failure for $D_{ij}(t) = L_{ij}$.

When degradation and failure time data are both available, the stochastic nature of component failure times can be modelled using random failure thresholds, as used in Hong and Meeker (2013a). Component$_{ij}$ fails at time T'_{ij} when the degradation reaches a random threshold, L_{ij}, that varies from subsystem to subsystem and between components within a subsystem. That is, $L_{ij} = D_{ij}(T'_{ij})$, where T'_{ij} is the failure time of Component$_{ij}$. The degradation threshold, L_{ij}, has the CDF $F_{L_{ij}}(D_{ij}(t); \boldsymbol{\theta}_{L_{ij}})$, where $\boldsymbol{\theta}_{L_{ij}}$ is a vector of

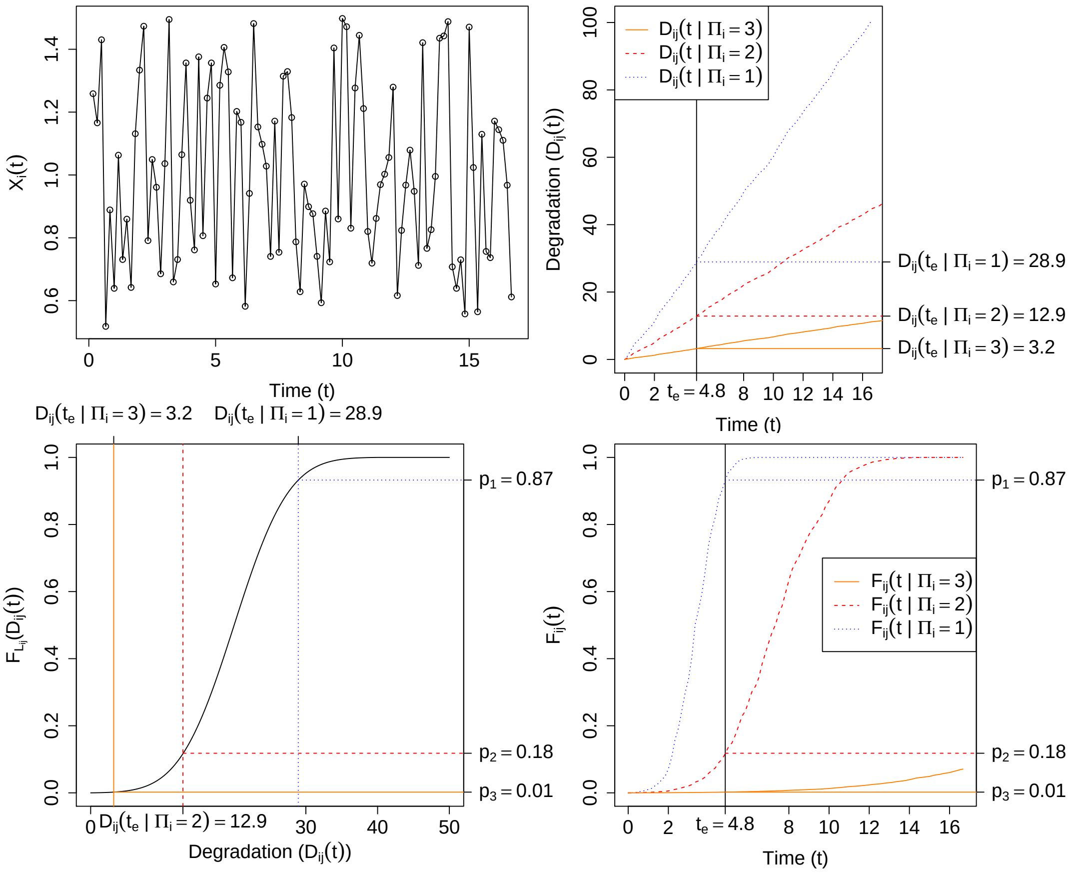

Figure 6.4: Illustration of the relationships between subsystem workload, degradation, random degradation thresholds, and component failure times for a three component subsystem (subsystem i).

model parameters which relates the degradation level of Component$_{ij}$ at time t to the probability of failure at time t. The cdf of the failure time T'_{ij}, can be expressed as

$$F_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, \boldsymbol{\theta}_{L_{ij}}, \boldsymbol{\theta}_{X_i}) = \Pr(T'_{ij} \leq t) = \Pr(L_{ij} \leq D_{ij}(t)) = F_{L_{ij}}(D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, \boldsymbol{\theta}_{X_i}); \boldsymbol{\theta}_{L_{ij}}). \tag{6.3.5}$$

In words, the probability that Component$_{ij}$ fails before time t is equivalent to the probability that the degradation level, $D_{ij}(t)$, is greater than the random failure threshold, L_{ij}.

Figure 6.4 illustrates the relationship between the covariate process (i.e., subsystem workload) and the failure time of a component for a three component subsystem (subsystem i with $n_i = 3$). Figure 6.4 (top left) shows an example workload process for subsystem i. In this illustrative example, we assume that all components in subsystem i are identical, that is they degrade at the same rate assuming a fixed workload and the probability of failure, for a fixed level of degradation, is the same - that is, we assume all components have the same failure threshold distribution. We assume the workload is split evenly between all working components.

Figure 6.4 (top right) shows the degradation of Component$_{ij}$ over time, $D_{ij}(t)$; the orange (bold) line shows the degradation of Component$_{ij}$ if Component$_{ij}$ degraded at the baseline rate (assuming the workload is split evenly between the components), the red (dashed) line shows the degradation of Component$_{ij}$ if two components are in a working state (assuming the workload is split evenly between the two working components), and the blue (dotted) line shows the degradation of Component$_{ij}$ if only one component is in a working state. We can see that the degradation at t_e varies depending on the workload received (indicated by the number of working components). The cdf for the failure threshold, L_{ij}, is shown in Figure 6.4 (bottom left), assuming each component in subsystem i is identical and hence all components have the same failure threshold distribution. From this figure we can see that each degradation level has an associated probability of failure. The probability of Component$_{ij}$ failing at time t_e is 0.01 if Component$_{ij}$ degraded at the baseline rate, compared to 0.87 if Component$_{ij}$ was the only functioning component in subsystem i.

The CDFs for the failure time T'_{ij} are shown in Figure 6.4 (bottom right). The bold (dashed, dotted) line shows the cdf for T'_{ij} for the degradation process shown by a bold (dashed, dotted) line in Figure 6.4 (top right). The probability of Component$_{ij}$ failing before time t_e is equivalent to the probability of $D_{ij}(t_e)$ exceeding the failure threshold, L_{ij}. The degradation at time t_e depends on the covariate history $\boldsymbol{X}_{ij}(0, t_e) = \{X_{ij}(s), 0 < s \leq t_e\}$, where

$$X_{ij}(s) = \frac{X_i(s)}{\Pi_i(s)}, \tag{6.3.6}$$

and $\Pi_i(s)$ is the number of working components at time s.

We note that the workload was assumed to be split evenly between all components in the example, but the workload does not have to be split evenly in practice.

Figure 6.5 illustrates the differences between the deterministic approach, the fixed threshold approach, and the random threshold approach. The red (dashed) line in Figure 6.5 depicts an example RUL distribution of a new component under the random threshold approach. This approach represents the RUL distribution for a component taking into account all failure time uncertainty. The blue (dotted) line depicts an example RUL distribution of a new component under the fixed threshold approach. Components are assumed to fail when the degradation level reaches a fixed threshold, $L = F_L^{-1}(0.5)$, where F_L is the failure time distribution of the red (dashed) distribution. The black (dashed) line at $t = 29.4$ shows the expected failure time of a new component (i.e., the RUL distribution

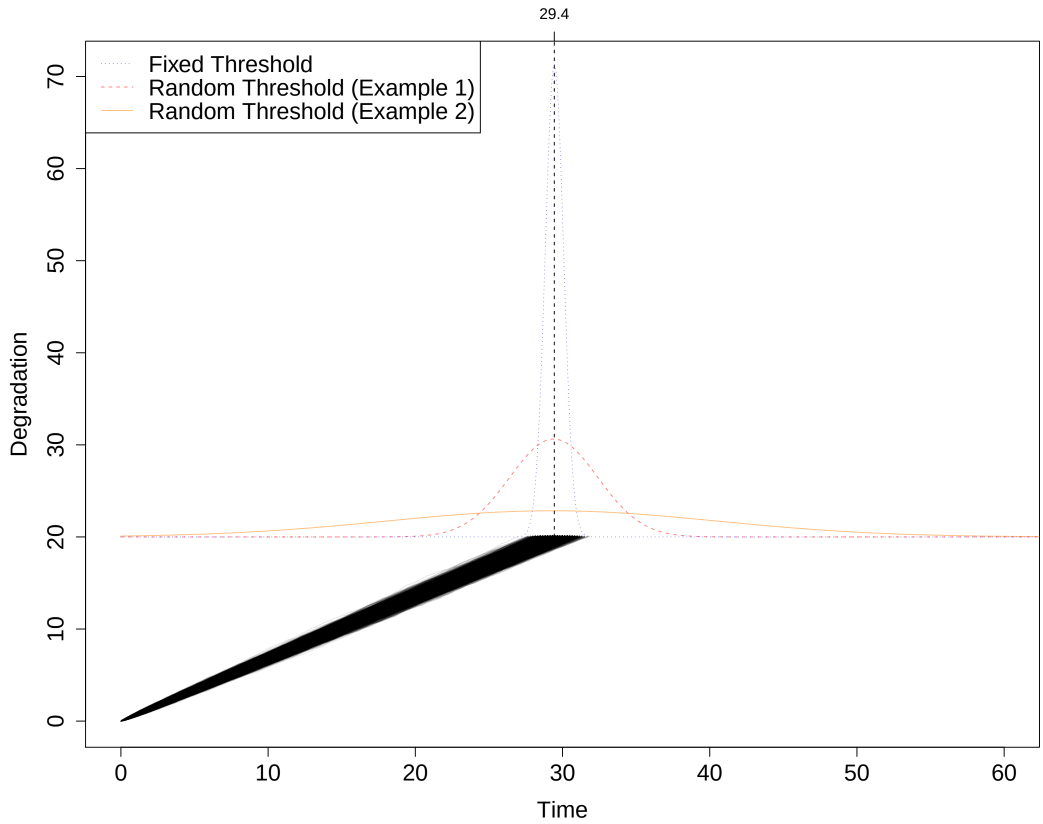

Figure 6.5: RUL distributions for a new component under different approaches. The blue (dotted) line shows the RUL distribution under the fixed threshold approach, the red (dashed) and orange (bold) lines show the RUL distributions under the random threshold approach, assuming small and large amounts of stochastic variation for the random failure threshold, respectively. The black (dashed) line at $t = 29.4$ shows the expected failure time of a new component (i.e., the RUL distribution under the deterministic approach). The black (bold) lines depict the degradation distributions at each time point.

under the deterministic approach). From Figure 6.5 we can see that the fixed threshold and deterministic approaches do not fully capture the RUL distributions of components, and will often lead to suboptimal maintenance decisions.

The orange (bold) line in Figure 6.5 depicts an example RUL distribution of a new component under the random threshold approach, with a larger amount of stochastic variation for the random failure threshold. Again, the fixed threshold is taken to be $L = F_L^{-1}(0.5)$ and the expected failure time is $t = 29.4$. This example highlights further the disadvantages of using the fixed threshold and deterministic approaches. The larger the amount of stochastic variation in the random failure threshold, the more information the fixed threshold and deterministic approaches are ignoring.

6.4 CONCLUSIONS

In this chapter we introduced the reliability and maintenance concepts required for the subsequent chapters. We introduced system reliability for series, parallel, and series-parallel systems. We then defined economic and stochastic dependence and described how multi-component systems can be economically dependent through fixed set-up costs and stochastically dependent through failure-based load sharing. Finally, we defined workload and described how workload causes components to degrade. Finally, we defined random failure thresholds and described the differences between RUL distributions under the random threshold, fixed threshold, and deterministic (expected failure time) approaches.

Chapter 7

Condition-Based Maintenance Policy

In this chapter we investigate a CBM policy for making maintenance decisions. We begin by illustrating the penalties incurred by multi-component systems as a result of economic dependence, through a fixed set-up cost, and stochastic dependence, through failure-based load sharing. We then propose a novel CBM policy that incorporates a loss-based utility function, which is a combination of interpretable penalties that encapsulate the costs of economic and stochastic dependence, in a sequential Bayesian decision framework. In Chapter 8, we compare the loss-based utility function presented in this chapter to the commonly used cost per unit time utility function. In addition, in Chapter 8, we will compare a random-threshold approach to a fixed-threshold approach and an expected failure time approach to highlight the importance of incorporating all uncertainty when making maintenance decisions.

7.1 Penalties for multi-component systems with economic and stochastic dependence

Before we derive the loss-based utility function, we illustrate, using examples, the penalties observed for multi-component systems with economic and stochastic dependence. First, we define maintenance opportunities and the settings in which our utility function is applicable.

7.1.1 Maintenance opportunities

We consider the case where all component failure times are maintenance opportunities. Component failures increase the degradation rate of other components, hence there is an incentive to replace components as soon as they fail. On the other hand, set-up costs are involved when replacing a component. Every maintenance opportunity has the option of no replacements unless the system is in a failed state.

We do not consider predictive maintenance, as we decide which components to replace at maintenance opportunities - defined as a time a component fails. At each maintenance opportunity, we used a predictive model to identify the components to replace (which could involve components that are not in a failed state). For example, consider Figure 6.2 and assume Component_1 has failed and all other components are in a working state. The failure time of Component_1 is an opportunity to perform maintenance and we want to decide which components to replace at this maintenance opportunity (if any at all) and this could include the failed component, Component_1, and other components that are not in failed states.

System Settings

We consider systems where:

(i) Each component is continuously monitored.

(ii) Compared to the time period between two maintenance actions, the duration of maintenance is negligible. Hence, there is no system downtime cost.

(iii) Replacements are perfect and restore components to a "good-as-new" condition.

(iv) There is no system failure cost.

The key properties of the maintenance policy we will develop could be extended to systems with non-negligible maintenance times and to systems with a system failure cost.

For illustration, we consider a three-component parallel system (see Figure 6.2 with $n = 3$) and assume the three components are identical (meaning the components cost the same and degrade at the same rate, assuming a fixed workload). We will compare three different ways of scheduling three component replacements to illustrate the penalties incurred by multi-component systems with economic and stochastic dependence.

For illustration, we assume that workload enters the parallel system and that this workload is split evenly between the working components. More specifically,

$$X_i(t) = \frac{X_{\text{sys}}(t)}{\Pi(t)}, \tag{7.1.1}$$

for $i = 1, 2, 3$, where $X_i(t)$ is the workload for component$_i$ at time t, $X_{\text{sys}}(t)$ is the system workload at time t, and $\Pi(t)$ is the number of functioning components at time t. The workload will cause the system components to degrade. The system can function as long as at least one component is functioning.

7.1.2 ILLUSTRATION ONE

A HYPOTHETICAL EXAMPLE

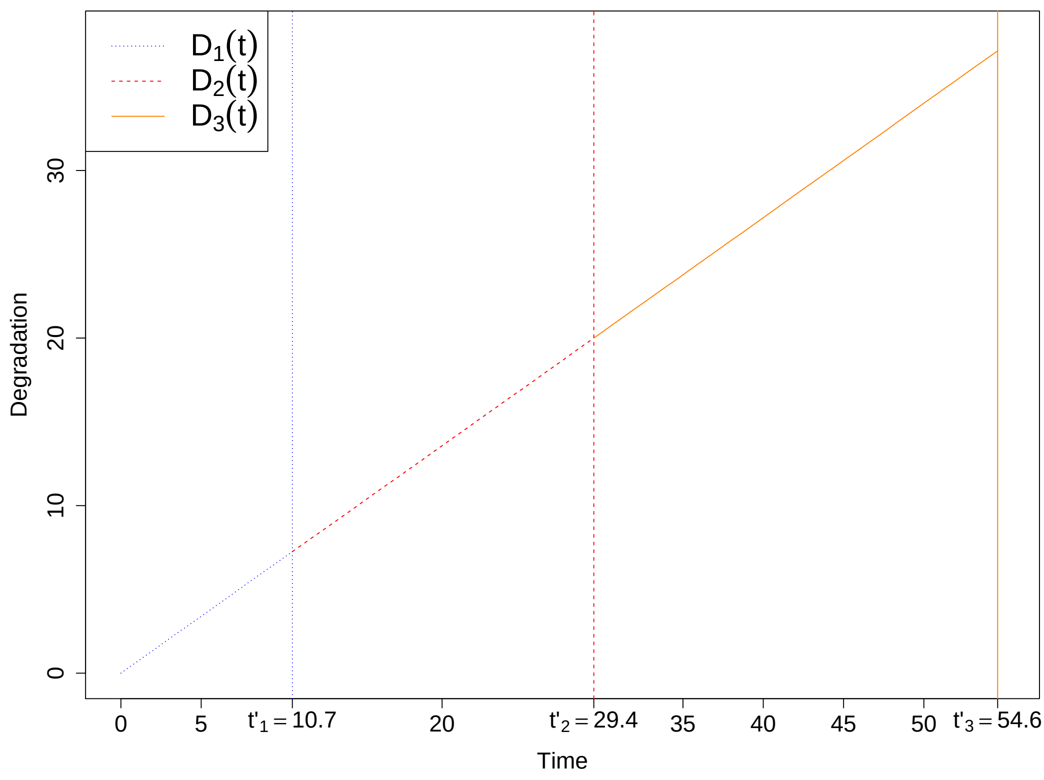

Figure 7.1: A hypothetical system where load sharing does not increase the degradation rate of the working components. The blue (dotted) line shows the degradation path of Component$_1$, the red (dashed) line shows the degradation path of Component$_2$, and the orange (bold) line shows the degradation path of Component$_3$. Component$_1$ is observed to fail at t'_1, Component$_2$ is observed to fail at t'_2, and Component$_3$ is observed to fail at t'_3. We note that the degradation paths for all components are the same for $t \in [0, t'_1]$, and the degradation paths for Component$_2$ and Component$_3$ are the same for $t \in (t'_1, t'_2]$.

Figure 7.1 provides example degradation paths for three new components. In this example, all three components are new at time $t = 0$ and workload continuously enters the system, and hence all components continuously degrade. At time $t'_1 = 10.7$ Component$_1$ is observed to fail. We note that the degradation for all components is the same up to time t'_1 since we assume the workload is split evenly between all components and that the components degrade at the same rate. Under our definition of maintenance opportunities, the failure of Component$_1$ at t'_1 is an opportunity to perform maintenance.

For illustrative reasons, let us assume that we can leave Component$_1$ in a failed state without the remaining working components degrading more quickly (i.e., Component$_2$ and Component$_3$ sustain the full system workload but degrade at the baseline rate). In this, hypothetical scenario, there are no advantages to replacing Component$_1$, only disadvantages (for example, paying a maintenance team to replace Component$_1$). Hence, we leave Component$_1$ in a failed state and both Component$_2$ and Component$_3$ continue to degrade (at the baseline rate).

From Figure 7.1, we can see that Component$_2$ and Component$_3$ continue to degrade until the failure of Component$_2$ at $t'_2 = 29.4$. In practice, this would be an opportunity to perform maintenance. However, for illustrative purposes, we will assume that we can leave Component$_1$ and Component$_2$ in failed states without any impact to Component$_3$ (i.e., Component$_3$ sustains the full system workload but degrades at the baseline rate).

From Figure 7.1, we can see that Component$_3$ continues to degrade until failure at $t'_3 = 54.6$. Now the whole system is in a failed state and hence maintenance must be performed. We make all three component replacements at once at $t'_3 = 54.6$ and pay one set-up cost, S. This example is hypothetical because, in practice, leaving components in failed states will result in the remaining working components degrading more quickly. The hypothetical example is used as a baseline for comparison and illustration reasons.

In this hypothetical scenario, we replaced three components on three separate occasions, and hence paid the set-up cost three times. Hence, the total penalty for replacing components upon failure is

$$\Lambda_{\mathrm{H}} = \Lambda_{\mathrm{H}}(\phi, \phi, \{\mathrm{Component}_1, \mathrm{Component}_2, \mathrm{Component}_3\}) = S, \tag{7.1.2}$$

where S is the set-up cost.

Three practical examples

Replacing components upon failure

In practice, leaving components in failed states will result in the remaining working components degrading more quickly. Suppose we replace Component$_1$ at $t'_1 = 10.7$. This will prevent load sharing between the components in the system. Then suppose the next component failure is Component$_2$ at $t'_2 = 29.4$ and suppose we replace Component$_2$ upon failure. Again, this will prevent load redistribution between the components. Finally, suppose the next failure is Component$_3$ at $t'_3 = 54.6$, and we replace Component$_3$ at t'_3. This is one possible way of making three component replacements.

In this scenario, we replaced three components on three separate occasions, and hence paid the set-up cost three times. Moreover, relative to the hypothetical example, we replaced Component$_1$ $(t'_3 - t'_1)$ time units earlier, and we replaced Component$_2$ $(t'_3 - t'_2)$ time units earlier. Hence, the total penalty, relative to the hypothetical example, for replacing components upon failure is

$$\Lambda_1 = \Lambda_1(\{\text{Component}_1\}, \{\text{Component}_2\}, \{\text{Component}_3\}) = 3S + \bar{c}(t'_3 - t'_1) + \bar{c}(t'_3 - t'_2), \tag{7.1.3}$$

where S is the set-up cost and $\bar{c}$ is the expected cost per unit time for a component in the system (noting the expected cost per unit time for each component is the same since we assumed all components are identical).

Intuitively, Λ_1 penalises replacing components upon failure because this approach results in three set-up costs, and requires Component$_1$ to be replaced $(t'_3 - t'_1)$ time units earlier and Component$_2$ to be replaced $(t'_2 - t'_1)$ time units earlier, than in the hypothetical example.

The cost of replacing Component$_1$ $(t'_3 - t'_1)$ time units earlier than in the hypothetical example is $\bar{c}(t'_3 - t'_1)$. This is the expected cost of using a component for $(t'_3 - t'_1)$ time units. Since we replaced Component$_1$ $(t'_3 - t'_1)$ time units earlier, than in the hypothetical example, we also expect Component$_1$ to fail $(t'_3 - t'_1)$ time units earlier.

Replacing a component prior to failure

Suppose we replace Component$_1$ and Component$_2$ at $t'_1 = 10.7$. Then suppose the next failure is Component$_3$ at $t'_3 = 54.6$, and we replace Component$_3$ at t'_3. This is another way of making three component replacements.

In this scenario, we replaced three components on two separate occasions, and hence paid the set-up cost twice. Moreover, relative to the hypothetical example, we replaced Component$_1$ $(t'_3 - t'_1)$ time units earlier and Component$_2$ $(t'_3 - t'_1)$ time units earlier. Hence, the total penalty, relative to the hypothetical example, is

$$\Lambda_2 = \Lambda_2(\{\text{Component}_1, \text{Component}_2\}, \{\text{Component}_3\}) = 2S + 2\bar{c}(t'_3 - t'_1). \tag{7.1.4}$$

Replacing all components at system failure

Suppose we postpone maintenance until all three components have failed and we replace all components at system failure. This is another way of making three component replacements.

Figure 7.2 illustrates example degradation paths observed when we take into account the effects of load sharing. Comparing Figures 7.1 and 7.2, we can see the impacts of load sharing. The failure of Component$_1$ causes Component$_2$ and Component$_3$ to degrade more quickly. In this scenario, Component$_2$ fails at $t'^L_2 = 19$, compared to $t'_2 = 29.4$ when degrading at the baseline rate. The increased load resulted in Component$_2$ failing 10.4 time units earlier. Furthermore, the degradation rate of Component$_3$ increases further at the failure of Component$_2$. We observe Component$_3$ to fail at $t'^L_3 = 21.8$; 32.8 time units earlier than if Component$_3$ degraded at the baseline rate.

Hence, the total penalty, relative to the hypothetical example, is

$$\Lambda_3 = \Lambda_3(\phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\}) = S + 3\bar{c}(t'_3 - t'^L_3), \tag{7.1.5}$$

where ϕ denotes the empty set or no maintenance.

The optimal maintenance schedule

We note that

$$\begin{aligned}
\Lambda_1 - \Lambda_2 &= S + \bar{c}(t'_1 - t'_2),\\
\Lambda_1 - \Lambda_3 &= 2S + \bar{c}(t'^L_3 - t'_3) + \bar{c}(t'^L_3 - t'_2) + \bar{c}(t'^L_3 - t'_1),\\
\Lambda_2 - \Lambda_3 &= S + \bar{c}(t'^L_3 - t'_3) + 2\bar{c}(t'^L_3 - t'_1).
\end{aligned} \tag{7.1.6}$$

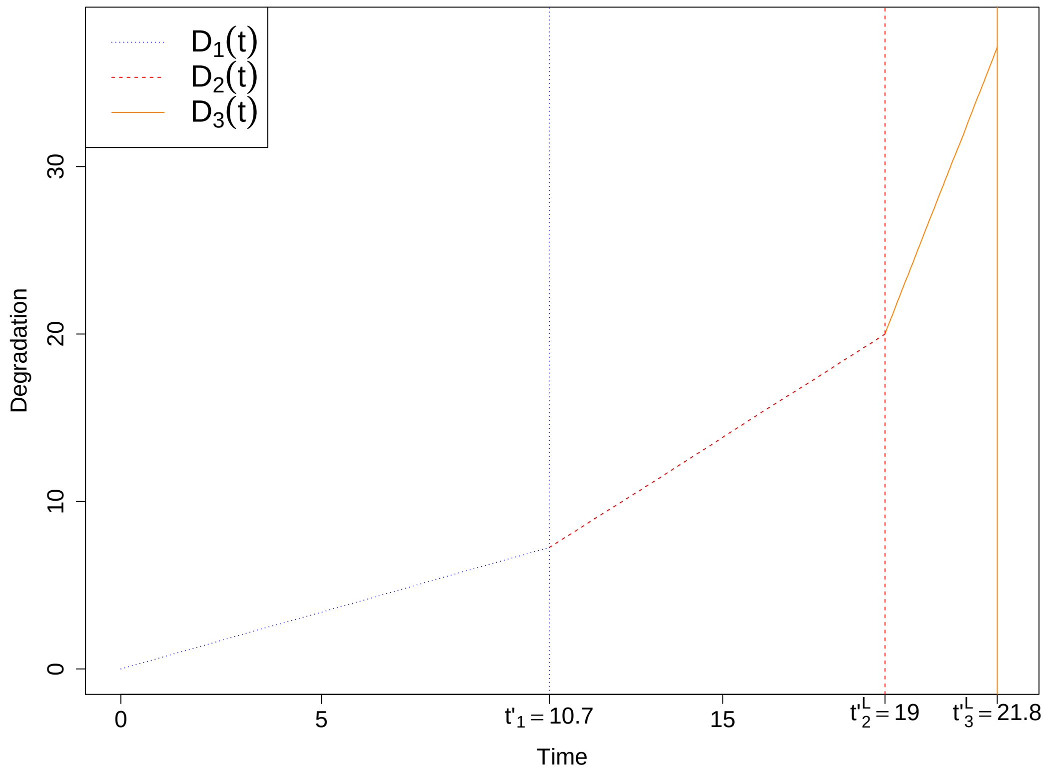

Figure 7.2: Observed degradation paths for each component when all components are replaced at system failure. The blue (dotted) line shows the degradation path of Component$_1$, the red (dashed) line shows the degradation path of Component$_2$, and the orange (bold) line shows the degradation path of Component$_3$. Component$_1$ is observed to fail at t'_1, Component$_2$ is observed to fail at t'^L_2, and Component$_3$ is observed to fail at t'^L_3. We note that the degradation paths for all components are the same for $t \in [0, t'_1]$, and the degradation paths for Component$_2$ and Component$_3$ are the same for $t \in (t'_1, t'^L_2]$.

The optimal way to make three component replacements is the action set with the smallest penalty. In other words, if $\Lambda_1 - \Lambda_2 < 0$ and $\Lambda_1 - \Lambda_3 < 0$, then approach one (replacing components upon failure) is the optimal choice.

The differences in penalties can be interpreted relative to each other. For example, $\Lambda_1 - \Lambda_3$, is the difference in penalties between approach one and approach three. In approach one, we replaced one component at times t'_1, t'_2, and t'_3. In approach three, we replaced three components at t'^L_3. In approach one, we pay two extra set-up costs, and we expect Component$_3$ to fail $(t'_3 - t'^L_3)$ time units later, Component$_2$ to fail $(t'_2 - t'^L_3)$ time units later, and Component$_1$ to fail $(t'^L_3 - t'_1)$ time units earlier, relative to approach three.

7.1.3 Illustration two

Let us view the same example considered in Section 7.1.2 but from a different perspective.

Replacing components upon failure

Consider replacing components upon failure. We replace Component$_1$ at t'_1, Component$_2$ at t'_2, and Component$_3$ at t'_3. Replacing components upon failure prevents load redistribution between components. In this scenario, we replaced three components on three separate occasions, and hence paid the set-up cost three times. The total penalty is

$$\Lambda_1^* = \Lambda_1^*(\{\text{Component}_1\}, \{\text{Component}_2\}, \{\text{Component}_3\}) = 3S, \tag{7.1.7}$$

since we paid three set-up costs.

Replacing a component prior to failure

Suppose we replace Component$_1$ and Component$_2$ at $t'_1 = 10.7$. Then suppose the next failure is Component$_3$ at $t'_3 = 54.6$, and suppose we replace Component$_3$ at t'_3. The total penalty is

$$\Lambda_2^* = \Lambda_2^*(\{\text{Component}_1, \text{Component}_2\}, \{\text{Component}_3\}) = 2S + \bar{c}(t'_2 - t'_1). \tag{7.1.8}$$

We paid two set-up costs and are penalised for this. In addition, we replaced a working component and hence did not utilise some RUL; this penalty is the cost of wasting the RUL of Component$_2$.

Replacing all components at system failure

Suppose we postpone maintenance until all three components have failed and suppose we replace all components at system failure. Figure 7.3 depicts the loss in life, of Component$_2$ and Component$_3$, between t'_1 (the first maintenance opportunity) and t'^L_2 (the subsequent maintenance opportunity).

From the point of view of t'_1 (the first maintenance opportunity), not replacing Component$_1$ results in Component$_2$ and Component$_3$ degrading more quickly until the

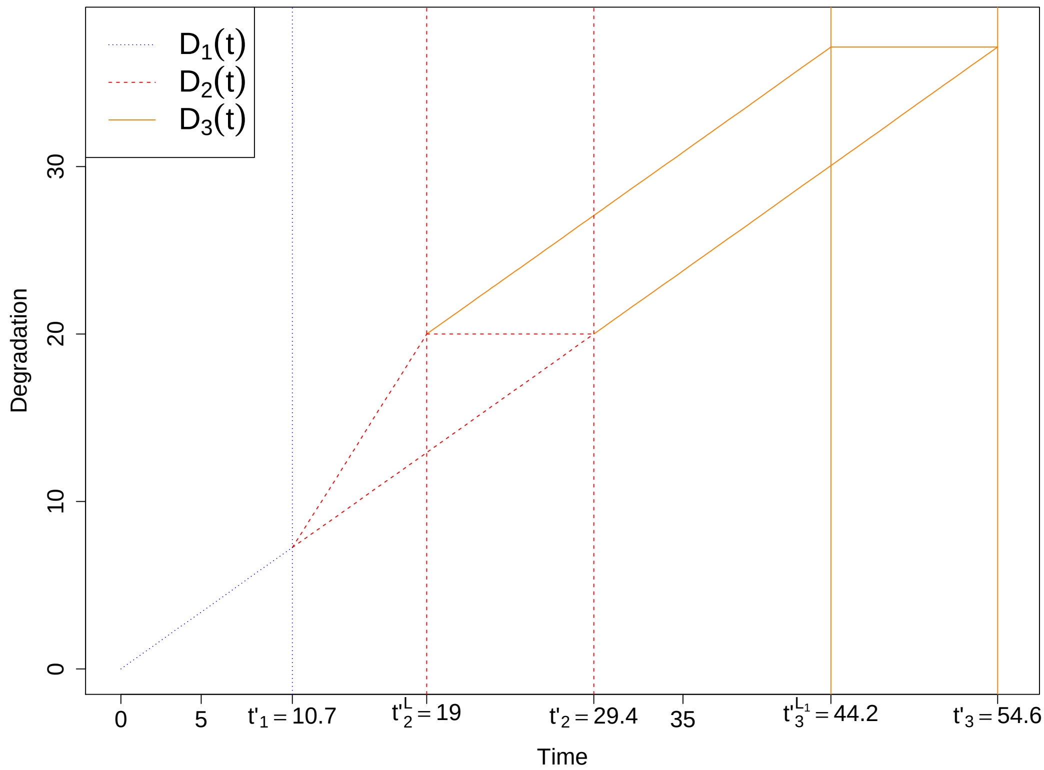

Figure 7.3: Degradation paths illustrating the impacts of load sharing between t'_1 and t'^L_2. Component$_1$ is observed to fail at t'_1. Between t'_1 and t'^L_2 Component$_2$ and Component$_3$ degrade more quickly. Component$_2$ is observed to fail at t'^L_2 because of load sharing, and is observed to fail at t'_2 if Component$_2$ degrades at the baseline rate. Component$_3$ is observed to fail at t'_3 if Component$_3$ degrades at the baseline rate. Component$_3$ is observed to fail at $t'^{L_1}_3$ due to the impacts of load sharing between t'_1 and t'^L_2. We note that the degradation paths for all components are the same for $t \in [0, t'_1]$, and the degradation paths for Component$_2$ and Component$_3$ are the same for $t \in (t'_1, t'^L_2]$, under the load sharing scenario, and the degradation paths for Component$_2$ and Component$_3$ are the same for $t \in (t'_1, t'_2]$, under the baseline rate scenario.

subsequent maintenance opportunity. The total penalty due to not performing maintenance at t'_1 is

$$\Lambda^*_3(\phi) = \bar{c}(t'_2 - t'^L_2) + \bar{c}(t'_3 - t'^{L_1}_3) - \bar{c}(t'^L_2 - t_1). \tag{7.1.9}$$

This first two terms in this penalty correspond to the loss, in terms of cost, due to load sharing between the current maintenance opportunity and the subsequent maintenance opportunity; $(t'_2 - t'^L_2)$ is the loss in life of Component$_2$, and $(t'_3 - t'^{L_1}_3)$ is the loss in life of Component$_3$. The final term is a reward, since we postponed the maintenance of

Component$_1$. Not replacing Component$_1$ results in a loss, due to load sharing, but also a reward, since we postponed the replacement of Component$_1$ by $(t_2'^L - t_1)$ time units. Replacing Component$_1$ at $t_2'^L$ instead of t_1' will save $(t_2'^L - t_1)$ time units of the lifetime of Component$_1$. In other words, if we replace Component$_1$ at $t_2'^L$ we expect Component$_1$ to fail $(t_2'^L - t_1)$ time units later than if we replaced Component$_1$ at t_1'.

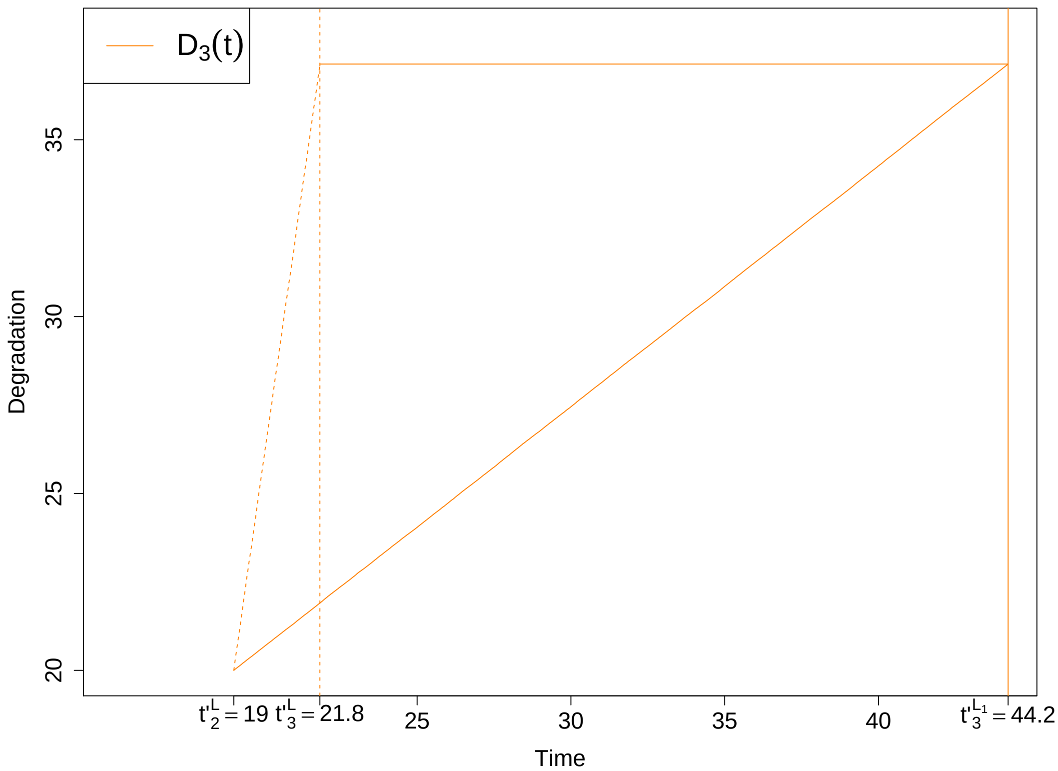

Figure 7.4: Degradation paths illustrating the impacts of load sharing between $t_2'^L$ and $t_3'^L$. Component$_3$ is observed to fail at $t_3'^{L_1}$ if Component$_3$ degrades at the baseline rate after $t_2'^L$. Component$_3$ is observed to fail at $t_3'^L$ due to the impacts of load sharing between $t_2'^L$ and $t_3'^L$.

At $t_2'^L$ (the next maintenance opportunity - see Figure 7.4) we perform no maintenance. Hence, Component$_3$ has to sustain the whole system workload. This results in further loss due to load sharing. The total loss, due to performing no maintenance at t_1' and no maintenance at $t_2'^L$, is

$$\Lambda_3^*(\phi, \phi) = \Lambda_3^*(\phi) + \bar{c}(t_3'^{L_1} - t_3'^L) - 2\bar{c}(t_3'^L - t_2'^L). \qquad (7.1.10)$$

This first term is the penalty due to performing no maintenance at t_1'. The second term is the loss, in terms of cost, due to load sharing between $t_2'^L$ and $t_3'^L$ (the failure time of

Component$_3$ and hence the failure of the system). The final term is the reward due to not replacing Component$_1$ and Component$_2$ at $t_2'^L$. Replacing Component$_1$ and Component$_2$ at $t_3'^L$ instead of at $t_2'^L$ saves $(t_3'^L - t_2'^L)$ time units of life for both Component$_1$ and Component$_2$. In other words, if we replace Component$_1$ and Component$_2$ at $t_3'^L$ we expect Component$_1$ and Component$_2$ to fail $(t_3'^L - t_2'^L)$ time units later than if we replaced both components at $t_2'^L$.

Finally, at $t_3'^L$ we replace all three components, and hence pay a set-up cost. The total penalty is

$$\begin{aligned}\Lambda_3^* &= \Lambda_3^*(\phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\}) \\ &= S + \bar{c}(t_2' - t_2'^L) + \bar{c}(t_3' - t_3'^L) - \bar{c}(t_3'^L - t_1') - \bar{c}(t_3'^L - t_2'^L).\end{aligned} \tag{7.1.11}$$

We paid one set-up cost and are penalised for this. In addition, we postponed the maintenance of Component$_1$ from t_1' to $t_3'^L$ and the maintenance of Component$_2$ from $t_2'^L$ to $t_3'^L$. We are rewarded for this (terms 4 and 5 in Equation (7.1.11)). Furthermore, leaving Component$_1$ in a failed state at t_1' resulted in Component$_2$ and Component$_3$ degrading more quickly due to load sharing between t_1' and $t_2'^L$; and then leaving Component$_1$ and Component$_2$ in failed states at $t_2'^L$ resulted in Component$_3$ sustaining the full system workload from $t_2'^L$ until the failure of the system at $t_3'^L$. We are penalized for components failing sooner due to load sharing (terms 2 and 3 in Equation (7.1.11)).

The penalties due to economic and stochastic dependence

We note that

$$\begin{aligned}\Lambda_1^* - \Lambda_2^* &= S + \bar{c}(t_1' - t_2') = \Lambda_1 - \Lambda_2, \\ \Lambda_1^* - \Lambda_3^* &= 2S + \bar{c}(t_3'^L - t_3') + \bar{c}(t_3'^L - t_2') + \bar{c}(t_3'^L - t_1') = \Lambda_1 - \Lambda_3, \\ \Lambda_2^* - \Lambda_3^* &= S + \bar{c}(t_3'^L - t_3') + 2\bar{c}(t_3'^L - t_1') = \Lambda_2 - \Lambda_3.\end{aligned} \tag{7.1.12}$$

Therefore, both ways of visualising the penalties result in the same outcome. In summary, there are four different types of penalties:

1. The set-up cost, S.
2. A penalty for replacing working components, and hence wasting RUL.
3. A load sharing penalty, due to components degrading more quickly because of load redistribution.

4. A reward, at the expense of load sharing, due to postponing maintenance of failed components.

In some scenarios, replacing components upon failure will be beneficial. For example, if the set-up cost is small relative to the loss due to load sharing. It may also be beneficial to replace a component in a working state. For example, suppose Component$_1$ is in a failed state, Component$_2$ is heavily degraded, and Component$_3$ is relatively new. It may be beneficial to replace a working component (Component$_2$) alongside Component$_1$ (the failed component). In this scenario, we prevent load sharing and also prevent another (shortly expected) maintenance opportunity at the failure of Component$_2$ (a heavily degraded component).

If, at the failure of Component$_1$, Component$_2$ and Component$_3$ are both heavily degraded and expected to fail soon, it may be beneficial to postpone maintenance until system failure and to replace all components together. This will result in only one set-up cost and will synchronise component failure times (by synchronising component degradation levels).

In contrast, for the three-component parallel system with identical components, replacing components upon failure will desynchronise component failure times. In the short term, it may seem beneficial to replace a component to prevent load sharing. In the long term, it may be beneficial to take a large loss (due to load sharing) to synchronise component failure times (possibly resulting in better maintenance groupings in the future). For example, if we replace a component upon failure to prevent load sharing, we may end up having to do this again and again to prevent load sharing at future opportunities.

For example, suppose $S = 10$ and $\bar{c} = 1$, we obtain

$$\Lambda_1^* - \Lambda_3^* = 2S + \bar{c}(t_3'^L - t_3') + \bar{c}(t_3'^L - t_2') + \bar{c}(t_3'^L - t_1') = \Lambda_1 - \Lambda_3 = -9.3. \tag{7.1.13}$$

This suggests that approach one is superior to approach three. Approach one may have the smallest penalty, but we considered a relatively short forecast horizon (i.e, the optimal way to make three component replacements). If we forecasted further ahead (for example, considering the optimal way to make 6, 9, 12, etc, component replacements) the optimal action to perform at t_1' may not be to replace Component$_1$.

Figure 7.5 illustrates the degradation levels of all three components at time t_3' (if we replace components upon failure). From Figure 7.5 we can see that at time t_3', upon replacing Component$_3$, the degradation level of Component$_1$ is 29.8, the degradation level of Component$_2$ is 17, and the degradation level of Component$_3$ is 0. In other words, Component$_1$ is heavily degraded, Component$_2$ is moderately degraded, and Component$_3$

is new. Therefore, the expected failure times of the components are out of sync. This may result in expensive maintenance costs in the future. For example, suppose Component$_1$ fails shortly after t_3'. This presents an opportunity to perform maintenance. Postponing maintenance of Component$_1$ to group component replacements at a later time, in order to pay the set-up cost less frequently, could be expensive in this scenario since Component$_2$ and Component$_3$ are not expected to fail within a relatively short time horizon. Therefore, Component$_2$ and Component$_3$ may have to sustain the whole system workload for a relatively long period of time, resulting in a relatively large amount of loss because of load redistribution due to Component$_1$ being in a failed state.

Replacing Component$_1$ at this maintenance opportunity will prevent load sharing. However, the next maintenance opportunity is expected to be at the failure of Component$_2$, and the expected failure times of the components are expected to stay out of sync. Once again, replacing Component$_2$ may result in the system having one new, one moderately degraded and one relatively new component (hence the components remain out of sync). Finally, replacing Component$_1$ and Component$_2$ would result in wasting RUL of Component$_2$. All possible maintenance options may be unappealing. Replacing each component upon failure may be the optimal short term decision, but it may result in expensive maintenance at a later time point.

On the other hand (see Figure 7.2), by replacing all components at the same time (at $t_3'^L$) all three components are new and are expected to fail at the same time. This could be beneficial in the long-term. For example (see Figure 7.5), Component$_1$ failing at $t_1' = 10.7$ and Component$_3$ failing at $t_3' = 54.6$ (while degrading at the baseline rate) may have been unlikely events. I.e., the component failure times may, on average, be expected to be clustered closely together, and we just sampled an unlikely set of failure times. If this is the case, it may be better to postpone maintenance and synchronize component failure times. Forecasting further ahead could provide this valuable information, even though this may come with initial (relatively) large penalties due to leaving components in failed states for (relatively) long periods of time.

The impact of desynchronising component failure times depends on several factors, including the stochastic variation in component failure times. The smaller the stochastic variation in component failure times, the more likely component failure times will be clustered closely together. This provides more incentive to cluster component replacement times (even if, in the short term, it appears more beneficial to replace a failed component). In contrast, if component failure times are highly stochastic, they are less likely to be clustered closely together. This gives more incentive to replace components upon failure. This highlights the importance of a sequential maintenance policy that takes into account long-term impacts of maintenance decisions along with the uncertainty in component failure times (as well as other uncertainties).

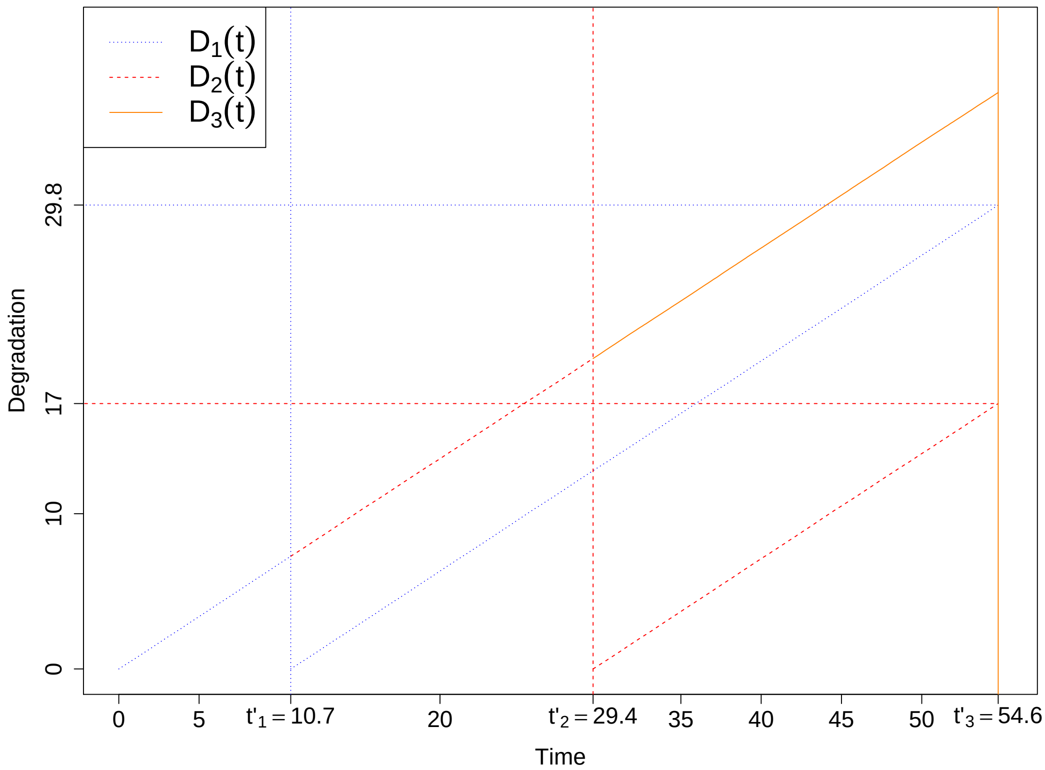

Figure 7.5: Example degradation paths if Component$_1$ fails and is replaced at t_1', Component$_2$ fails and is replaced at t_2', and Component$_3$ fails and is replaced at t_3'. The blue (dotted) line shows the degradation path of Component$_1$, the red (dashed) line shows the degradation path of Component$_2$, and the orange (bold) line shows the degradation path of Component$_3$.

To choose the optimal set of components to replace at each maintenance opportunity we need to consider all possible penalties and the impact maintenance replacements have on future maintenance decisions (for example, a large penalty, due to load sharing, at a maintenance opportunity may help with future groupings). Moreover, the decisions will depend on the uncertainties in component failure times.

This discussion highlights that our decision framework needs to be sequential, to prevent optimal short term decisions resulting in higher long term maintenance costs, and needs to incorporate uncertainty.

We will now derive a one-step ahead maintenance policy which incorporates a loss-based utility function in a Bayesian framework. We then extend the policy to a sequential decision policy that minimises the long-term expected loss of a sequence of maintenance actions.

7.2 CONDITION-BASED MAINTENANCE POLICY

The following general maintenance policy is proposed to be used at maintenance opportunities. The failure time of a component is an opportunity to perform maintenance. Maintenance must be performed if the system is in a failed state. We consider a system with m subsystems. Subsystem i has n_i potentially non-identical components (see Figure 6.3). The replacement cost for Component$_{ij}$ is c_{ij}.

7.2.1 PROBLEM SETUP

In this section we describe the state of each component immediately prior to, and immediately following, a maintenance action at a maintenance opportunity.

Suppose we are at the kth maintenance opportunity, at time t_k, and have observed

$$\begin{aligned} \boldsymbol{x}(t_k) &= (x_1(t_k), \ldots, x_m(t_k)), \\ \boldsymbol{y}(t_k) &= (y_{11}(t_k), y_{12}(t_k), \ldots, y_{1n_1}(t_k), \ldots, y_{m1}(t_k), \ldots, y_{mn_m}(t_k)), \end{aligned} \tag{7.2.1}$$

where $x_i(t_k) > 0$ is the observed workload at t_k for subsystem i and $y_{ij}(t_k) = 1$ if Component$_{ij}$ is in a working state at t_k and $y_{ij}(t_k) = 0$ if Component$_{ij}$ is in a failed state at t_k.

Thus, we can obtain

$$\boldsymbol{D}(t_k) = (D_{11}(t_k), D_{12}(t_k), \ldots, D_{1n_1}(t_k), \ldots, D_{m1}(t_k), \ldots, D_{mn_m}(t_k)), \tag{7.2.2}$$

where $D_{ij}(t_k)$ is the degradation of Component$_{ij}$ at t_k and is given by

$$D_{ij}(t_k) = \int_{t_{ij}^*}^{t_k} \Delta D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, x_{ij}(t)) dt, \tag{7.2.3}$$

where t_{ij}^* is the time of the most recent replacement of Component$_{ij}$, $x_{ij}(t)$ is the workload of Component$_{ij}$ at time t, $\Delta D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, x_{ij}(t))$ is the degradation increment for Component$_{ij}$ at time t due to workload $x_{ij}(t)$, and $\boldsymbol{\theta}_{D_{ij}}$ is a vector of model parameters that characterises how much the workload entering Component$_{ij}$ at time t affects the degradation level of Component$_{ij}$.

We need to decide which action, $u_{t_k} \in U_{t_k}$, to perform at the current time point, t_k, where u_{t_k} is a set of components to be replaced and can be empty unless the system is in a failed state, and U_{t_k} is the set of all possible actions.

After applying u_{t_k} we observe

$$\boldsymbol{y}^+(t_k) = (y^+_{11}(t_k), y^+_{12}(t_k), \ldots, y^+_{1n_1}(t_k), \ldots, y^+_{m1}(t_k), \ldots, y^+_{mn_m}(t_k)), \tag{7.2.4}$$

and obtain

$$\boldsymbol{D}^+(t_k) = (D^+_{11}(t_k), D^+_{12}(t_k), \ldots, D^+_{1n_1}(t_k), \ldots, D^+_{m1}(t_k), \ldots, D^+_{mn_m}(t_k)), \tag{7.2.5}$$

where

$$y^+_{ij}(t_k) = \begin{cases} 1 & \text{if Component}_{ij} \in u_{t_k}, \\ y_{ij}(t_k) & \text{if Component}_{ij} \notin u_{t_k}, \end{cases} \tag{7.2.6}$$

and

$$D^+_{ij}(t_k) = \begin{cases} 0 & \text{if Component}_{ij} \in u_{t_k}, \\ D_{ij}(t_k) & \text{if Component}_{ij} \notin u_{t_k}. \end{cases} \tag{7.2.7}$$

The term $D_{ij}(t_k)$ represents the degradation of Component_{ij} at t_k prior to maintenance and $D^+_{ij}(t_k)$ represents the degradation of Component_{ij} at t_k after applying u_{t_k}. The degradation of Component_{ij} after applying u_{t_k} will remain unchanged if Component_{ij} is not replaced and will be restored to zero if Component_{ij} is replaced. The term $y_{ij}(t_k)$ indicates whether Component_{ij} is in a working state at t_k prior to maintenance. Similarly, the term $y^+_{ij}(t_k)$ indicates whether Component_{ij} is in a working state at t_k after maintenance. If Component_{ij} is not replaced the state of Component_{ij} will remain unchanged after maintenance; if Component_{ij} is replaced then Component_{ij} will be in a working state after maintenance.

The number of functioning components in subsystem i after applying u_{t_k} is

$$\pi^+_i(t_k) = \sum_{j=1}^{n_i} y^+_{ij}(t_k), \tag{7.2.8}$$

for $i = 1, \ldots, m$.

7.2.2 Short-Sighted Maintenance Policy

In this section we introduce the penalties that occur as a result of a maintenance action, and propose a decision framework to decide which maintenance action to perform, at a maintenance opportunity, based on one-step ahead forecasts.

Applying u_{t_k} produces a corresponding penalty. The penalty due to applying u_{t_k} is given by

$$\begin{aligned}\Lambda(u_{t_k} \mid T_{k+1}, &\boldsymbol{X}(t_k,\infty), \boldsymbol{D}(t_k,\infty), \boldsymbol{\theta}) = SI_{\{|u_{t_k}|\neq 0\}} \\ &+ \sum_{i=1}^{m}\sum_{j=1}^{n_i} \bar{c}_{ij}(\tau_{ij}^{-} - t_k) I_{\{\text{Component}_{ij} \in u_{t_k}, y_{ij}(t_k)=1\}} \\ &+ \sum_{i=1}^{m}\sum_{j=1}^{n_i} \bar{c}_{ij}(\tau_{ij}^{+} - \tau_{ij}^{+,L}) - \sum_{i=1}^{m}\sum_{j=1}^{n_i} \bar{c}_{ij}(T_{k+1} - t_k) I_{\{\text{Component}_{ij} \notin u_{t_k}, y_{ij}(t_k)=0\}},\end{aligned} \tag{7.2.9}$$

where T_{k+1} is the time of the $(k+1)$st maintenance opportunity, $\boldsymbol{X}(t_k,\infty) = (\boldsymbol{X}_1^T(t_k,\infty), \ldots, \boldsymbol{X}_m^T(t_k,\infty))^T$, where the ith row of $\boldsymbol{X}(t_k,\infty)$ is denoted $\boldsymbol{X}_i(t_k,\infty) = \{X_i(t;\boldsymbol{\theta}_{X_i}), t_k < t < \infty\}$, and represents the workload of subsystem i for times $t > t_k$, $\boldsymbol{D}(t_k,\infty) = (\boldsymbol{D}_{11}^T(t_k,\infty), \ldots, \boldsymbol{D}_{1n_1}^T(t_k,\infty), \ldots, \boldsymbol{D}_{mn_m}^T(t_k,\infty))^T$, where $\boldsymbol{D}_{ij}^T(t_k,\infty) = \{D_{ij}(t;\boldsymbol{\theta}_{D_{ij}}, \boldsymbol{X}_{ij}(t_k,t;\boldsymbol{\theta}_{X_i})), t_k < t < \infty\}$, and represents the degradation of Component$_{ij}$ for times $t > t_k$, $\boldsymbol{\theta} = (\boldsymbol{\theta}_{D_{11}}, \boldsymbol{\theta}_{L_{11}}, \boldsymbol{\theta}_{X_1}, \ldots, \boldsymbol{\theta}_{D_{mn_m}}, \boldsymbol{\theta}_{L_{mn_m}}, \boldsymbol{\theta}_{X_m})$ is a vector of model parameters, S is the set-up cost, $|u_{t_k}|$ is the number of components in u_{t_k}, $\bar{c}_{ij}$ is the expected cost per unit time for Component$_{ij}$, τ_{ij}^{-} is the expected failure time for Component$_{ij}$, obtained immediately prior to u_{t_k} being performed, if Component$_{ij}$ degraded at the baseline rate after t_k, τ_{ij}^{+} is the expected failure time for Component$_{ij}$, obtained immediately after u_{t_k} is performed, if Component$_{ij}$ degraded at the baseline rate after t_k, $\tau_{ij}^{+,L}$ is the expected failure time for Component$_{ij}$, obtained immediately after u_{t_k} is performed, taking into account potential load sharing between t_k and T_{k+1}, and

$$I_A = \begin{cases} 1, & \text{if event } A \text{ occurs,} \\ 0, & \text{otherwise.} \end{cases} \tag{7.2.10}$$

More specifically, $(\tau_{ij}^{-} - t_k)I_{\{\text{Component}_{ij} \in u_{t_k}, y_{ij}(t_k)=1\}}$ is the expected wasted life of Component$_{ij}$ if Component$_{ij}$ is a working component that is replaced at t_k; $(\tau_{ij}^{+} - \tau_{ij}^{+,L})$ is the expected loss in life of Component$_{ij}$ between t_k and T_{k+1} due to load sharing; and

$(T_{k+1} - t_k)I_{\{\text{Component}_{ij} \notin u_{t_k}, y_{ij}(t_k)=0\}}$ is the idle time of a failed component that is not replaced at t_k.

The first term in Equation (7.2.9) is the set-up cost and is only paid if the action, u_{t_k}, is non-empty. The second term is only paid if one or more components in u_{t_k} are in a working state at t_k prior to maintenance. More specifically, the second term is the total loss, in terms of cost, due to replacing working components. The third term is the loss, in terms of cost, due to load sharing between t_k and T_{k+1}. The final term is the reward for postponing maintenance of failed components and is hence subtracted from the total penalty.

The expected failure time τ_{ij}^{-} is given by

$$\tau_{ij}^{-} = \begin{cases} \int_{t_k}^{\infty} t f_{L_{ij}}(D_{ij}(t; \boldsymbol{X}_{ij}(t_k, t)))dt = \int_{t_k}^{\infty} [1 - F_{L_{ij}}(D_{ij}(t; \boldsymbol{X}_{ij}(t_k, t)))]dt, \\ \text{if } y_{ij}(t_k) = 1, \\ t'_{ij}, \\ \text{if } y_{ij}(t_k) = 0, \end{cases} \tag{7.2.11}$$

where $f_{L_{ij}}(\cdot; \boldsymbol{\theta}_{L_{ij}})$ is the pdf for the degradation threshold, L_{ij}, $F_{L_{ij}}(\cdot; \boldsymbol{\theta}_{L_{ij}})$ is the cdf for the degradation threshold, $\boldsymbol{\theta}_{L_{ij}}$ is a vector of model parameters associated with $f_{L_{ij}}(\cdot)$ and $F_{L_{ij}}(\cdot)$, $\boldsymbol{X}_{ij}(t_k, t) = \{X_{ij}(s) : t_k < s \le t\}$ is the workload of Component$_{ij}$ from time t_k to time t conditional on all components in subsystem i being in a working state (i.e., the baseline or "no load sharing" workload), and t'_{ij} is the most recent failure time for Component$_{ij}$. In addition, the degradation of Component$_{ij}$ at time $t > t_k$ is

$$D_{ij}(t) = D_{ij}(t_k) + \int_{t_k}^{t} \Delta D_{ij}(s; \boldsymbol{\theta}_{D_{ij}}, X_{ij}(s; \boldsymbol{\theta}_{X_i}))ds. \tag{7.2.12}$$

The expected failure time τ_{ij}^{+} is given by

$$\tau_{ij}^{+} = \begin{cases} \int_{t_k}^{\infty} t f_{L_{ij}}(D_{ij}(t; \boldsymbol{X}_{ij}(t_k, t)))dt = \int_{t_k}^{\infty} [1 - F_{L_{ij}}(D_{ij}(t; \boldsymbol{X}_{ij}(t_k, t)))]dt, \\ \text{if } y_{ij}^{+}(t_k) = 1, \\ t'_{ij}, \\ \text{if } y_{ij}^{+}(t_k) = 0. \end{cases} \tag{7.2.13}$$

In this case, the degradation of Component$_{ij}$ at time $t > t_k$ is

$$D_{ij}(t) = D_{ij}^{+}(t_k) + \int_{t_k}^{t} \Delta D_{ij}(s; \boldsymbol{\theta}_{D_{ij}}, X_{ij}(s; \boldsymbol{\theta}_{X_i}))ds. \tag{7.2.14}$$

The expected failure time $\tau_{ij}^{+,L}$ is given by

$$\tau_{ij}^{+,L} = \begin{cases} \int_{t_k}^{\infty} t f_{L_{ij}}(D_{ij}(t; \boldsymbol{X}_{ij}^{+}(t_k, t)))dt = \int_{t_k}^{\infty} [1 - F_{L_{ij}}(D_{ij}(t; \boldsymbol{X}_{ij}^{+}(t_k, t)))]dt, \\ \text{if } y_{ij}^{+}(t_k) = 1, \\ t'_{ij}, \\ \text{if } y_{ij}^{+}(t_k) = 0. \end{cases} \tag{7.2.15}$$

In this case, the degradation of Component$_{ij}$ at time $t > t_k$ is

$$D_{ij}(t) = D_{ij}^{+}(t_k) + \int_{t_k}^{t} \Delta D_{ij}(s; \boldsymbol{\theta}_{D_{ij}}, X_{ij}^{+}(s; \boldsymbol{\theta}_{X_i}))ds, \tag{7.2.16}$$

where $\boldsymbol{X}_{ij}^{+}(t_k, t) = \{X_{ij}^{+}(s) : t_k < s \leq t\}$ is the workload of Component$_{ij}$ from time t_k to time t, taking into account potential load sharing between t_k and T_{k+1}. The workload of Component$_{ij}$ at time s, $X_{ij}^{+}(s)$, for $t_k < s \leq T_{k+1}$, is the workload for Component$_{ij}$ after potential load redistribution and $X_{ij}^{+}(s)$, for $T_{k+1} < s < \infty$, is the baseline (or "no load sharing") workload for Component$_{ij}$. If $\pi_i^{+}(t_k) = n_i$, where n_i is the number of components in subsystem i, then there is no load redistribution since all components in subsystem i are in a working state. Moreover, it is possible that the workload for Component$_{ij}$ does not change even if $\pi_i^{+}(t_k) \neq n_i$; it may be the case that no extra load is distributed to Component$_{ij}$. We just require $\sum_{j=1}^{n_i} X_{ij}^{+}(t) = X_i(t)$, where $X_i(t)$ is the workload of system i at time t. Furthermore, $\boldsymbol{X}_{ij}^{+}(T_{k+1}, \infty) = \boldsymbol{X}_{ij}(T_{k+1}, \infty)$ (the baseline or "no load sharing" workload). I.e., $\boldsymbol{X}_{ij}^{+}(t_k, \infty)$ is used to measure the impact of load sharing between t and T_{k+1} and only depends on the potential load redistribution between t_k and T_{k+1}.

For example, if the workload in subsystem i, $X_i(s; \boldsymbol{\theta}_{X_i})$, for $s > t_k$, is split evenly between all working components, we have

$$\begin{aligned} X_{ij}(s) = X_{ij}(s; \boldsymbol{\theta}_{X_i}) &= \frac{X_i(s; \boldsymbol{\theta}_{X_i})}{n_i}, \\ X_{ij}^{+}(s) = X_{ij}^{+}(s; \boldsymbol{\theta}_{X_i}) &= \frac{X_i(s; \boldsymbol{\theta}_{X_i})}{\Pi_i(s)}, \end{aligned} \tag{7.2.17}$$

for $j = 1, \ldots, n_i$, where n_i is the number of components in subsystem i, and

$$\Pi_i(s) = \begin{cases} \pi_i^+(t_k), & \text{for } t_k < s \leq T_{k+1}, \\ n_i, & \text{for } s > T_{k+1}, \end{cases} \tag{7.2.18}$$

where $\pi_i^+(t_k)$ is the number of functioning components within subsystem i between t_k and T_{k+1}.

τ_{ij}^- is the expected failure time of Component$_{ij}$ assuming no maintenance is performed on Component$_{ij}$ at the current maintenance opportunity at time t_k, if Component$_{ij}$ degrades at the baseline rate at times $t > t_k$ (i.e., the rate when all components in subsystem i are in a working state). Hence, $\tau_{ij}^- - t_k$ is the expected RUL of Component$_{ij}$ assuming Component$_{ij}$ degrades at the baseline rate at times $t > t_k$. If Component$_{ij}$ is a working component that is replaced, $\tau_{ij}^- - t_k$ is the wasted life of Component$_{ij}$.

τ_{ij}^+ is the expected failure time of Component$_{ij}$ obtained immediately after the maintenance opportunity at t_k, if Component$_{ij}$ degrades at the baseline rate at times $t > t_k$. If Component$_{ij}$ is not replaced at t_k, then $\tau_{ij}^+ = \tau_{ij}^-$. If Component$_{ij}$ is replaced at t_k, the degradation level for Component$_{ij}$ is restored to zero ($D_{ij}^+(t_k) = 0$).

$\tau_{ij}^{+,L}$ is the expected failure time of Component$_{ij}$ obtained immediately after the maintenance opportunity at t_k, if Component$_{ij}$ degrades at the load-sharing rate between t_k and T_{k+1} and the baseline rate at times $t > T_{k+1}$. $\tau_{ij}^{+,L} - \tau_{ij}^+$ measures the expected loss in RUL due to load sharing between t_k and T_{k+1} (see Figure 7.3). If $\pi_i^+(t_k) = n_i$ (i.e, if all components in subsystem i are in a working state), then $\tau_{ij}^{+,L} = \tau_{ij}^+$. As expected, when there are no failed components in subsystem i, there will be no loss due to load sharing within subsystem i. In addition, if $\pi_i^+(t_k) \neq n_i$ and if Component$_{ij}$ receives no extra workload then $\tau_{ij}^{+,L} = \tau_{ij}^+$. In both scenarios, $X_{ij}^+(s; \boldsymbol{\theta}_{X_i}) = X_{ij}(s; \boldsymbol{\theta}_{X_i})$ and hence $\tau_{ij}^{+,L} = \tau_{ij}^+$.

The optimal action to be performed at the current time point, t_k, is the action with the smallest expected penalty, and is given by

$$u_{t_k} = \underset{u_{t_k} \in U_{t_k}}{\operatorname{argmin}} \left\{ \mathbb{E}\left[\Lambda(u_{t_k} \mid T_{k+1}, \boldsymbol{X}(t_k, \infty), \boldsymbol{D}(t_k, \infty), \boldsymbol{\theta}) \right] \right\}, \tag{7.2.19}$$

where the expectation is taken with respect to the joint distribution, $(T_{k+1}, \boldsymbol{X}(t_k, \infty), \boldsymbol{D}(t_k, \infty), \boldsymbol{\theta})$.

7.2.3 SEQUENTIAL MAINTENANCE POLICY

In this section we extend the policy from Section 7.2.2 to a sequential maintenance policy. The goal of the sequential maintenance policy is to consider the optimal way to make ξ component replacements. In order to capture all possible ways of making ξ component replacements we need to forecast $\mathcal{K}$-steps ahead.

7.2.3.1 FORECASTING $\mathcal{K}$-STEPS AHEAD

In this section we construct the penalty function that corresponds to performing $\mathcal{K}$ maintenance actions. Similarly to policies in the literature that group maintenance activities (Vu et al., 2012, 2020), the planning horizon should guarantee that all components are maintained at least once. The degradation of Component$_{ij}$ at T_{k+1} is given by

$$D_{ij}(T_{k+1}) = D^{+}_{ij}(t_k) + \int_{t_k}^{T_{k+1}} \Delta D_{ij}(t; \boldsymbol{\theta}_{D_{ij}}, X^{+}_{ij}(t; \boldsymbol{\theta}_{X_i}))dt. \tag{7.2.20}$$

At T_{k+1} we can apply $u_{T_{k+1}} \in U_{T_{k+1}}$, where $u_{T_{k+1}}$ is a set of components to be replaced and can be empty unless the system is in a failed state, and $U_{T_{k+1}}$ is the set of all possible actions. The penalty associated with $u_{T_{k+1}}$ is given by

$$\begin{aligned}\Lambda(u_{T_{k+1}} \mid T_{k+1}, T_{k+2}, \boldsymbol{X}(T_{k+1},\infty), \boldsymbol{D}(T_{k+1},\infty), \boldsymbol{\theta}) &= SI_{\{|u_{T_{k+1}}|\neq 0\}} \\ &+ \sum_{i=1}^{m}\sum_{j=1}^{n_i} \bar{c}_{ij}(\tau^{-}_{ij} - T_{k+1}) I_{\{\text{Component}_{ij} \in u_{T_{k+1}}, Y_{ij}(T_{k+1})=1\}} \\ &+ \sum_{i=1}^{m}\sum_{j=1}^{n_i} \bar{c}_{ij}(\tau^{+}_{ij} - \tau^{+,L}_{ij}) - \sum_{i=1}^{m}\sum_{j=1}^{n_i} \bar{c}_{ij}(T_{k+2} - T_{k+1}) I_{\{\text{Component}_{ij} \notin u_{T_{k+1}}, Y_{ij}(T_{k+1})=0\}},\end{aligned} \tag{7.2.21}$$

where T_{k+2} is the time of the $(k+2)$nd maintenance opportunity, and $Y_{ij}(T_{k+1})$ is the failure state of Component$_{ij}$ prior to maintenance at T_{k+1}. The expected failure times, τ^{-}_{ij}, τ^{+}_{ij}, and $\tau^{+,L}_{ij}$ are obtained as in Section 7.2.2 but replacing t_k with T_{k+1} and T_{k+1} with T_{k+2}.

The total penalty for the set $(u_{t_k}, u_{T_{k+1}})$ is

$$\Lambda(u_{t_k}, u_{T_{k+1}} \mid T_{k+1}, T_{k+2}, \boldsymbol{X}(t_k,\infty), \boldsymbol{D}(t_k,\infty), \boldsymbol{\theta}) = \Lambda(u_{t_k} \mid T_{k+1}, \boldsymbol{X}(t_k,\infty), \boldsymbol{D}(t_k,\infty), \boldsymbol{\theta}) \\ + \Lambda(u_{T_{k+1}} \mid T_{k+1}, T_{k+2}, \boldsymbol{X}(T_{k+1},\infty), \boldsymbol{D}(T_{k+1},\infty), \boldsymbol{\theta}). \tag{7.2.22}$$

The process can be repeated to obtain all sets of length $\mathcal{K}$, $(u_{t_k}, \ldots, u_{T_{k+\mathcal{K}-1}})$. The associated penalty is

$$\Lambda(u_{t_k}, \ldots, u_{T_{k+\mathcal{K}-1}} \mid T_{k+1}, \ldots, T_{k+\mathcal{K}}, \boldsymbol{X}(t_k,\infty), \boldsymbol{\theta}) = \Lambda(u_{t_k} \mid T_{k+1}, \boldsymbol{X}(t_k,\infty), \boldsymbol{D}(t_k,\infty), \boldsymbol{\theta}) \\ + \sum_{i=1}^{\mathcal{K}-1} \Lambda(u_{T_{k+i}} \mid T_{k+i}, T_{k+i+1}, \boldsymbol{X}(T_{k+i},\infty), \boldsymbol{D}(T_{k+i},\infty), \boldsymbol{\theta}). \tag{7.2.23}$$

7.2.3.2 Bayesian Sequential Decision Framework

The optimal action to perform at t_k is the action with the smallest expected penalty, and is given by

$$u_{t_k} = \underset{u_{t_k} \in U_{t_k}}{\operatorname{argmin}} \Bigg\{ \mathbb{E}\Bigg[\min_{(u_{T_{k+1}}, \ldots, u_{T_{k+\mathcal{K}-1}})} \Bigg(\Lambda(u_{t_k}, u_{T_{k+1}}, \ldots, u_{T_{k+\mathcal{K}-1}} | T_{k+1}, \ldots, T_{k+\mathcal{K}}, \\ \boldsymbol{X}(t_k,\infty), \boldsymbol{D}(t_k,\infty), \boldsymbol{\theta}, |u_{T_{k+\mathcal{K}}}| \neq 0, |u_{t_k}| + |u_{T_{k+1}}| + \cdots + |u_{T_{k+\mathcal{K}}}| = \xi) \Bigg) \Bigg] \Bigg\}, \tag{7.2.24}$$

where the expectation is taken with respect to the joint distribution, $(T_{k+1}, \ldots, T_{k+\mathcal{K}}, \boldsymbol{X}(t_k,\infty), \boldsymbol{D}(t_k,\infty), \boldsymbol{\theta})$. Performing action $u_{t_k} \in U_{t_k}$ at t_k is expected to give the smallest long term penalty.

7.3 Conclusions

We began this chapter by identifying four different types of penalties observed for multi-component systems with economic dependence through a fixed set-up cost and stochastic dependence through failure-based load sharing. We then derived a one-step ahead CBM policy which incorporated a loss-based utility function in a Bayesian framework. We then extended the CBM policy to a sequential decision policy that minimises the long-term expected loss of a sequence of maintenance actions. In Chapter

8, we compare the loss-based utility function presented in this chapter to the commonly used cost per unit time utility function. In addition, in Chapter 8, we will compare a random-threshold approach to a fixed-threshold approach and an expected failure time approach to highlight the importance of incorporating all uncertainty when making maintenance decisions.

Chapter 8

Simulation Studies

In this chapter we implement the CBM policy proposed in Chapter 7. We consider two different systems; a three-component parallel system with identical components (Section 8.2), and a six-component series-parallel power plant system (Section 8.3) from Vu et al. (2020) (the degradation processes and model parameters used in our study are different).

The CBM policy is compared to alternative policies to highlight the importance of a sequential CBM policy that incorporates both economic and stochastic dependence and all types of uncertainty present; and to investigate whether it has advantages over the cost per unit time policy.

8.1 Alternative Maintenance Policies

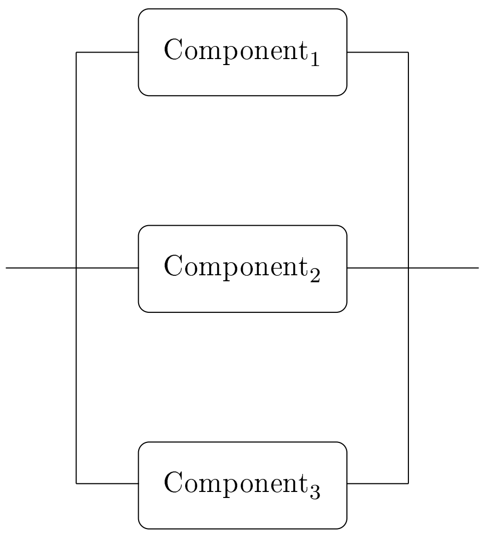

Figure 8.1: A parallel system with three identical components.

The first alternative maintenance policy, referred to as the individual maintenance policy, replaces every component immediately upon failure. By replacing components immediately upon failure, the individual maintenance policy results in no loss, in terms of cost, due to load sharing. However, the individual maintenance policy will pay the set-up cost, S, with the highest frequency.

The second policy, referred to as the simultaneous maintenance policy, replaces all components at the same time, at system failure. The simultaneous maintenance policy will pay the set-up cost, S, as infrequently as possible. However, the simultaneous maintenance policy will result in the maximum loss, in terms of cost, due to load sharing.

The third policy, referred to as the CPT policy, decides which action to perform by minimising cost per unit time (Vu et al., 2020). This utility is commonly used in the literature to schedule maintenance. We adapt the approach used by Vu et al. (2020) to calculate the cost per unit time of a sequence of actions. The cost per unit time is the sum of the set-up costs and the cost of ξ component replacements over the time interval $I = [a, b]$. The start of the time interval, a, is the time of the current maintenance opportunity, and the end of the time interval, b, is the time of the first component failure after the final maintenance opportunity. Figure 8.2 gives two sequences of $\xi = 6$ component replacements, for the system shown in Figure 8.1, from the point of view of the current maintenance opportunity. The interval definitions will allow us to compare how the cost per unit time over the interval $I = [a, b]$ compares to the total penalty over $I = [a, b]$. We note that b is random and will vary from sequence to sequence. The first sequence is

$$\begin{aligned}\zeta_1 = \big\{&\phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\},\\ &\phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\}\big\},\end{aligned} \tag{8.1.1}$$

and the second sequence is

$$\zeta_2 = \big\{\text{Component}_1, \text{Component}_2, \text{Component}_3, \text{Component}_3, \text{Component}_1, \text{Component}_2\big\}. \tag{8.1.2}$$

In Figure 8.2, b_1 is the end of the time interval for the sequence ζ_1 and is the time of the first component failure after the final maintenance opportunity (where all three components were replaced at system failure); b_2 is the end of the time interval for the sequence ζ_2 and is the time of the first component failure after the final maintenance opportunity (where Component$_2$ was replaced at $t'_{2,2}$).

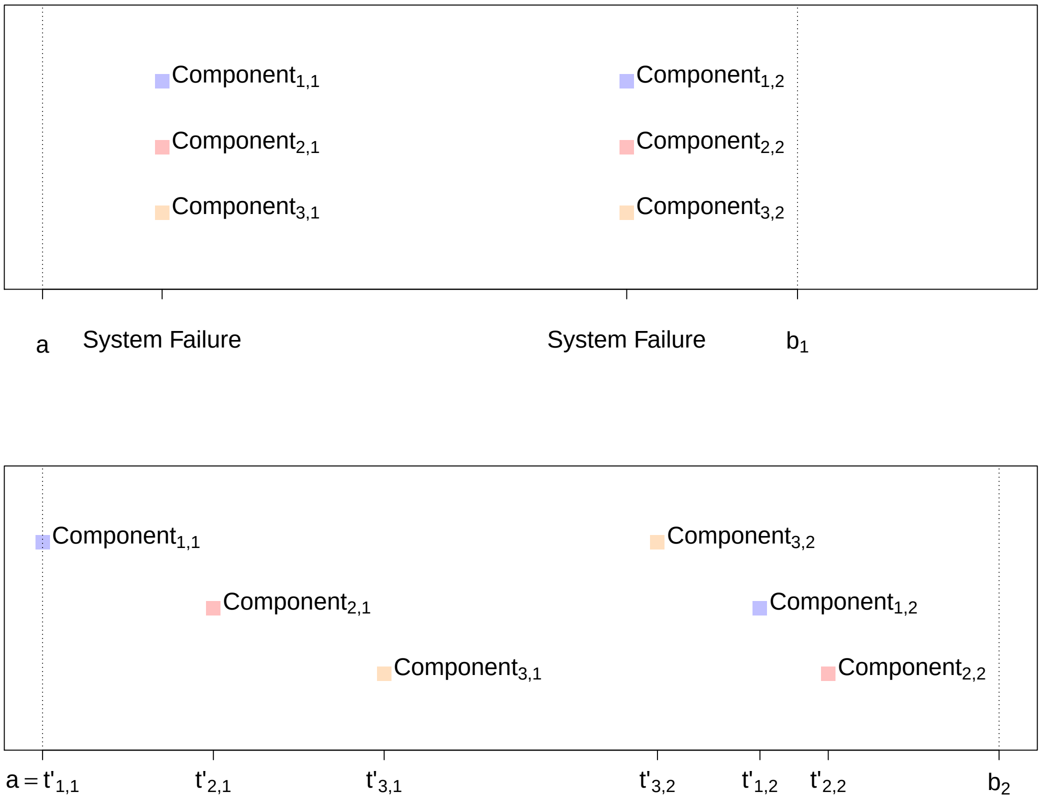

Figure 8.2: Two possible ways of making $\xi = 6$ component replacements, for the system shown in Figure 8.1, from the point of view of the failure of Component$_1$ (at a). Component$_{i,j}$ refers to the jth replacement of Component$_i$, and $t'_{i,j}$ refers to the jth failure time of Component$_i$.

The fourth policy, referred to as the threshold policy, assumes components fail when their degradation reaches a fixed threshold. The threshold policy is identical to the CBM policy, but under the threshold policy we assume components fail when their degradation reaches a fixed threshold.

Finally, several papers in the literature use the expected failure time of a component when scheduling maintenance opportunities (Vu et al., 2012, 2020). Vu et al. (2012, 2020) schedule tentative (expected) maintenance times for each component in the system, before grouping components to save expenses. This is comparable to having fixed, expected maintenance opportunities (defined using component expected maintenance times) for each component which are then grouped to minimise the penalties incurred through economic and stochastic dependence. To this end, the fifth alternative policy, referred to as the deterministic policy, will schedule maintenance times for each component based on their expected failure times, before grouping the components to minimise the total penalty. The deterministic policy is identical to the CBM policy, but under the deterministic policy we assume components fail at their expected maintenance times.

8.2 A Three-Component Parallel System with Identical Components

Figure 8.1 presents a parallel system with three identical components, labeled Component$_i$ for $i = 1, 2, 3$ (the components cost the same and degrade at the same rate, assuming a fixed workload and have the same probability of failure (distribution) assuming a fixed degradation level). The components are placed in a parallel setting, which means the system functions as long as at least one component functions. Since the components are identical, the cost of replacing Component$_i$ is c for $i = 1, 2, 3$.

8.2.1 Deterioration Process

The system receives workload continuously, $X_{\text{sys}}(t; \boldsymbol{\theta}_X)$, and this workload is split evenly between all working components. The following model for $X_{\text{sys}}(t; \boldsymbol{\theta}_X)$ is used

$$X_{\text{sys}}(t; \boldsymbol{\theta}_X) \sim \text{TN}(\mu_X, \sigma_X, \alpha_X, \beta_X), \tag{8.2.1}$$

where TN is the truncated normal distribution and $\boldsymbol{\theta}_X = (\mu_X, \sigma_X, \alpha_X, \beta_X)$, where μ_X is the mean, σ_X is the standard deviation and α_X and β_X are lower and upper bounds, respectively.

The increase in degradation of Component$_i$ due to $X_{\text{sys}}(t; \boldsymbol{\theta}_X)$ is modelled using redundant dependency, as in Keizer et al. (2018) and Yu et al. (2007). More specifically, the increase in degradation of Component$_i$ due to $X_{\text{sys}}(t; \boldsymbol{\theta}_X)$ is

$$\Delta D_i(t) = \Delta D_i(t; \boldsymbol{\theta}_D, X_{\text{sys}}(t; \boldsymbol{\theta}_X), \Pi(t)) = \frac{X_{\text{sys}}(t; \boldsymbol{\theta}_X)}{\Pi(t)^\rho}, \tag{8.2.2}$$

for $i = 1, \ldots, 3$, where $X_{\text{sys}}(t; \boldsymbol{\theta}_X)$ is the system workload at time t, $\Pi(t)$ is the number of working components at time t, and $\boldsymbol{\theta}_D = \rho \geq 0$ is a load sharing factor. The larger the value of ρ the stronger the influence a failed component has on the failure rates of the remaining components. When all components are functioning the system degrades at a rate that is proportional to $1/3^\rho$. This is defined as the baseline rate.

The total degradation at time t is

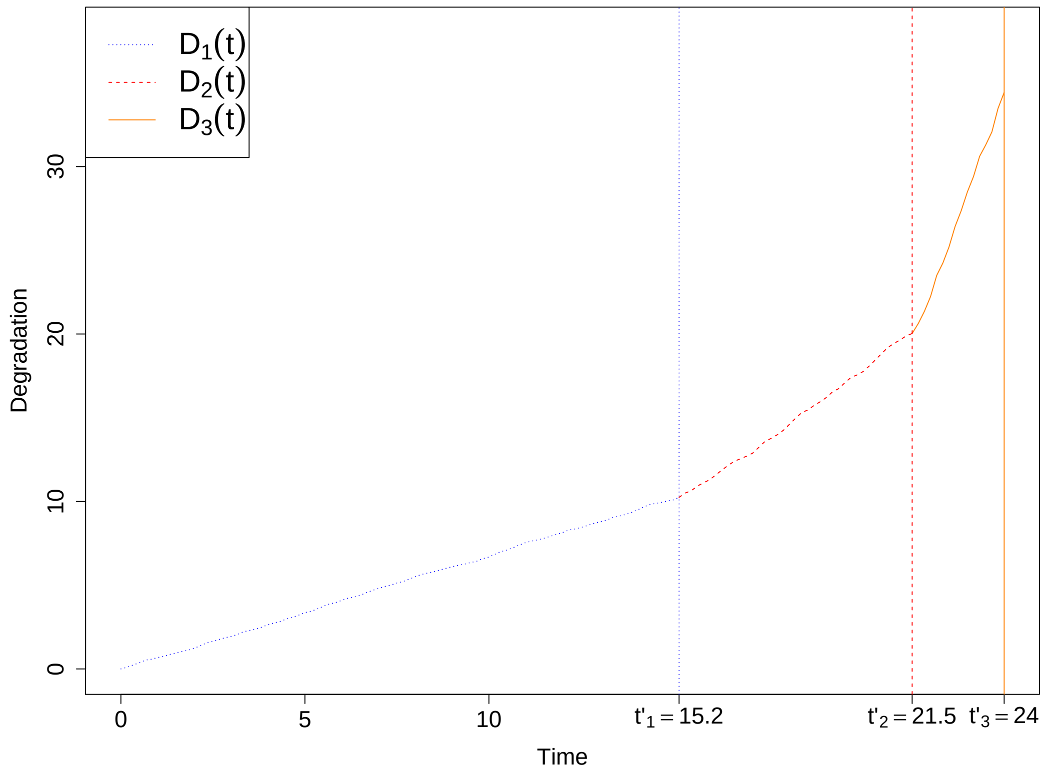

Figure 8.3: Observed degradation paths for each component. The blue (dotted) line shows the degradation path of Component$_1$, the red (dashed) line shows the degradation path of Component$_2$, and the orange (bold) line shows the degradation path of Component$_3$. Component$_1$ is observed to fail at t'_1, Component$_2$ is observed to fail at t'_2, and Component$_3$ is observed to fail at t'_3. We note that the degradation paths for all components are the same for $t \in [0, t'_1]$, and the degradation paths for Component$_2$ and Component$_3$ are the same for $t \in (t'_1, t'_2]$.

$$D_i(t) = \int_{t_i^*}^{t} \Delta D_i(s) ds, \tag{8.2.3}$$

for $i = 1, 2, 3$, where t_i^* is the time of the most recent replacement of Component$_i$.

8.2.2 SIMULATION STUDY DETAILS

μ_X	σ_X	α_X	β_X	ρ	μ_L	σ_L	α_L	β_L	c	S	ξ
1	1	0.5	1.5	2	20	$\{2, 4, 6, 8\}$	0	40	150	$\{20, 40, 60, 80\}$	$\{3, 6\}$

Table 8.1: Model parameters, component cost, c, set-up cost, S, and ξ.

Table 8.1 summarises the values of the model parameters used in this simulation study alongside component cost, c, the set-up cost, S, and the number of replacements considered in the CBM policy, ξ. The value of the set-up cost, S, the standard deviation of the random threshold, σ_L, and the number of component replacements considered in the sequential maintenance policy, ξ, will be varied in the simulation study.

The parameter σ_L describes the stochastic variation in the random failure threshold, L. As σ_L tends to 0, the random failure threshold tends to a fixed failure threshold. As σ_L increases, the stochastic variation in the random failure threshold increases. In general, even identical components will not fail at the same degradation level and simulating data with varying amounts of stochastic variation in the random failure threshold will represent certain real life systems. We expect policies that use fixed thresholds and policies that use expected maintenance times to perform poorly, compared to policies that incorporate random thresholds, as σ_L increases. We vary σ_L in this simulation study to highlight the importance of a random failure threshold.

In addition, the set-up cost, S, will vary from system to system. For example, maintenance of a failed hard drive requires an on-site maintenance worker to simply remove the failed drive and replace it with a new one. On the other hand, maintenance of off-shore wind farms, subsea power cables, or underground pipelines, are more complicated operations and require specialist equipment. The benefits of grouping component replacements will depend crucially on the set-up cost, S, which will be application dependent. We vary the set-up cost, S, in this simulation study to compare the benefits of grouping component replacements for systems with relatively expensive set-up costs and for systems with relatively cheap set-up costs.

Furthermore, the number of component replacements considered in the sequential maintenance policy, ξ, will be varied in the simulation study to examine the sequential nature of maintenance. Policies that do not consider the long term impacts of maintenance decisions can result in expensive long term maintenance costs.

Figure 8.3 gives an illustration of the component degradation process between system initiation and system failure (assuming no maintenance actions are taken) using the parameters shown in Table 8.1, with $\sigma_L = 6$. The degradation paths for all components are the same for $t \in [0, t'_1]$. At t'_1 Component$_1$ is observed to fail and hence the degradation rates increase for Component$_2$ and Component$_3$. Similarly, Component$_2$ is observed to fail at t'_2, and consequently, the degradation rate for Component$_3$ increases again. Component$_3$, and hence the system, are observed to fail at t'_3.

Component$_i$, $i = 1, 2, 3$, fails at time T'_i when the degradation reaches a random threshold L. All components have the same random threshold distribution since all components are

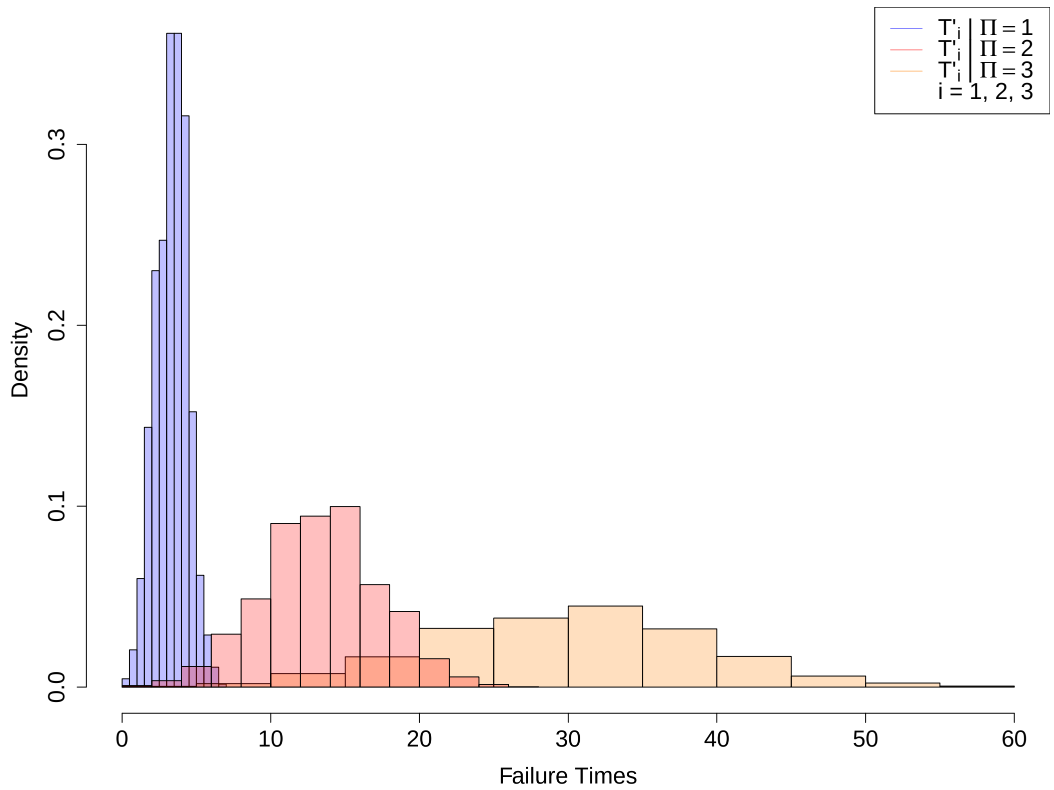

Figure 8.4: RUL distributions for a new component. The blue (left) histogram shows the RUL distribution for a new component conditional on only one component functioning in the three-component parallel system. The red (middle) histogram shows the RUL distribution for a new component conditional on two components functioning. The orange (right) histogram shows the RUL distribution for a new component degrading at the baseline rate.

identical. That is, $L = D_i(T_i')$, where T_i' is the failure time of Component$_i$, for $i = 1, 2, 3$. The cdf of the degradation threshold, L, is given by

$$F_L(D_i(t; \boldsymbol{\theta}_D, \boldsymbol{\theta}_X); \boldsymbol{\theta}_L) = F_L(D_i(t); \mu_L, \sigma_L, \alpha_L, \beta_L), \tag{8.2.4}$$

where $\boldsymbol{\theta}_L = (\mu_L, \sigma_L, \alpha_L, \beta_L)$, and $F_L(\cdot)$ is the cdf of the truncated normal distribution, with mean μ_L, standard deviation, σ_L, and lower and upper bounds α_L and β_L, respectively.

Figure 8.4 shows the RUL distribution for an example new component based on 10,000 simulations using the parameters shown in Table 8.1 (with $\sigma_L = 6$) and Equations (8.2.1) - (8.2.4). The blue (left) histogram shows the RUL distribution for a new component conditional on only one component functioning in the three-component parallel system. The red (middle) histogram shows the RUL distribution for a new component conditional

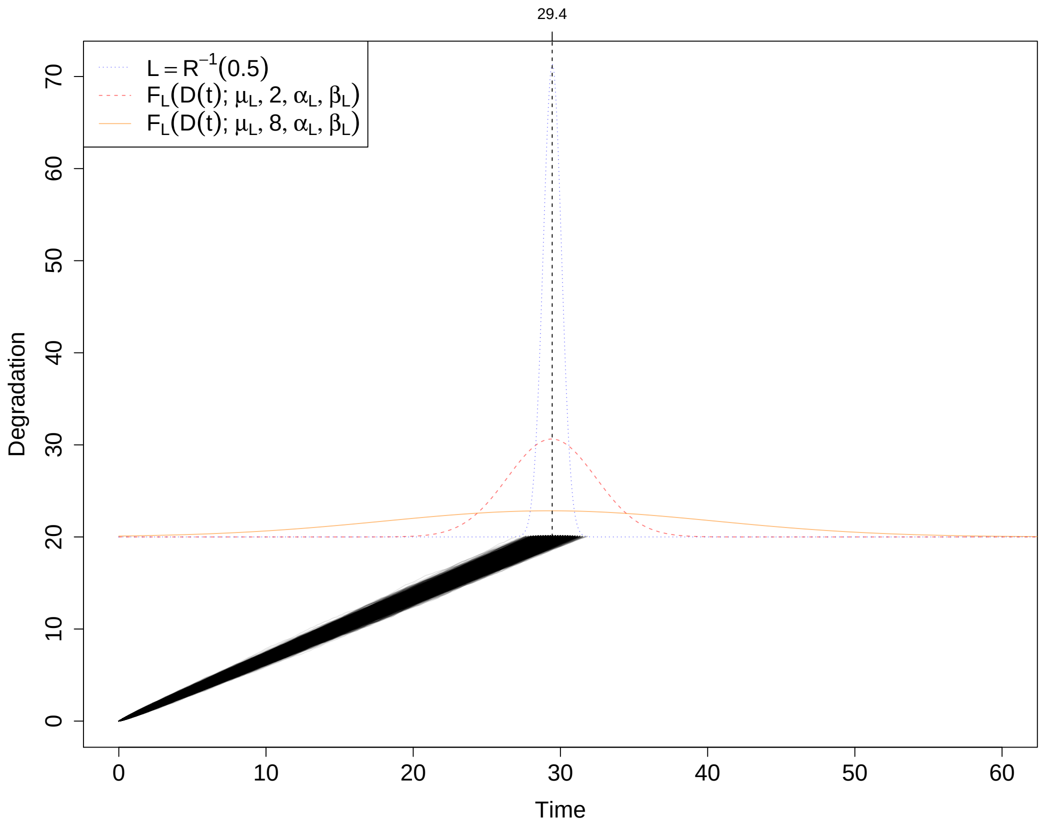

Figure 8.5: RUL distributions for a new component under different policies. The blue (dotted) line shows the RUL distribution under the threshold policy, the red (dashed) line shows the RUL distribution under the CBM policy when $\sigma_L = 2$, and the orange (bold) line shows the RUL distribution under the CBM policy when $\sigma_L = 8$. The black (dashed) line at $t = 29.4$ shows the expected failure time of a new component (i.e., the RUL distribution under the deterministic policy). The black (bold) lines depict the degradation distributions at each time point.

on two components functioning. The orange (right) histogram shows the RUL distribution for a new component degrading at the baseline rate.

Given the parameter values chosen, the expected lifetime for a new component degrading at the baseline rate is 30.1 time units, the expected lifetime for a new component conditional on two components functioning is 13.4 time units, and the expected lifetime for a new component conditional on only one component functioning is 3.3 time units. Consequently, we obtain the expected component cost per unit time, $\bar{c} = c/30.1 = 5.0$. Note, if the workload or environment change dynamically then the expected cost per unit time, $\bar{c}$, will need to be updated on a regular basis. For example, $\bar{c}$ can be updated before every maintenance decision (see Algorithm A.1).

Algorithm A.1, in Appendix A.2, is used to obtain the optimal action to be performed at a maintenance opportunity.

8.2.3 Results

We compare the long-term system cost per unit time under the CBM policy to the five alternative policies. Each simulation is run until 45 component replacements have been made in order to give an idea of the long-term cost per unit time under each policy. One thousand simulations are performed for each combination of parameters. The system cost per unit time for the proposed CBM policy (policy 1), the individual maintenance policy (policy 2), the simultaneous maintenance policy (policy 3), the CPT policy (policy 4), the threshold policy (policy 5), and the deterministic policy (policy 6) are denoted by $\mathrm{CPT}_1(\xi)$, CPT_2, CPT_3, CPT_4, CPT_5, and CPT_6, respectively.

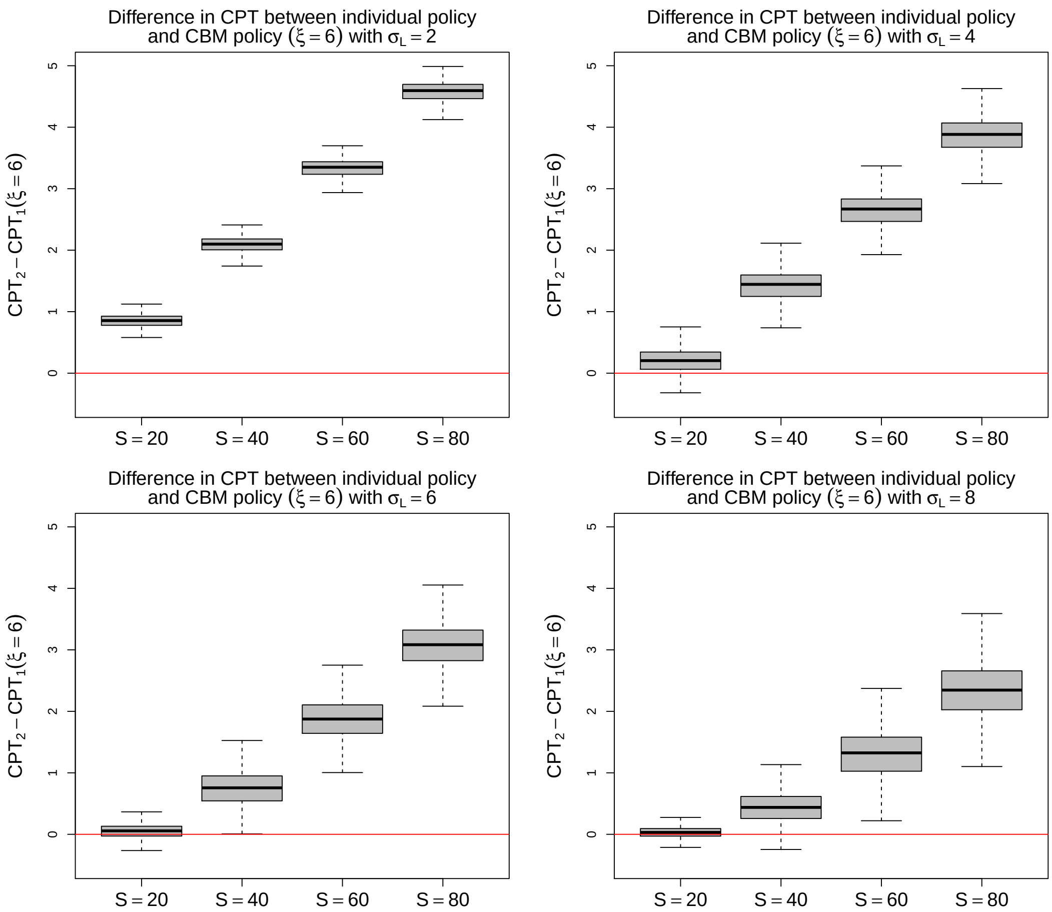

Figure 8.6: Boxplots showing the difference in CPT between the individual maintenance policy (CPT_2) and the CBM policy ($\mathrm{CPT}_1(\xi = 6)$) for different values of S, with $\sigma_L = 2$ (top left), $\sigma_L = 4$ (top right), $\sigma_L = 6$ (bottom left), and $\sigma_L = 8$ (bottom right).

Table 8.2 shows the average run-time for a maintenance decision under each policy. Each maintenance decision was run using R parallel computing with 20 cores with Intel CPU (Xeon, E5-2699 v4, base frequency 2.2GHz) through the Slurm workload manager (Slurm, 2022) on Newcastle University's high performance computing service, Rocket.

Policy	Average run-time (seconds)
CBM policy ($\xi = 6$)	35.7
CBM policy ($\xi = 3$)	4.4
CPT policy	35.3
Threshold policy	5.3
Deterministic policy	105.1

Table 8.2: Average run-time for a maintenance decision under each policy.

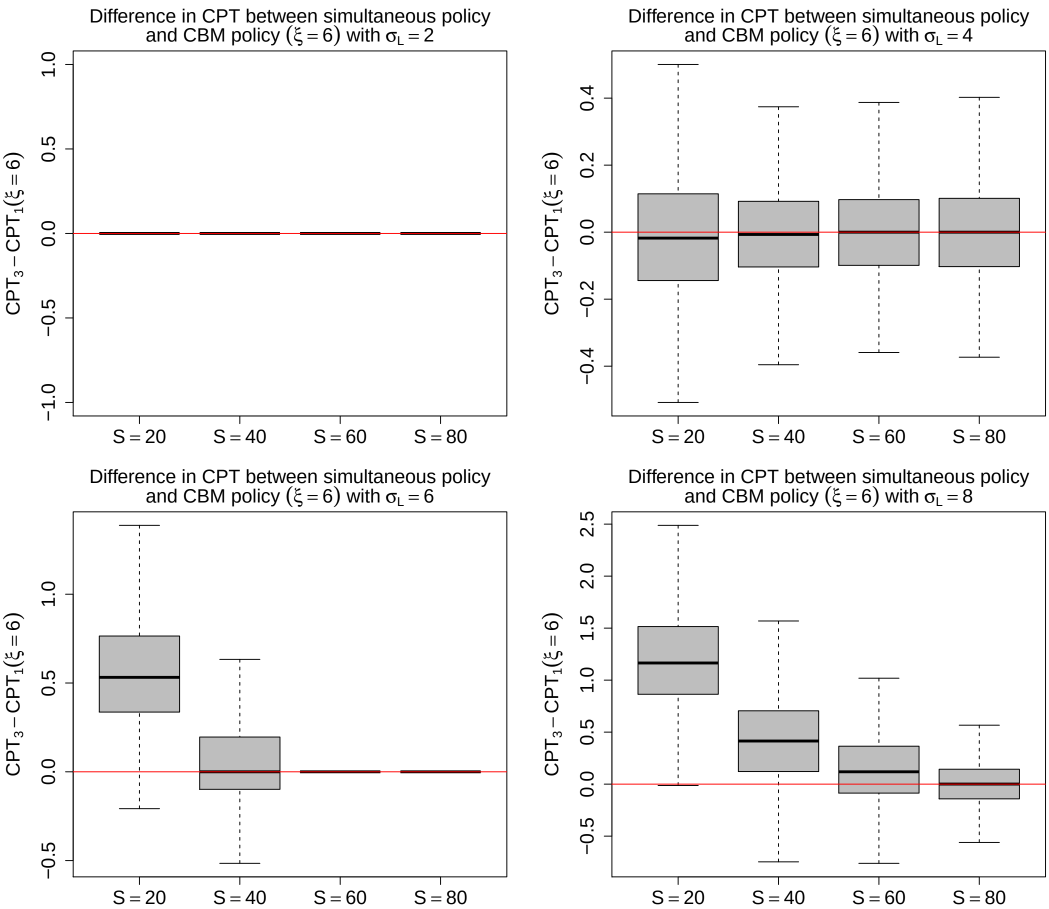

Figure 8.7: Boxplots showing the difference in CPT between the simultaneous maintenance policy (CPT_3) and the CBM policy ($\text{CPT}_1(\xi = 6)$) for different values of S, with $\sigma_L = 2$ (top left), $\sigma_L = 4$ (top right), $\sigma_L = 6$ (bottom left), and $\sigma_L = 8$ (bottom right). The scales of the y-axis vary between plots.

8.2.3.1 CBM Policy ($\xi = 6$) versus Individual Maintenance Policy

Figure 8.6 shows the difference in CPT between the individual maintenance policy and the CBM policy for different values of S, with $\sigma_L = 2$ (top left), $\sigma_L = 4$ (top right), $\sigma_L = 6$ (bottom left), and $\sigma_L = 8$ (bottom right).

From Figure 8.6 we can see that the difference in CPT between the individual maintenance

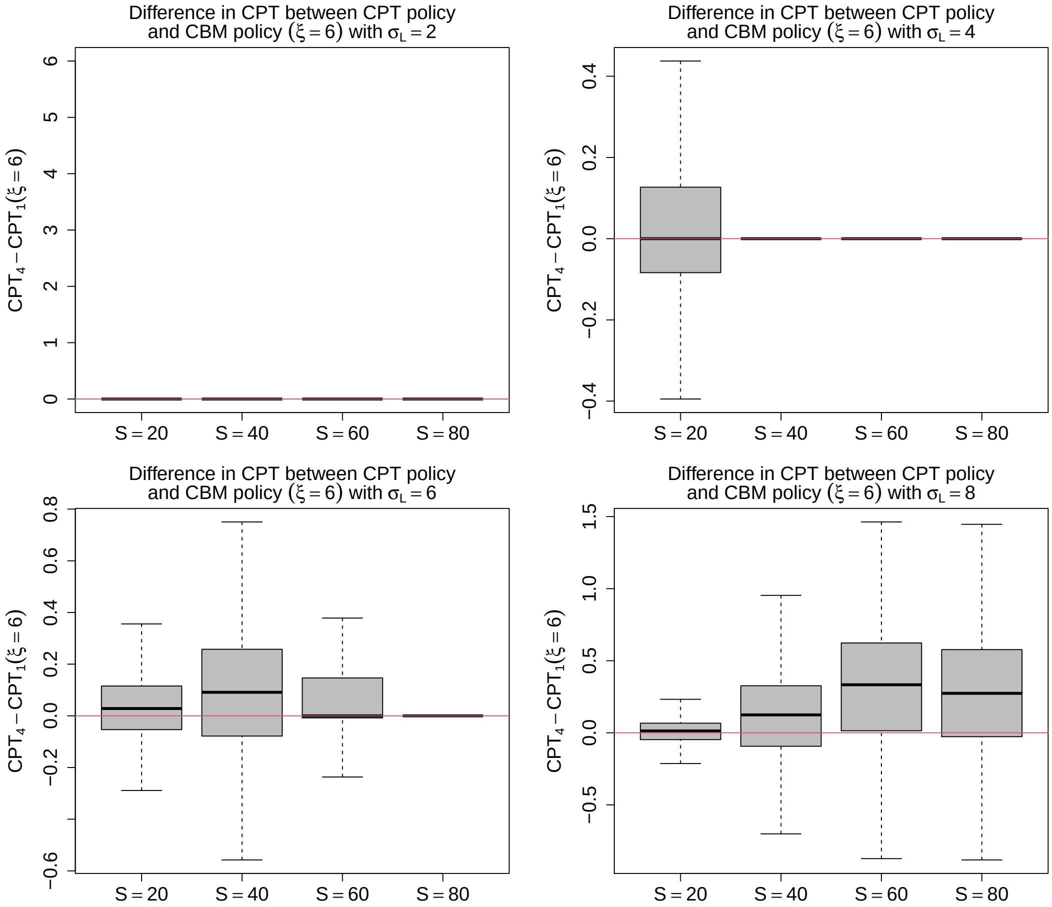

Figure 8.8: Boxplots showing the difference in CPT between the CPT policy (CPT_4) and the CBM policy ($\mathrm{CPT}_1(\xi = 6)$) for different values of S, with $\sigma_L = 2$ (top left), $\sigma_L = 4$ (top right), $\sigma_L = 6$ (bottom left), and $\sigma_L = 8$ (bottom right). The scales of the y-axis vary between plots.

policy and the CBM policy increases as S increases, for all values of σ_L, as expected. As S increases there is more incentive to cluster replacements in order to pay the set-up cost, S, as infrequently as possible. Consequently, the CBM policy outperforms the individual policy, which replaces components as soon as they fail.

The difference in CPT between the two policies decreases as σ_L increases. When $S = 20$, the policies perform approximately the same when $\sigma_L = 8$, however, the CBM policy clearly outperforms the individual policy when $\sigma_L = 2$. This is because when $\sigma_L = 2$ there is a smaller amount of stochastic variation in the component failure times. Consequently, component failure times are clustered closely together (see Figure 8.5). This gives incentive to replace components at the same time, at system failure, to result in paying the set-up cost, S, as infrequently as possible. In addition, since component failure times are clustered closely together, this results in only small amounts of loss due to load sharing.

In contrast, when $\sigma_L = 8$, component failure times are highly stochastic (see Figure 8.5), and hence are less likely to be clustered together. Component failure times will often be far apart and leaving components in a failed state will result in large amounts of loss due to load sharing. Consequently, when $\sigma_L = 8$ (and $S = 20$), there is incentive to replace components as soon as they fail, since the set-up cost is small relative to the load sharing penalties.

8.2.3.2 CBM Policy ($\xi = 6$) versus Simultaneous Maintenance Policy

Figure 8.7 shows the difference in CPT between the simultaneous maintenance policy and the CBM policy for different values of S and σ_L.

From Figure 8.7 we can see that the difference in CPT between the simultaneous maintenance policy and the CBM policy increases as S decreases, for $\sigma_L = 6, 8$, as expected. As S decreases there is more incentive to replace components as soon as they fail in order to prevent loss due to load sharing. Consequently, the CBM policy outperforms the simultaneous policy, which replaces components at system failure.

There is less incentive to cluster component replacements at system failure when there is a large amount of stochastic variation in component failure times. In contrast, when $\sigma_L = 2, 4$, the policies perform approximately equally for all values of S. This is because it is more cost effective to pay the set-up cost, S, as infrequently as possible when component failure times are clustered more closely together (see Figure 8.5).

8.2.3.3 CBM Policy ($\xi = 6$) versus CPT Policy

Figure 8.8 shows the difference in CPT between the CPT policy and the CBM policy for different values of S and σ_L. From Figure 8.8 we can see that the CBM policy and the CPT policy perform approximately the same for all values of S for $\sigma_L = 2, 4$. The CBM policy outperforms the CPT policy for $\sigma_L = 6$ and $S = 40$ and for $\sigma_L = 8$ and $S = 40, 60, 80$. Moreover, as mentioned in Section 1.2 there are potential benefits to writing the loss of each type of dependence separately in the loss-based utility function.

In addition, the loss (or reward) of a maintenance action is well defined. The loss depends on expected component failure times (immediately before and immediately after maintenance), the time of the subsequent maintenance opportunity and the loss between the current and subsequent maintenance opportunity (see Equation (7.2.9)); which can all be calculated at each maintenance opportunity. Hence the loss of a maintenance action is revealed at the time the action is performed. Consider the three-component parallel system in Figure 8.1, and consider the following sequence of $\xi = 6$ component replacements

$$\zeta = \big\{\phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\},\\ \phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\}\big\}. \tag{8.2.5}$$

This sequence of maintenance actions is depicted in Figure 8.9. At time $t'_{1,1}$, Component$_1$ fails and we perform no maintenance, and obtain a loss, $\Lambda_1(\phi)$. The next failure is Component$_2$ at $t'^{L}_{2,1}$ and we perform no maintenance. The total loss is $\Lambda(\{\phi,\phi\}) = \Lambda_1(\phi) + \Lambda_2(\phi)$. This process continues until time $t'^{L}_{2,2}$ when we replace all components $\{\text{Component}_1, \text{Component}_2, \text{Component}_3\}$. The total loss is $\Lambda(\zeta)$. At every decision, the loss is instantly revealed since it is independent of future actions; the loss due to performing ϕ at $t'_{1,1}$ is $\Lambda_1(\phi)$, regardless of what decision is made at the subsequent maintenance opportunity. In addition, the total loss is additive. At each maintenance opportunity, we are penalised if we pay the set-up cost, if we replace components in a working state, for loss due to failure-based load sharing, and we are rewarded due to postponing maintenance of failed components.

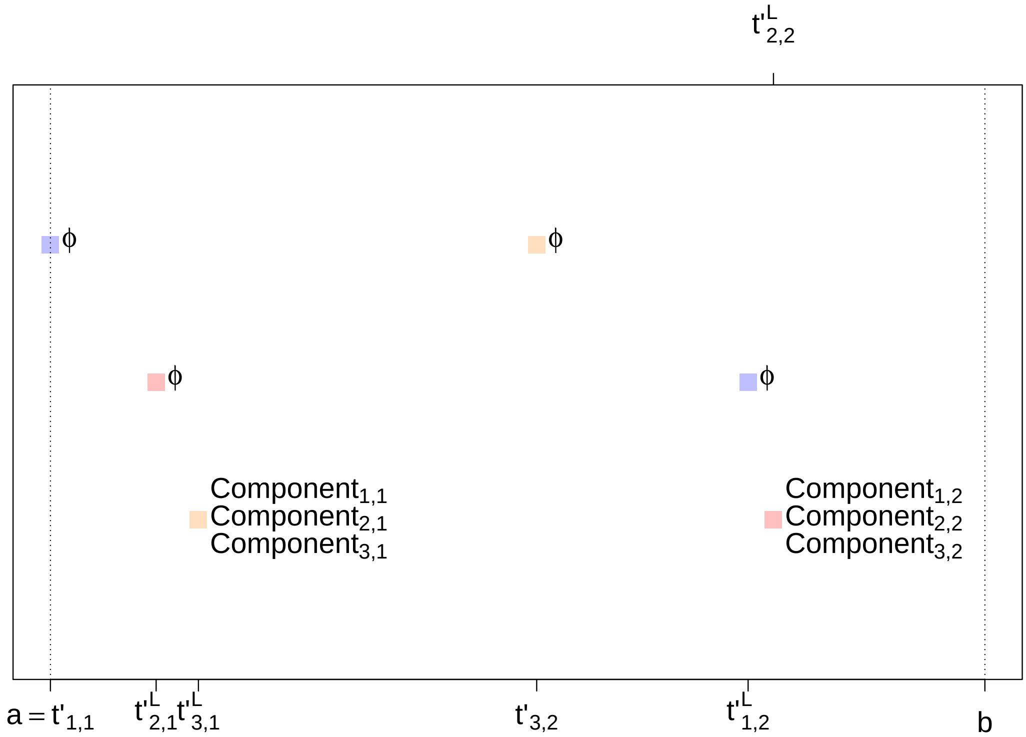

Figure 8.9: One possible way of making $\xi = 6$ component replacements, for the system shown in Figure 8.1, from the point of view of the failure of Component$_1$ (at a). Component$_{i,j}$ refers to the jth replacement of Component$_i$, and $t'_{i,j}$ refers to the jth failure time of Component$_i$.

The cost-per-unit time utility isn't as intuitive. For example, the cost of ζ is $2S + 6c$ (since we replace six identical components and pay two-set up costs). But what is the "reward" of paying this cost? The reward, under the cost-per-unit time utility, is the amount of system time we gain from paying a maintenance cost. The "system time" that we gain from paying the maintenance cost is difficult to define. What is the time reward due to performing ζ? We define the time interval, $I = [a, b]$, where a is the time of the first maintenance opportunity in the sequence ($t'_{1,1}$), and b is the end of the time interval, taken to be the first failure time after the final maintenance opportunity in the sequence, see Figure 8.9. This is consistent with the interval considered in the loss-based approach. We obtain

$$\mathrm{CPT}(\zeta) = \frac{2S + 6C}{b - a}. \tag{8.2.6}$$

This is not as intuitive as the loss-based penalty because the time-reward due to replacing all components at $t'_{3,2}$ has not yet been fully revealed, as two of the replaced components are still in a working state. The reward due to replacing all components at $t'_{3,2}$ requires us to forecast further ahead, but to forecast further ahead we need to make a decision at each future maintenance opportunity. Since the full reward of maintenance is not revealed this may result in suboptimal decisions being made since the proportion of the full reward that is revealed can vary for different sequences of length ξ; or it may require us to forecast further ahead to reveal the benefit of a sequence of actions. It can also be viewed as the full time-reward of replacing all components at $t'_{3,1}$ being fully revealed at time $t'_{3,2}$, whereas the loss due to replacing all components at $t'_{3,1}$ only requires us to forecast until the subsequent maintenance opportunity (which is well defined) and to calculate the expected component failure times (immediately before and immediately after maintenance). This property of the loss-based utility function could make it appealing to DRL algorithms. The reason the CPT policy performs worse for some values of σ_L and S may be due to the CPT policy not "fully revealing" the reward of maintenance.

Moreover, using a fixed-time interval, as in Vu et al. (2020), for every sequence of maintenance actions would not work in the simulation study considered in this section. Comparing different decisions using a fixed-time interval reduces to choosing the plan with the lowest cost. In this simulation study, using a fixed-time interval would result in the policy choosing the action with the smallest number of set-up costs, which would be identical to the simultaneous maintenance policy. Thus, we found that the decision that minimises cost-per-unit time over a fixed interval, will not, in general, be the optimal decision. Figure 8.10 compares two possible ways of making $\xi = 6$ component replacements, for the system shown in Figure 8.1, from the point of view of the failure of Component$_1$ (at a) using a fixed-time interval (analogous to the fixed-time intervals considered in Vu et al. (2020)).

The first sequence is

$$\begin{aligned}\zeta_1 = \big\{&\phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\},\\ &\phi, \phi, \{\text{Component}_1, \text{Component}_2, \text{Component}_3\}\big\},\end{aligned} \tag{8.2.7}$$

and the second sequence is

$$\zeta_2 = \big\{\text{Component}_1, \text{Component}_2, \text{Component}_3, \text{Component}_3, \text{Component}_1, \text{Component}_2\big\}. \tag{8.2.8}$$

The cost of ζ_1 is $2S + 6C$ and the cost of ζ_2 is $6S + 6C$. We find $\zeta_1 \leq \zeta_2$ with equality only when $S = 0$. Therefore, for any $S > 0$ the fixed-time interval approach will opt to replace all components at system failure rather than replacing components upon failure. The cost-per-unit time approach here does not take into account that the components in Figure 8.10 (top) may be more heavily degraded in the time interval compared to the components in Figure 8.10 (bottom) and may fail shortly after b, whereas $\text{Component}_{2,2}$ is brand new at b. This detail could make replacing components at system failure less appealing. Furthermore, when $S = 0$, the fixed-time interval policy would not be able to distinguish between replacing all components at system failure and replacing each component upon failure. However, when $S = 0$ replacing all components together at system failure becomes less appealing. Clustering component replacements becomes more beneficial as S increases. The drawbacks of using cost-per-unit time could be considered further in future research alongside potential issues with using a loss-based approach to obtain the maintenance decisions; since the goal for many companies will be to minimise the long-term cost - and the loss-based approach does not directly minimise this.

In summary, when using the CPT policy, the reward of a sequence of maintenance actions will, in general, not be fully revealed until further into the future (which requires further decisions to be made). In contrast, using the loss-based utility proposed in Chapter 7, the full loss (or full reward) of a maintenance action is obtained at the time an action is performed since it is independent of future actions. In the sequence, ζ, of length ξ, the third decision is to replace all components and we obtain a loss for this decision $\Lambda_3(\{\text{Component}_1, \text{Component}_2, \text{Component}_3\})$ - there will be no further loss because of this action in the future; however, we need to forecast further ahead to obtain the "system time" reward due to replacing these components.

This property of the loss-based utility could identify the optimal sequence of actions for smaller values of ξ compared to the cost-per-unit time policy. This could be considered in future work. Is it beneficial to choose a utility function (or reward function) whose rewards are instantly revealed (at the the time the action is performed) and are independent of future actions? Could this reduce computational time? Could it reduce the value of ξ

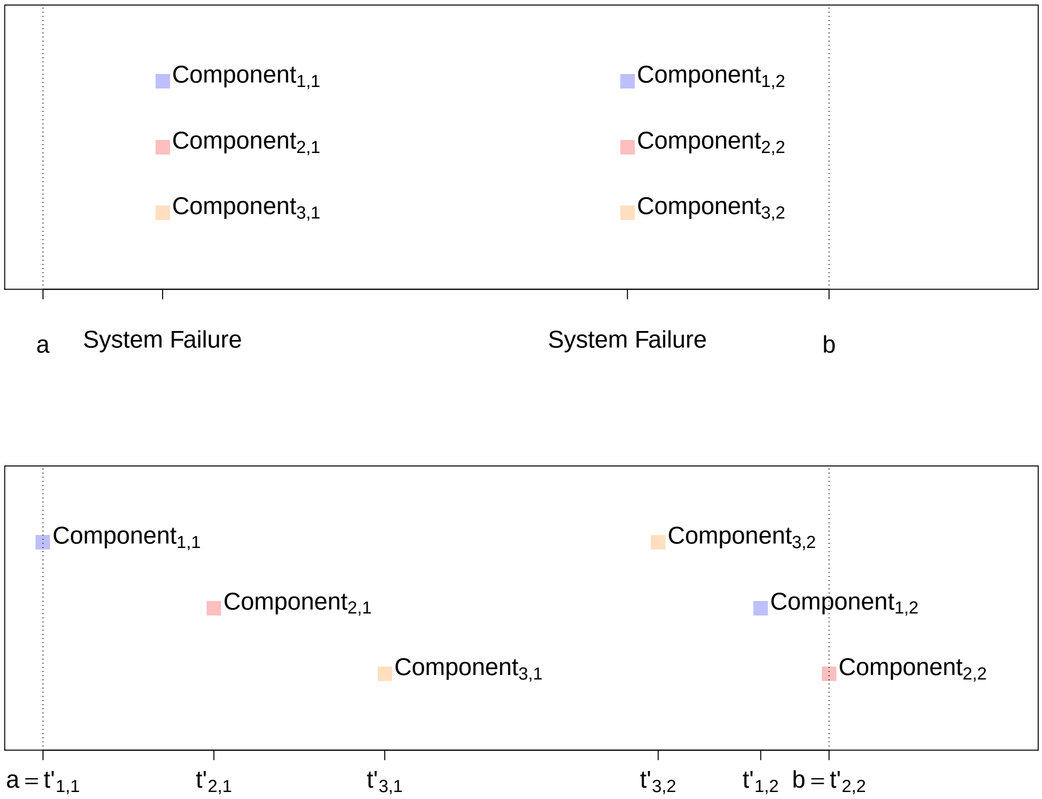

Figure 8.10: Two possible ways of making $\xi = 6$ component replacements, for the system shown in Figure 8.1, from the point of view of the failure of Component$_1$ (at a) using a fixed-time interval. Component$_{i,j}$ refers to the jth replacement of Component$_i$, and $t'_{i,j}$ refers to the jth failure time of Component$_i$.

required to obtain the optimal decision? In addition, the loss-based utility performed as well or better than the CPT utility for all values of S and all values of σ_L. Furthermore, it is not clear from the cost-per-unit time given by Equation (8.2.6), how much each type of dependence (economic dependence and failure-based load sharing dependence) contributes to the overall cost-per-unit time; whereas we can see the contribution of each type of dependence to the loss using the loss-based utility given by Equation (7.2.9).

8.2.3.4 CBM Policy ($\xi = 6$) versus Threshold Policy/Deterministic Policy

Figure 8.11 shows the difference in CPT between the threshold policy and the CBM policy for different values of S and σ_L. The CBM policy performs at least as well as the threshold policy for all values of S and all values of σ_L.

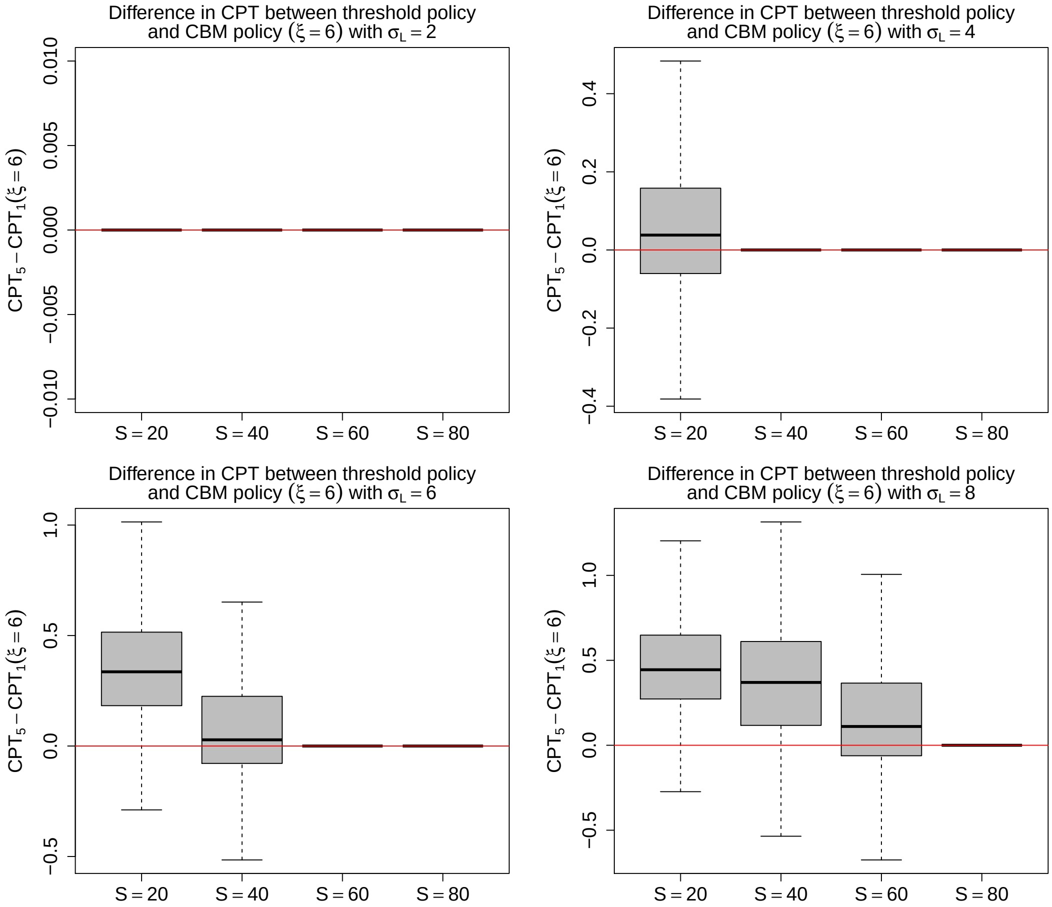

Figure 8.11: Boxplots showing the difference in CPT between the threshold policy (CPT_5) and the CBM policy ($\text{CPT}_1(\xi = 6)$) for different values of S, with $\sigma_L = 2$ (top left), $\sigma_L = 4$ (top right), $\sigma_L = 6$ (bottom left), and $\sigma_L = 8$ (bottom right). The scales of the y-axis vary between plots.

The threshold policy assumes components fail when their degradation reaches a fixed threshold (see Figure 8.5). Consequently, it ignores some of the stochastic variation in component failure times. The fixed threshold is taken to be $L = F_L^{-1}(0.5)$, where $F_L^{-1}(\cdot)$ is the inverse of the degradation threshold cdf given by Equation (8.2.4). The fixed threshold is the median degradation level at component failure times.

Under the threshold policy, it appears more beneficial to cluster all component failure times because it believes that all components will fail at approximately the same time once all degradation levels are synchronised (equal). The threshold policy naively believes that once all component degradation levels are synchronised, components can always be replaced at system failure, resulting in the set-up cost being paid as infrequently as possible, with only small amounts of load sharing (since the threshold policy predicts that all component failure times will be clustered closely together, see Figure 8.5). This is the optimal policy when component failure times are tightly clustered together.

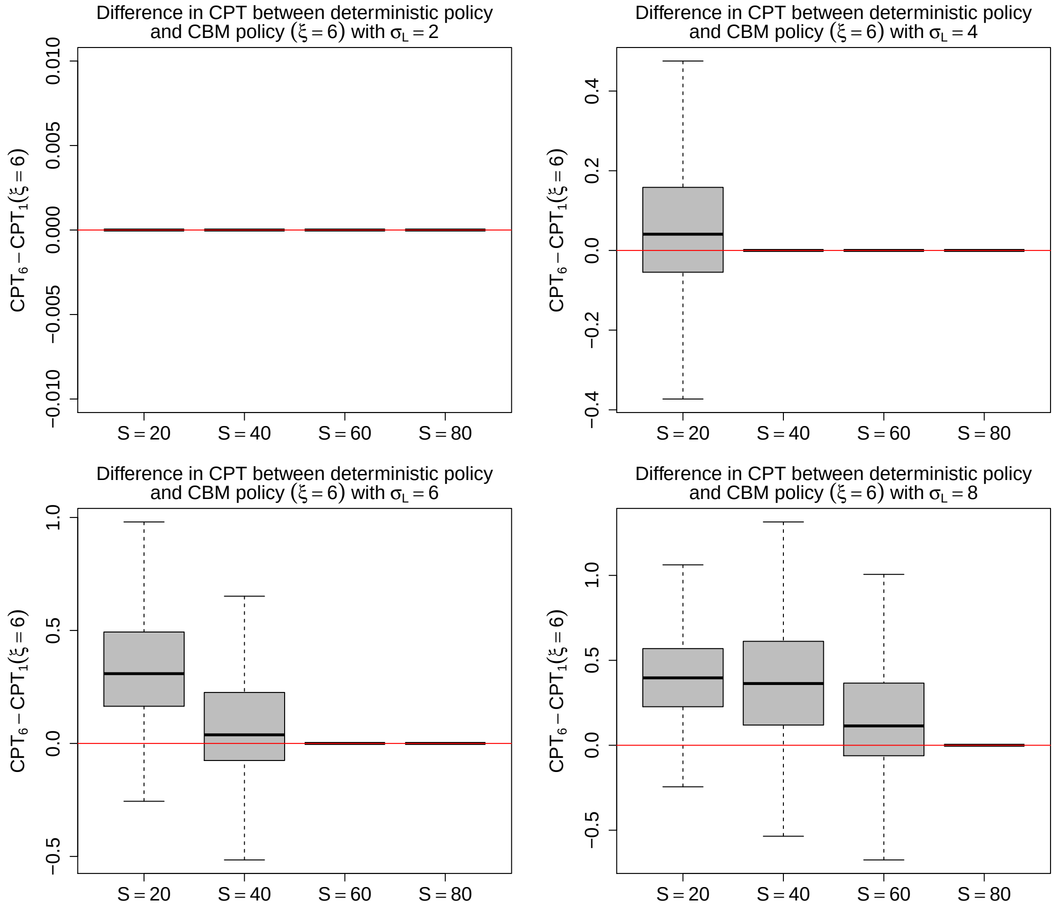

Figure 8.12: Boxplots showing the difference in CPT between the deterministic policy (CPT_6) and the CBM policy ($CPT_1(\xi = 6)$) for different values of S, with $\sigma_L = 2$ (top left), $\sigma_L = 4$ (top right), $\sigma_L = 6$ (bottom left), and $\sigma_L = 8$ (bottom right). The scales of the y-axis vary between plots.

From Figure 8.11 we can see that the threshold policy and the CBM policy perform identically when $\sigma_L = 2$. This is because when $\sigma_L = 2$, the stochastic variation in the random failure threshold is small, and hence ignoring this variation has no effect on the optimal maintenance decision at maintenance opportunities. As σ_L increases, the difference in CPT between the threshold policy and the CBM policy increases.

From Figure 8.5 we can see that the threshold RUL distribution does not capture the stochastic variation in component failure times when $\sigma_L = 8$ and hence results in a suboptimal policy. In contrast, when $\sigma_L = 2$, the threshold RUL distribution closely matches the RUL distribution with $\sigma_L = 2$.

The threshold policy is invariant to changes in σ_L and hence results in a suboptimal policy as σ_L increases. When $\sigma_L = 2$ component failure times are clustered closely together. This gives incentive to replace components at the same time, at system failure, to result in paying the set-up cost, S, as infrequently as possible. In contrast, when $\sigma_L = 8$, component failure times are highly stochastic. Leaving components in failed states will often result in large amounts of loss due to load sharing.

Figure 8.12 shows the difference in CPT between the deterministic policy and the CBM policy for different values of S and σ_L. The CBM policy performs at least as well as the deterministic policy for all values of S and all values of σ_L.

The deterministic policy assumes that all components will fail at exactly the same time once all degradation levels are synchronised (see Figure 8.5). Under the deterministic policy, it appears more beneficial to cluster all component failure times to synchronise all component failure times. Similarly to the threshold policy, the deterministic policy naively believes that once all component degradation levels are synchronised (equal), components can always be replaced at system failure, resulting in the set-up cost being paid as infrequently as possible, with no load sharing. This would be the optimal policy if there was no stochastic variation in component failure times.

Similarly to the threshold policy, the deterministic policy and the CBM policy perform identically when $\sigma_L = 2$. As σ_L increases, the difference in CPT between the deterministic policy and the CBM policy increases.

8.2.3.5 CBM Policy ($\xi = 6$) versus CBM Policy ($\xi = 3$)

Figure 8.13 shows the difference in CPT between the CBM policy with $\xi = 3$ and the CBM policy with $\xi = 6$ for different values of S and σ_L.

From Figure 8.13 we can see that the CBM policy with $\xi = 3$ and the CBM policy with $\xi = 6$ perform almost identically for all values of S and all values of σ_L. The CBM policy with $\xi = 3$ is able to identify that it is more beneficial to postpone maintenance when there is a small amount of stochastic variation in component failure times and can recognise that it is advantageous to replace components upon failure when component failure times are highly stochastic.

8.3 A Six-Component Series-Parallel System

Figure 8.14 presents a reliability block diagram of a power plant composed of six components arranged in four subsystems from Vu et al. (2020) (the degradation processes and model parameters used in our study are different). Subsystem one consists of two components, pump A and pump B, arranged in parallel; subsystem two consists of two components, pulveriser A and pulveriser B, arranged in parallel; subsystem three is a single component, the boiler; and subsystem four is a single component, the generator.

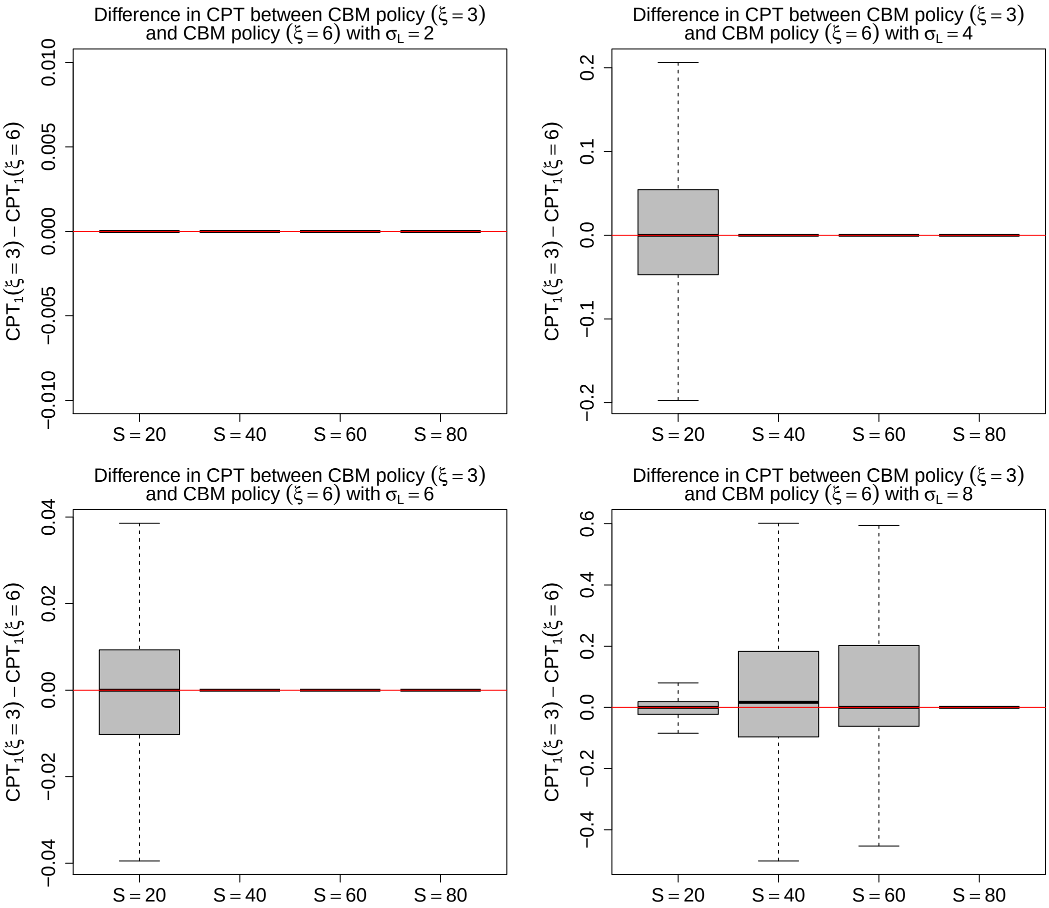

Figure 8.13: Boxplots showing the difference in CPT between the CBM policy with $\xi = 3$ ($\mathrm{CPT}_1(\xi = 3)$) and the CBM policy with $\xi = 6$ ($\mathrm{CPT}_1(\xi = 6)$) for different values of S, with $\sigma_L = 2$ (top left), $\sigma_L = 4$ (top right), $\sigma_L = 6$ (bottom left), and $\sigma_L = 8$ (bottom right). The scales of the y-axis vary between plots.

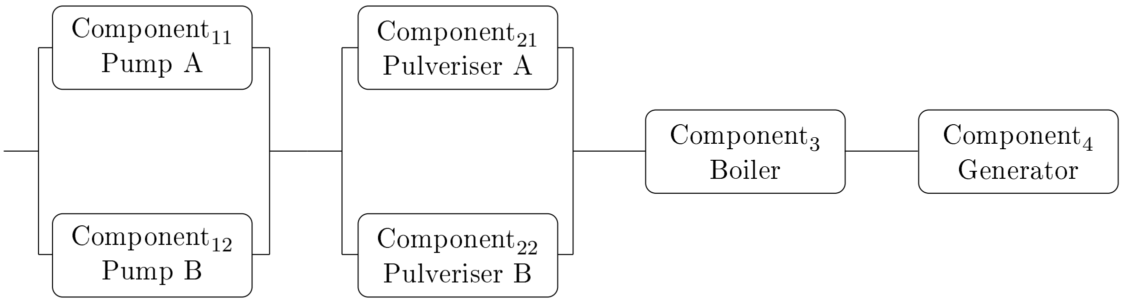

Figure 8.14: Reliability block diagram of a power plant composed of six components arranged in four subsystems from (Vu et al., 2020). Subsystem one consists of two components, pump A and pump B, arranged in parallel; subsystem two consists of two components, pulveriser A and pulveriser B, arranged in parallel; subsystem three is a single component, the boiler; and subsystem four is a single component, the generator.

Pump A, pump B, pulveriser A, and pulveriser B are non-critical components since the power plant can still operate when one of the pumps/pulverisers fails. The boiler and the generator are critical to the operation of the power plant. The number of critical components can change over time. For example, if pump A fails, pump B becomes a critical component.

In other words, subsystems one and two will function as long as one component functions. The failure of the boiler or the generator results in a system failure. The failure of a component in subsystem one or two will result in the remaining component having to sustain the whole subsystem workload; thus increasing the failure rate of the remaining working component. The failure of a component is an opportunity to perform maintenance. The system is in a failed state if at least one subsystem is in a failed state.

The pumps in subsystem one are identical (the pumps cost the same and degrade at the same rate, assuming a fixed workload) and the cost of replacing a pump is denoted c_1, but the pulverisers in subsystem two are not identical; the cost of replacing pulveriser A is c_{21} and the cost of replacing pulveriser B is c_{22}. The cost of replacing the boiler is c_3 and the cost of replacing the generator is c_4.

In this section we compare the CBM policy to the short-sighted maintenance policy for this power plant to highlight the importance of a sequential maintenance policy when making maintenance decisions.

8.3.1 Degradation Processes

Gamma processes have been widely used to describe the degradation of systems. A characteristic of Gamma processes is that they are strictly monotone increasing, which is the behavior observed in most physical deterioration processes. Moreover, the paths are discontinuous and can be thought of as the accumulation of an infinite number of small shocks.

Gamma processes will be used to model the degradation of the components in the six-component series-parallel system shown in Figure 8.14. More specifically, the increase in degradation of Component$_i$, for $i = 3, 4$, at time t is

$$\Delta D_i(t) \sim \mathrm{Ga}(\alpha_i, \beta_i), \tag{8.3.1}$$

where α_i is the shape parameter and β_i is the rate parameter. Similarly, the increase in degradation of Component$_{1i}$, for $i = 1, 2$, at time t is

$$\Delta D_{1i}(t) \mid \Pi_1(t) = j \sim \mathrm{Ga}(\alpha_{1,j}, \beta_{1,j}), \tag{8.3.2}$$

for $j = 1, 2$, where $\alpha_{1,j}$, for $j = 1, 2$, are shape parameters and $\beta_{1,j}$, for $j = 1, 2$, are rate parameters. When both components in subsystem one are functioning, Component$_{1i}$, for $i = 1, 2$, degrades at the baseline rate. Finally, the increase in degradation of Component$_{2i}$, for $i = 1, 2$, at time t is

$$\Delta D_{2i}(t) \mid \Pi_2(t) = j \sim \mathrm{Ga}(\alpha_{2i,j}, \beta_{2i,j}), \tag{8.3.3}$$

for $j = 1, 2$, where $\alpha_{2i,j}$, for $i, j = 1, 2$, are shape parameters and $\beta_{2i,j}$, for $i, j = 1, 2$, are rate parameters. The total degradation at time t is

$$D_{ij}(t) = \int_{t^*_{ij}}^{t} \Delta D_{ij}(s) ds, \tag{8.3.4}$$

for $i = 1, \ldots, 4$, and $j = 1, \ldots, n_i$, where t^*_{ij} is the time of the most recent replacement of Component$_{ij}$, and n_i is the number of components in subsystem i.

Component$_{ij}$, for $i = 1, \ldots, 4$, and $j = 1, \ldots, n_i$, fails at time T'_{ij} when the degradation reaches a random threshold, L_{ij}. That is, $L_{ij} = D_{ij}(T'_{ij})$, where T'_{ij} is the failure time of Component$_{ij}$, for $i = 1, \ldots, 4$, and $j = 1, \ldots, n_i$. Both components in subsystem one have the same random threshold distribution since they are identical. The cdf of the degradation threshold, L_{ij}, is given by

$$F_{L_{ij}}(D_{ij}(t); \boldsymbol{\theta}_{L_{ij}}) = F_{L_{ij}}(D_{ij}(t); \mu_{L,ij}, \sigma_{L,ij}, \alpha_{L,ij}, \beta_{L,ij}), \tag{8.3.5}$$

where $\boldsymbol{\theta}_{L,ij} = (\mu_{L,ij}, \sigma_{L,ij}, \alpha_{L,ij}, \beta_{L,ij})$, and $F_{L_{ij}}(\cdot)$ is the cdf of the truncated normal distribution, with mean $\mu_{L,ij}$, standard deviation, $\sigma_{L,ij}$, and lower and upper bounds $\alpha_{L,ij}$ and $\beta_{L,ij}$, respectively.

$\alpha_{1,2}$	$\beta_{1,2}$	$\alpha_{1,1}$	$\beta_{1,1}$	$\alpha_{22,2}$	$\beta_{22,2}$	$\alpha_{22,1}$	$\beta_{22,1}$	$\alpha_{21,2}$	$\beta_{21,2}$	$\alpha_{21,1}$
10	10	120	30	45	20	120	30	90	80	80
$\beta_{21,1}$	α_3	β_3	α_4	β_4	$\mu_{L,1}$	$\sigma_{L,1}$	$\alpha_{L,1}$	$\beta_{L,1}$	$\mu_{L,22}$	$\sigma_{L,22}$
25	40	20	40	20	150	60	0	300	350	140
$\alpha_{L,22}$	$\beta_{L,22}$	$\mu_{L,21}$	$\sigma_{L,21}$	$\alpha_{L,21}$	$\beta_{L,21}$	$\mu_{L,3}$	$\sigma_{L,3}$	$\alpha_{L,3}$	$\beta_{L,3}$	
0	700	150	60	0	300	350	140	0	700	
$\mu_{L,4}$	$\sigma_{L,4}$	$\alpha_{L,4}$	$\beta_{L,4}$	c_1	c_{21}	c_{22}	c_3	c_4	ξ	
350	140	0	700	150	175	160	200	200	6	

Table 8.3: Model parameters, component costs, and ξ.

Table 8.3 summarises the values of the model parameters used in this simulation study alongside component costs, and the number of replacements considered in the CBM policy,

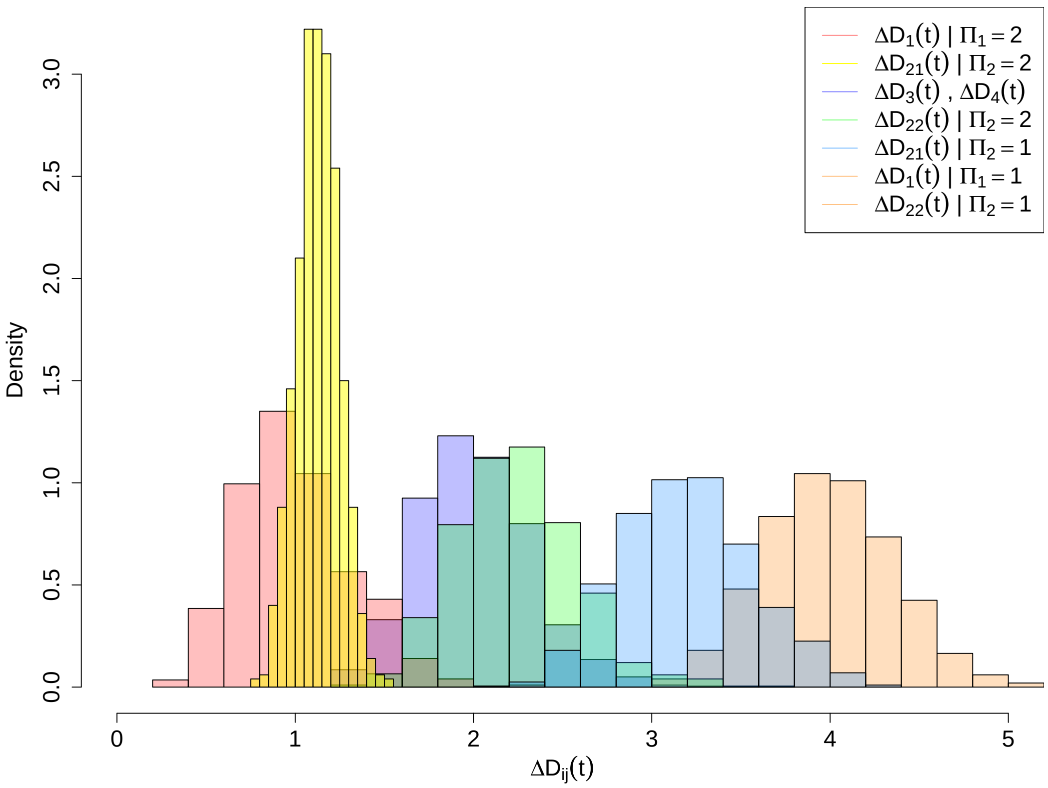

Figure 8.15: Degradation increment distributions. The degradation increment distribution of (from left to right): a pump (i.e., a component in subsystem one) conditional on both pumps working; pulveriser A conditional on both pulverisers working; the boiler/generator; pulveriser B conditional on both pulverisers working; pulveriser A conditional on pulveriser B being in a failed state; and a pump conditional on the other pump being in a failed state/pulveriser B conditional on pulveriser A being in a failed state.

ξ. The value of the set-up cost, S, will be varied in the simulation study. The parameter values are chosen to provide a large amount of uncertainty in component failure times, which will often be present in real systems as a result of the stochastic nature of component failure times and because of limited or noisy data.

Figure 8.15 shows the degradation increment distributions based on 10,000 simulations using the parameters shown in Table 8.3 and Equations (8.3.1) - (8.3.3). Figure 8.16 shows the RUL distributions for new components in each subsystem based on 10,000 simulations using the parameters shown in Table 8.3 and Equations (8.3.1) - (8.3.5). From Figure 8.16 (left) we can see that the pulverisers have approximately the same expected RUL when both pulverisers are functioning, but the effect of load sharing is more significant for pulveriser A. In addition, from Figure 8.16 we can see that the components in subsystem one/two degrade more quickly when the other component is in a failed state. Furthermore, from Figure 8.16 (right) we can see that the boiler and the generator have the longest expected

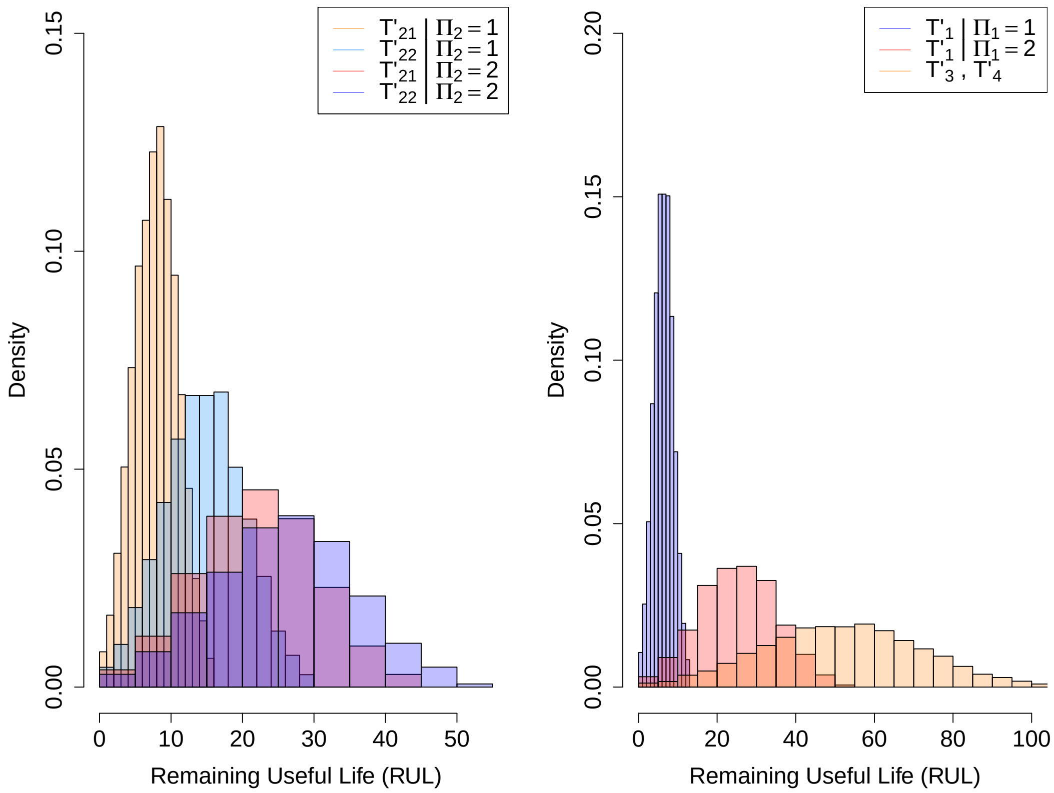

Figure 8.16: RUL distributions. Figure (left) shows the RUL distributions of (from left to right): pulveriser A conditional on pulveriser B being in a failed state; pulveriser B conditional on pulveriser A being in a failed state; pulveriser A conditional on both pulverisers working; and pulveriser B conditional on both pulverisers working. Figure (right) shows the RUL distributions of (from left to right): a pump conditional on the other pump being in a failed state; a pump conditional on both pumps working; and the boiler/generator.

lifetime; this is realistic since critical components should have the longest lifetimes.

8.3.2 Results

In this section we compare the long-term system cost per unit time under the CBM policy to the short-sighted maintenance policy. Each simulation is run until 18 component replacements have been made. One thousand simulations are performed for each value of S. The system cost per unit time for the proposed CBM policy, and the short-sighted maintenance policy, are denoted by CPT($\xi = 6$), and CPT(SS), respectively.

Table 8.4 shows the average run-time for a maintenance decision under each policy. Each maintenance decision was run using R parallel computing with 20 cores with Intel CPU

(Xeon, E5-2699 v4, base frequency 2.2GHz) through the Slurm workload manager (Slurm, 2022) on Newcastle University's high performance computing service, Rocket.

Policy	Average run-time (seconds)
CBM policy ($\xi = 6$)	163
Short-sighted policy	4

Table 8.4: Average run-time for a maintenance decision under each policy.

8.3.2.1 CBM Policy versus Short-Sighted Policy

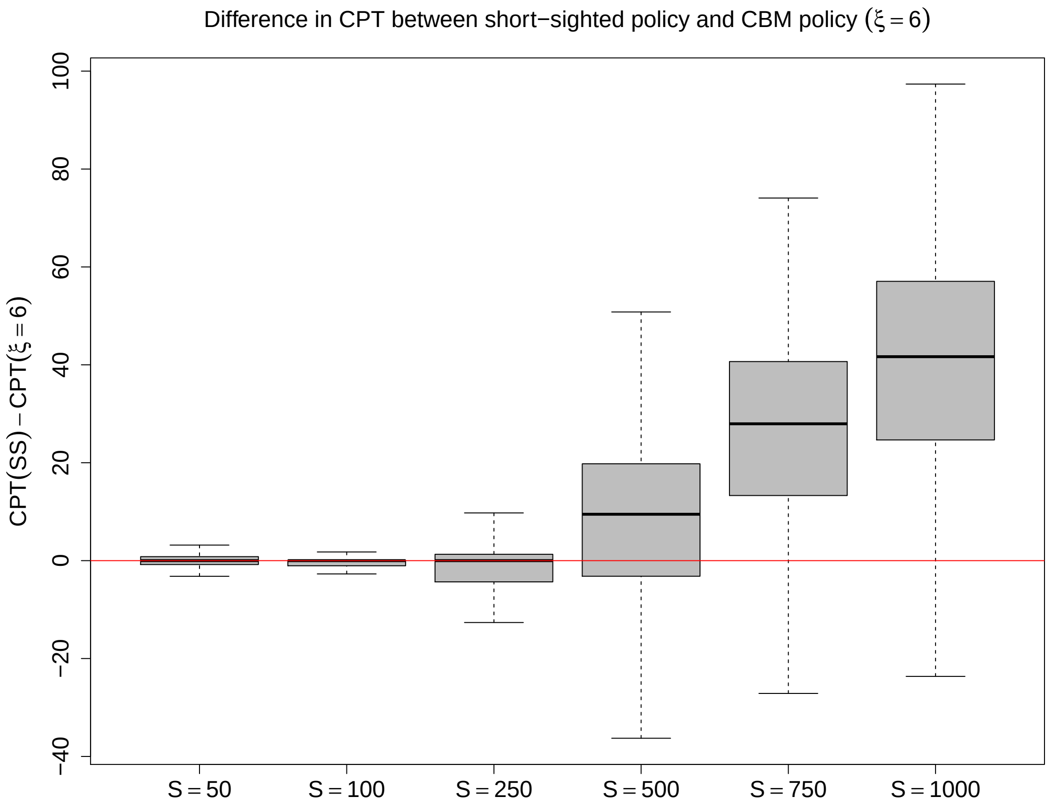

Figure 8.17: Difference in CPT between short-sighted maintenance policy (CPT(SS)) and CBM policy (CPT($\xi = 6$)) for different values of S.

Figure 8.17 shows the difference in CPT between the short-sighted maintenance policy and the CBM policy for different values of S. From Figure 8.17 we can see that the short-sighted policy and the CBM policy perform approximately the same when $S = \{50, 100, 250\}$. When the set-up cost is small it is usually beneficial to simply replace components upon failure, or to replace failed components upon system failure. Both policies are able to recognise that this is the optimal maintenance strategy. However, as S increases the difference in CPT between the short-sighted policy and the CBM policy increases.

The failure of any of the subsystems results in a system failure. Suppose the system is in a failed state and pump A, pulveriser A, and the boiler are in failed states. When S is large, it may be beneficial, in the long-term, to replace all components at this system failure. This will result in a large initial penalty since we will be replacing three components that are not in failed states (pump B, pulveriser B, and the generator), however, replacing all components upon system failure will result in paying the large set-up cost, S, as infrequently as possible in the long term. Replacing working, but partially degraded, components will increase the (expected) time to the next system failure, resulting in less frequent maintenance. This results in paying the expensive set-up cost less frequently.

The short-sighted policy would naively opt to replace only the failed components (pump A, pulveriser A, and the boiler), since this results in the smallest short-term penalty. Continuing to only replace failed components at system failure will always result in the smallest short-term penalty (when S is large) but does not result in the smallest long-term penalty. Forecasting further ahead reveals that a large initial penalty results in a smaller long-term penalty.

This example highlights the importance of a sequential maintenance policy. Policies that minimise short-term penalties or cost per unit time, or maximise short-term profit (Liu et al., 2014), may result in a suboptimal long-term strategy.

8.4 CONCLUSIONS

In this chapter we implemented the sequential maintenance policy proposed in Chapter 7. We considered two different systems, a three-component parallel system with identical components, and a six-component series-parallel power plant system. We compared the CBM policy to alternative policies to highlight benefits of a loss-based approach and including all uncertainty when making maintenance decisions.

The loss-based utility performed as well or better than the CPT utility and we highlighted the benefits of a loss-based approach over the cost-per-unit time utility and suggested areas for future research. We compared the random-threshold approach to the fixed-threshold approach (the threshold policy) and the deterministic policy to highlight that not incorporating all uncertainty can lead to suboptimal decisions and to encourage policies to incorporate parameter uncertainty. Finally, we compared a short-sighted policy to a sequential policy to highlight that short-sighted policies can lead to suboptimal decisions.

Chapter 9

Conclusions and Future Work

This thesis is concerned with reliability and maintenance optimisation in the age of data-centric engineering and has two main contributions. One is in the field of reliability for HDDs with automatic data-collecting devices and one is in the field of CBM for complex multi-component systems with dependencies. The first contribution, described in detail in Part I, provides a novel way to model the survival probabilities and failure ages of hard drives using dynamic data collected by the drives. The methodology allows us to specify the impact of critical attributes on the failure age of a hard drive. The work described in Part I has been accepted for part of a special issue on degradation and maintenance, modelling and analysis in Applied Stochastic Models in Business and Industry (ASMBI).

The second contribution, described in detail in Part II, provides a novel loss-based utility (or reward) function, that is incorporated in a Bayesian sequential decision framework, to decide which components are to be maintained at maintenance opportunities for continuously monitored multi-component systems that are subject to economic and failure-based load sharing dependence (a type of stochastic dependence). The work described in Part II was published in January 2022 as part of a special issue on maintenance planning in Reliability Engineering & System Safety (Oakley et al., 2022).

9.1 Part I

The aim of this work was to study the survival probabilities and failure ages of hard drives, using data collected by SMART, and to examine the impact of critical attributes on hard drive failure ages. The problem of predicting hard drive failure ages is of critical importance and has been extensively studied over the past decades. Predicting drive failures before they occur can inform us to take action in advance. Our work was motivated by the lack of papers in the literature focusing on hard drive failure prediction from a probabilistic or

RUL point of view. Furthermore, we sought to concretely answer the question “How much do critical attributes impact the failure age of a hard drive”?

As a starting point, we focused on modelling the survival probability of hard drives without incorporating the (covariate) attributes collected by SMART. More specifically, following from Mittman et al. (2019), we proposed a two-state model to obtain the survival probabilities of hard drives using left-truncated and right-censored failure age data. A generalised limited failure population model was used to describe the failure age of hard drives. This model captured the early failure mode and the wear-out failure mode of hard drives. In Chapter 3, we described in detail the left-truncated and right-censored data observed and derived the likelihood and the DRLs under the two-state model. We completed the model by specifying prior distributions for model parameters and estimated the model parameters in RStan.

We then set out to extend the two-state model to incorporate the attributes collected by SMART. Following from Ma et al. (2015), Rincón et al. (2017), and Backblaze (2022b), we reduced the number of SMART attributes to consider to five attributes; SMART 5, the reallocated sectors count; SMART 187, the reported uncorrectable errors; SMART 188, command timeout; SMART 197, the current pending sector count; and SMART 198, the uncorrectable sector count. We named these attributes critical attributes. A parametric model for each critical attribute is needed for the purpose of prediction. However, the erratic nature of these poorly understood processes made it difficult to predict their values over time.

Consequently, we proposed an illness-death model to obtain the survival probabilities of hard drives, where we defined a drive to be in the illness state, termed the critical state, if at least one of the critical attributes is nonzero. This is an immediate extension of the two-state model, and it allowed us to incorporate the critical attributes collected by SMART without having to forecast the process for any of the critical attributes. Instead, we needed to forecast the probability of entering the critical state; a more manageable problem. We characterised the illness-death model using state specific transition hazards and proposed a parametric model for each transition hazard.

We then extended the illness-death model to a four-state multi-state model, named the multi-state model. This model incorporated two intermediate states: the critical 1 state and the critical 2 state. We defined a hard drive to be in the critical 1 state if one of the critical attributes is nonzero. We defined a drive to be in the critical 2 state if at least two of the critical attributes are nonzero.

We illustrated how to obtain DRLs under the multi-state model and the two-state model. We found that the RUL for drives in the critical 1 state is lower than the RUL for drives in the healthy state; and the RUL for drives in the critical 2 state is lower than the

RUL for drives in the critical 1 state. The multi-state model suggested that drives with multiple critical attributes are more prone to failure than drives with only one critical attribute and drives with one critical attribute are more prone to failure than drives without any critical attributes. The RUL distribution, for a drive conditional on surviving until age $\tau_i = 5000, 10000, 15000, 20000$, obtained under the two-state model appeared to be a weighted mixture of the three RUL distributions (corresponding to the RUL distributions for drives in the healthy, critical 1 and critical 2 states) obtained under the multi-state model. This indicated that the two-state model may be underestimating the RUL of drives in the healthy state and overestimating the RUL of drives in the critical states.

Furthermore, we illustrated how to obtain survival probabilities under the multi-state model. We showed how the survival posterior predictive distributions can be used to compare the probabilities of failure of drives in the healthy state to drives in the critical states, and drives in the critical 1 state to drives in the critical 2 state, within a forecast horizon of interest. This allows us to concretely define the impact of a single critical attribute and the impact of multiple critical attributes on the survival probabilities of hard drives; which in turn allows us to examine the impact of a single critical attribute and the impact of multiple critical attributes on the RUL of hard drives. This approach allowed us to coherently answer the question "How much do critical attributes impact the failure age of a hard drive?"

We assessed the performance of the two-state model, the illness-death model, and the multi-state model in a simulation study, using the AUC (discrimination) and the PE (calibration) described in Chapter 4. We performed Monte Carlo cross-validation, splitting the data into training (60%) and validation (40%) data. For each split, we fitted the two-state model, the illness-death model, and the multi-state model to the training data. We obtained the AUC and the PE every four weeks (672 hours), i.e., at calendar times $\tau = 0, 672, \dots,$ assuming relevant time intervals of $s = 672, 1344, 2016$ hours for prediction. We found that the multi-state model and the illness-death model outperformed the two-state model. Furthermore, we found that the multi-state model outperformed the illness-death model. The results illustrated the importance of incorporating the attributes collected by SMART, and the multi-state model provided a framework to do this.

The differences in the AUC and the PE between the multi-state model and the illness-death model were not as large as the differences between the multi-state model and the two-state model or between the illness-death model and the two-state model. This suggested that more complex models, for example a multi-state model with five states, may not be superior to the four-state multi-state model presented in Chapter 5. In addition, this model would have more parameters and be more challenging to train (due to fewer drives transitioning between each state). Furthermore, we found that command timeout, SMART 188, did not appear to be a critical attribute.

We stated in Section 5.1.2.1: Backblaze uses five SMART attributes as a means of helping determine if a drive is going to fail (Rincón et al., 2017; Backblaze, 2022b). Namely, SMART 5, the reallocated sectors count; SMART 187, the reported uncorrectable errors; SMART 188, command timeout; SMART 197, the current pending sector count; and SMART 198, the uncorrectable sector count. When the raw value for at least one of these five attributes is greater than zero, Backblaze has a reason to investigate (Rincón et al., 2017; Backblaze, 2022b). In Part I of this thesis we found that command timeout, SMART 188, did not appear to be a critical attribute. In addition, we showed how to use posterior predictive survival distributions and posterior predictive RUL distributions (see Figure 5.5, Table 5.2 and Figure 5.6) to concretely examine the impact of critical attributes on hard drive survival probabilities and failure ages. The novel approach to modelling hard drive survival probabilities can be used to monitor the risk of data loss.

Future work should consider extending the illness-death model and the multi-state model to hierarchical models to model the failure ages of the entire Backblaze population consisting of different but similar subpopulations (drive-brands). The limited amount of data available for many of the subpopulations may prohibit fitting the multi-state models to each drive-brand separately. Modelling subpopulation-specific parameters hierarchially, borrowing strength across subpopulations, could make fitting the multi-state models more feasible. Moreover, in this thesis we use GLFP and Weibull hazard functions for the two-state, illness-death and multi-state models. Future work could consider alternative hazard functions, such as Cox proportional hazards and semi-Markov models (Suresh et al., 2017; Lee et al., 2021). In addition, a joint modelling approach for longitudinal (the SMART attributes) and time-to-event data could be considered and compared to the multi-state models presented in Part I of this thesis.

9.2 Part II

The research in Part II is motivated by gaps in the maintenance literature. There are only a small number of papers in the literature that consider stochastic dependence (Keizer et al., 2017; de Jonge and Scarf, 2019). The majority of studies on multi-component systems consider a single type of dependence, implying that ample research opportunities exist that incorporate multiple dependencies (de Jonge and Scarf, 2019). Moreover, as noted by de Jonge and Scarf (2019), only a limited number of studies take parameter uncertainty into account. We contribute to all of these areas in Part II of this thesis. Furthermore, we highlight the benefits of sequential maintenance decisions over one-step ahead decisions.

It is common in the maintenance literature to make maintenance decisions by minimising the cost per unit time. In Part II of this thesis we proposed a loss-based utility (or reward) function, Λ, for multi-component systems with economic dependence (through a fixed set-up cost) and stochastic dependence (through failure-based load sharing). The utility, Λ, is a combination of interpretable penalties that encapsulate the costs of economic and stochastic dependence.

There are potential benefits to using a loss-based utility function in practice. First, writing the loss of each type of dependence separately shows the cost of each type of dependence. It could be beneficial, in terms of design or maintenance planning, to understand which dependence is the most expensive. For example, if we find that the cost of load-sharing is magnitudes larger than other costs, it may prompt us to add more redundancy to a system or to a particular subsystem; or if the cost of resource dependence is magnitudes larger than other costs, it may allow us to schedule more workers or to buy more spare parts in advance. The above insights may not be as clearly drawn when using the cost-per-unit-time utility. Separating the costs of each form of dependence may be of value to the maintenance literature, especially for large complex systems with multiple dependencies.

Second, in Section 8.2.3.3, we discussed that when using the cost-per-unit time utility, the reward of a sequence of maintenance actions will, in general, not be fully revealed until further into the future (which requires further decisions to be made). In contrast, using the loss-based utility proposed in Chapter 7, the full loss (or full reward) of a maintenance action is obtained at the time an action is performed since it is independent of future actions.

This property of the loss-based utility could identify the optimal sequence of actions for smaller values of ξ compared to the cost-per-unit time policy. This could be considered in future work. Is it beneficial to choose a utility function (or reward function) whose rewards are instantly revealed (at the the time the action is performed) and are independent of future actions? Could this reduce computational time? Could it reduce the value of ξ required to obtain the optimal decision? Moreover, as discussed in Section 8.2.3.3, using a fixed-time interval, as in Vu et al. (2020), for every sequence of maintenance actions would not work in the simulation study considered in Section 8.2.3.3. Comparing different decisions using a fixed-time interval reduces to choosing the plan with the lowest cost. In our simulation study, using a fixed-time interval would result in the policy choosing the action with the smallest number of set-up costs, which would be identical to the simultaneous maintenance policy. Thus, we found that the decision that minimises cost-per-unit time over a fixed interval, will not, in general, be the optimal decision. The drawbacks of using cost-per-unit time could be considered further in future research alongside potential issues with using a loss-based approach to obtain the maintenance decisions; since the goal for many companies will be to minimise the long-term cost - and the loss-based approach does not directly minimise this.

Before proposing the loss-based utility function (to be incorporated into a CBM policy) we illustrated, through examples, the penalties observed for multi-component systems subject to economic dependence, through a fixed set-up cost, and stochastic dependence, through failure-based load sharing. We defined four different types of penalty observed for multi-component systems subject to economic and stochastic dependence. Namely, the set-up cost; a penalty for replacing working components, and hence wasting remaining useful component life; a load sharing penalty, due to components degrading more quickly because of load redistribution; and a reward, at the expense of load sharing, due to postponing maintenance of failed components.

We then derived a one-step ahead maintenance policy which incorporated a loss-based utility function in a Bayesian framework. We then extended the policy to a sequential decision policy that minimises the long-term expected loss of a sequence of maintenance actions. The ideas proposed in Part II could be used to extended loss-based utility functions to systems with resource and structural dependence. In addition, the losses due to each type of dependence can be tailored to application.

In Chapter 8 we found that the loss-based utility performed as well or better than the CPT utility and we highlighted the benefits of a loss-based approach over the cost-per-unit time utility and suggested areas for future research. We compared the random-threshold approach to the fixed-threshold approach (the threshold policy) and the deterministic policy to highlight that not incorporating all uncertainty can lead to suboptimal decisions and to encourage policies to incorporate parameter uncertainty. Finally, we compared a short-sighted policy to a sequential policy to highlight that short-sighted policies can lead to suboptimal decisions.

We consider the case where all component failure times are maintenance opportunities. Further work could involve adapting the policy to consider preventative replacements. Moreover, for some systems the downtime cost will not be negligible, and a system shutdown cost could be incorporated into the analysis.

The proposed policy will encounter computational difficulties for larger systems. For this reason we suggest future work could consider using deep reinforcement learning methods to test the proposed methodology for large systems with economic and stochastic dependence. The negative of the proposed utility function could be used in deep reinforcement learning approaches as the reward function for maintenance actions (Huang et al., 2020; Wei et al., 2020; Zhang and Si, 2020). The properties of the loss-based utility, discussed in Section 8.2.3.3, could make it appealing to deep reinforcement learning algorithms. Moreover, the proposed policy should be tested on some real datasets and compared to the alternative policies introduced in Section 8.1 to test whether the policy can be applied in practice and if it works in practice. Future work could also consider the robustness of the loss-based utility to model misspecification.

Bibliography

J. Arts and R. Basten. Design of multi-component periodic maintenance programs with single-component models. *IISE Transactions*, 50(7):606–615, 2018. doi: 10.1080/24725854.2018.1437301.

Backblaze. Backblaze hard drive data sets, 2022a. URL `https://www.backblaze.com/b2/hard-drive-test-data.html`.

Backblaze. What smart stats tell us about hard drives, 2022b. URL `https://www.backblaze.com/blog/what-smart-stats-indicate-hard-drive-failures`.

M. Ben-Daya, U. Kumar, and D. P. Murthy. *Introduction to maintenance engineering: modelling, optimization and management*. John Wiley & Sons, 2016.

M. Betancourt. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.

M. M. Botezatu, I. Giurgiu, J. Bogojeska, and D. Wiesmann. Predicting disk replacement towards reliable data centers. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 39–48, 2016.

B. Brown, B. Liu, S. McIntyre, and M. Revie. Reliability analysis of load-sharing systems with spatial dependence and proximity effects. *Reliability Engineering & System Safety*, 221:108284, 2022.

B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76(1), 2017.

V. Chan and W. Q. Meeker. A failure-time model for infant-mortality and wearout failure modes. *IEEE Transactions on Reliability*, 48(4):377–387, 1999.

I. C. Chaves, M. R. P. De Paula, L. G. Leite, J. P. P. Gomes, and J. C. Machado. Hard disk drive failure prediction method based on a bayesian network. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2018.

T. R. Chhetri, A. Kurteva, J. G. Adigun, and A. Fensel. Knowledge graph based hard drive failure prediction. *Sensors*, 22(3):985, 2022.

B. de Jonge and P. A. Scarf. A review on maintenance optimization. *European Journal of Operational Research*, 2019. doi: 10.1016/j.ejor.2019.09.047.

C. Diallo, U. Venkatadri, A. Khatab, and Z. Liu. Optimal selective maintenance decisions for large serial k-out-of-n: G systems under imperfect maintenance. *Reliability Engineering & System Safety*, 175:234–245, 2018. doi: 10.1016/j.ress.2018.03.023.

P. Do and C. Bérenguer. Conditional reliability-based importance measures. *Reliability Engineering & System Safety*, 193:106633, 2020. doi: 10.1016/j.ress.2019.106633.

F. D. dos Santos Lima, G. M. R. Amaral, L. G. de Moura Leite, J. P. P. Gomes, and J. de Castro Machado. Predicting failures in hard drives with lstm networks. In *2017 Brazilian Conference on Intelligent Systems (BRACIS)*, pages 222–227. IEEE, 2017.

B. L. Egleston, D. O. Scharfstein, E. E. Freeman, and S. K. West. Causal inference for non-mortality outcomes in the presence of death. *Biostatistics*, 8(3):526–545, 2007.

W. Fauriat and E. Zio. Optimization of an aperiodic sequential inspection and condition-based maintenance policy driven by value of information. *Reliability Engineering & System Safety*, 204:107133, 2020. doi: 10.1016/j.ress.2020.107133.

J. P. Fine, H. Jiang, and R. Chappell. On semi-competing risks data. *Biometrika*, 88(4): 907–919, 2001.

D. P. Gaver and I. G. O'Muircheartaigh. Robust empirical bayes analyses of event rates. *Technometrics*, 29(1):1–15, 1987.

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 1995.

T. A. Gerds and M. Schumacher. Consistent estimation of the expected brier score in general survival models with right-censored event times. *Biometrical Journal*, 48(6): 1029–1040, 2006.

M. S. Hamada, H. F. Martz, C. S. Reese, and A. G. Wilson. *Bayesian reliability*, volume 15. Springer, 2008.

F. E. Harrell Jr, K. L. Lee, and D. B. Mark. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in medicine*, 15(4):361–387, 1996.

R. Henderson, P. Diggle, and A. Dobson. Identification and efficacy of longitudinal markers for survival. *Biostatistics*, 3(1):33–50, 2002.

Y. Hong and W. Q. Meeker. Field-failure predictions based on failure-time data with dynamic covariate information. *Technometrics*, 55(2):135–149, 2013a. doi: 10.1080/00401706.2013.765324. URL `https://doi.org/10.1080/00401706.2013.765324`.

Y. Hong and W. Q. Meeker. Field-failure predictions based on failure-time data with dynamic covariate information. *Technometrics*, 55(2):135–149, 2013b.

Y. Hong, Y. Duan, W. Q. Meeker, D. L. Stanley, and X. Gu. Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics*, 57(2):180–193, 2015.

J.-J. Hsieh, W. Wang, and A. Adam Ding. Regression analysis based on semicompeting risks data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):3–20, 2008.

J. Huang, Q. Chang, and J. Arinez. Deep reinforcement learning based preventive maintenance policy for serial production lines. *Expert Systems with Applications*, 160: 113701, 2020. doi: 10.1016/j.eswa.2020.113701.

C. A. Irawan, D. Ouelhadj, D. Jones, M. Stålhane, and I. B. Sperstad. Optimisation of maintenance routing and scheduling for offshore wind farms. *European Journal of Operational Research*, 256(1):76–89, 2017. doi: 10.1016/j.ejor.2016.05.059.

V. E. Johnson, A. Moosman, and P. Cotter. A hierarchical model for estimating the early reliability of complex systems. *IEEE Transactions on Reliability*, 54(2):224–231, 2005.

M. C. O. Keizer, S. D. P. Flapper, and R. H. Teunter. Condition-based maintenance policies for systems with multiple dependent components: A review. *European Journal of Operational Research*, 261(2):405–420, 2017. doi: 10.1016/j.ejor.2017.02.044.

M. C. O. Keizer, R. H. Teunter, J. Veldman, and M. Z. Babai. Condition-based maintenance for systems with economic dependence and load sharing. *International Journal of Production Economics*, 195:319–327, 2018. doi: 10.1016/j.ijpe.2017.10.030.

R. Laggoune, A. Chateauneuf, and D. Aissani. Impact of few failure data on the opportunistic replacement policy for multi-component systems. *Reliability Engineering & System Safety*, 95(2):108–119, 2010. doi: 10.1016/j.ress.2009.08.007.

L. Lakhal, L.-P. Rivest, and B. Abdous. Estimating survival and association in a semicompeting risks model. *Biometrics*, 64(1):180–188, 2008.

C. Lee, S. J. Lee, and S. Haneuse. Time-to-event analysis when the event is defined on a finite time interval. *Statistical methods in medical research*, 29(6):1573–1591, 2020.

C. Lee, P. Gilsanz, and S. Haneuse. Fitting a shared frailty illness-death model to left-truncated semi-competing risks data to examine the impact of education level on incident dementia. *BMC Medical Research Methodology*, 21(1):1–13, 2021.

K. H. Lee, S. Haneuse, D. Schrag, and F. Dominici. Bayesian semiparametric analysis of semicompeting risks data: investigating hospital readmission after a pancreatic cancer diagnosis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 64(2): 253–273, 2015.

N. Li, N. Gebraeel, Y. Lei, X. Fang, X. Cai, and T. Yan. Remaining useful life prediction based on a multi-sensor data fusion model. *Reliability Engineering & System Safety*, 208:107249, 2021. doi: 10.1016/j.ress.2020.107249.

B. Liu, Z. Xu, M. Xie, and W. Kuo. A value-based preventive maintenance policy for multi-component system with continuously degrading components. *Reliability Engineering & System Safety*, 132:83–89, 2014. doi: 10.1016/j.ress.2014.06.012.

B. Liu, M. Xie, and W. Kuo. Reliability modeling and preventive maintenance of load-sharing systemswith degrading components. *Iie Transactions*, 48(8):699–709, 2016.

B. Liu, R.-H. Yeh, M. Xie, and W. Kuo. Maintenance scheduling for multicomponent systems with hidden failures. *IEEE Transactions on reliability*, 66(4):1280–1292, 2017.

B. Liu, X. Zhao, G. Liu, and Y. Liu. Life cycle cost analysis considering multiple dependent degradation processes and environmental influence. *Reliability Engineering & System Safety*, 197:106784, 2020.

B. Liu, M. D. Pandey, X. Wang, and X. Zhao. A finite-horizon condition-based maintenance policy for a two-unit system with dependent degradation processes. *European Journal of Operational Research*, 295(2):705–717, 2021.

L. Liu, R. A. Wolfe, and X. Huang. Shared frailty models for recurrent events and a terminal event. *Biometrics*, 60(3):747–756, 2004.

S. Lu, B. Luo, T. Patel, Y. Yao, D. Tiwari, and W. Shi. Making disk failure predictions {SMARTer}! In *18th USENIX Conference on File and Storage Technologies (FAST 20)*, pages 151–167, 2020.

D. J. Lunn, A. Thomas, N. Best, and D. Spiegelhalter. Winbugs-a bayesian modelling framework: concepts, structure, and extensibility. *Statistics and computing*, 10(4):325–337, 2000.

A. Ma, R. Traylor, F. Douglis, M. Chamness, G. Lu, D. Sawyer, S. Chandra, and W. Hsu. Raidshield: characterizing, monitoring, and proactively protecting against disk failures. *ACM Transactions on Storage (TOS)*, 11(4):1–28, 2015.

I. Manousakis, S. Sankar, G. McKnight, T. D. Nguyen, and R. Bianchini. Environmental conditions and disk reliability in free-cooled datacenters. In *14th USENIX conference on file and storage technologies (FAST 16)*, pages 53–65, 2016.

R. M. Martinod, O. Bistorin, L. F. Castañeda, and N. Rezg. Maintenance policy optimisation for multi-component systems considering degradation of components and imperfect maintenance actions. *Computers & Industrial Engineering*, 124:100–112, 2018. doi: 10.1016/j.cie.2018.07.019.

J. J. McCall. Maintenance policies for stochastically failing equipment: a survey. *Management science*, 11(5):493–524, 1965. doi: 10.1287/mnsc.11.5.493.

W. Q. Meeker, L. A. Escobar, and F. G. Pascual. *Statistical methods for reliability data*. John Wiley & Sons, 2022.

E. Mittman, C. Lewis-Beck, and W. Q. Meeker. A hierarchical model for heterogenous reliability field data. *Technometrics*, 61(3):354–368, 2019.

R. K. Mobley. *An introduction to predictive maintenance*. Elsevier, 2002.

J. F. Murray, G. F. Hughes, K. Kreutz-Delgado, and D. Schuurmans. Machine learning methods for predicting failures in hard drives: A multiple-instance application. *Journal of Machine Learning Research*, 6(5), 2005.

K.-A. Nguyen, P. Do, and A. Grall. Multi-level predictive maintenance for multi-component systems. *Reliability engineering & system safety*, 144:83–94, 2015. doi: 10.1016/j.ress.2015.07.017.

J. L. Oakley, K. J. Wilson, and P. Philipson. A condition-based maintenance policy for continuously monitored multi-component systems with economic and stochastic dependence. *Reliability Engineering & System Safety*, 222:108321, 2022.

Z. Pang, X. Si, C. Hu, D. Du, and H. Pei. A bayesian inference for remaining useful life estimation by fusing accelerated degradation data and condition monitoring data. *Reliability Engineering & System Safety*, 208:107341, 2021. doi: 10.1016/j.ress.2020.107341.

G. Papageorgiou, M. M. Mokhles, J. J. Takkenberg, and D. Rizopoulos. Individualized dynamic prediction of survival with the presence of intermediate events. *Statistics in medicine*, 38(30):5623–5640, 2019.

L. Peng and J. P. Fine. Regression modeling of semicompeting risks data. *Biometrics*, 63(1):96–108, 2007.

E. Pinheiro, W.-D. Weber, and L. A. Barroso. Failure trends in a large disk drive population. 2007.

M. Plummer et al. Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*, volume 124, pages 1–10. Vienna, Austria., 2003.

C. Proust-Lima and J. M. Taylor. Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment psa: a joint modeling approach. *Biostatistics*, 10(3):535–549, 2009.

H. Putter, M. Fiocco, and R. B. Geskus. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in medicine*, 26(11):2389–2430, 2007.

N. Rasmekomen and A. K. Parlikad. Maintenance optimization for asset systems with dependent performance degradation. *IEEE Transactions on Reliability*, 62(2):362–367, 2013. doi: 10.1109/TR.2013.2257056.

C. A. Rincón, J.-F. Pâris, R. Vilalta, A. M. Cheng, and D. D. Long. Disk failure prediction in heterogeneous environments. In *2017 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, pages 1–7. IEEE, 2017.

D. Rizopoulos, G. Molenberghs, and E. M. Lesaffre. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical Journal*, 59(6):1261–1276, 2017.

M. Schemper and R. Henderson. Predictive accuracy and explained variation in cox regression. *Biometrics*, 56(1):249–255, 2000.

J. Shen, J. Wan, S.-J. Lim, and L. Yu. Random-forest-based failure prediction for hard disk drives. *International Journal of Distributed Sensor Networks*, 14(11):1550147718806480, 2018.

Slurm. Slurm: A highly scalable workload manager, 2022. URL `https://github.com/SchedMD/slurm`.

B. Snyder and M. J. Kaiser. Ecological and economic cost-benefit analysis of offshore wind energy. *Renewable Energy*, 34(6):1567–1578, 2009. doi: 10.1016/j.renene.2008.11.015.

Stan. Stan documentation, 2022. URL `https://mc-stan.org/users/documentation`.

K. Suresh, J. M. Taylor, D. E. Spratt, S. Daignault, and A. Tsodikov. Comparison of joint modeling and landmarking for dynamic prediction under an illness-death model. *Biometrical Journal*, 59(6):1277–1300, 2017.

E. J. Tchetgen Tchetgen. Identification and estimation of survivor average causal effects. *Statistics in medicine*, 33(21):3601–3628, 2014.

L. Thomas. A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. *Reliability Engineering*, 16(4):297–309, 1986. doi: 10.1016/0143-8174(86)90099-5.

B. Vakulenko-Lagun and M. Mandel. Comparing estimation approaches for the illness–death model under left truncation and right censoring. *Statistics in Medicine*, 35(9): 1533–1548, 2016.

A. Van Horenbeek, J. Buré, D. Cattrysse, L. Pintelon, and P. Vansteenwegen. Joint maintenance and inventory optimization systems: A review. *International Journal of Production Economics*, 143(2):499–508, 2013. doi: 10.1016/j.ijpe.2012.04.001.

R. Varadhan, Q.-L. Xue, and K. Bandeen-Roche. Semicompeting risks in aging research: methods, issues and needs. *Lifetime data analysis*, 20(4):538–562, 2014.

K. V. Vishwanath and N. Nagappan. Characterizing cloud computing hardware reliability. In *Proceedings of the 1st ACM symposium on Cloud computing*, pages 193–204, 2010.

H. C. Vu, P. Do Van, A. Barros, and C. Bérenguer. Maintenance activities planning and grouping for complex structure systems. In *ESREL 2012 - PSAM 2012 - the Annual European Safety and Reliability Conference and 11th International Probabilistic Safety Assessment and Management Conference*, page CDROM, Helsinki, Finland, June 2012. URL `https://hal.archives-ouvertes.fr/hal-00695783`.

H. C. Vu, P. Do, M. Fouladirad, and A. Grall. Dynamic opportunistic maintenance planning for multi-component redundant systems with various types of opportunities. *Reliability Engineering & System Safety*, 198:106854, 2020. doi: 10.1016/j.ress.2020.106854.

G. Waeyenbergh and L. Pintelon. A framework for maintenance concept development. *International journal of production economics*, 77(3):299–313, 2002. doi: 10.1016/S0925-5273(01)00156-6.

H. Wang. A survey of maintenance policies of deteriorating systems. *European journal of operational research*, 139(3):469–489, 2002. doi: 10.1016/S0377-2217(01)00197-7.

S. Wei, Y. Bao, and H. Li. Optimal policy for structure maintenance: A deep reinforcement learning framework. *Structural Safety*, 83:101906, 2020. doi: 10.1016/j.strusafe.2019.101906.

S. Wu, Y. Chen, Q. Wu, and Z. Wang. Linking component importance to optimisation of preventive maintenance policy. *Reliability Engineering & System Safety*, 146:26–32, 2016. doi: 10.1016/j.ress.2015.10.008.

J. Xu, J. D. Kalbfleisch, and B. Tai. Statistical analysis of illness–death processes and semicompeting risks data. *Biometrics*, 66(3):716–725, 2010.

M.-Y. You, F. Liu, W. Wang, and G. Meng. Statistically planned and individually improved predictive maintenance management for continuously monitored degrading systems. *IEEE Transactions on Reliability*, 59(4):744–753, 2010. doi: 10.1109/TR.2010.2085572.

H. Yu, C. Chu, Ė. Châtelet, and F. Yalaoui. Reliability optimization of a redundant system with failure dependencies. *Reliability Engineering & System Safety*, 92(12):1627–1634, 2007. doi: 10.1016/j.ress.2006.09.015.

N. Zhang and W. Si. Deep reinforcement learning for condition-based maintenance planning of multi-component systems under dependent competing risks. *Reliability Engineering & System Safety*, 203:107094, 2020. doi: 10.1016/j.ress.2020.107094.

N. Zhang, M. Fouladirad, and A. Barros. Optimal imperfect maintenance cost analysis of a two-component system with failure interactions. *Reliability Engineering & System Safety*, 177:24–34, 2018. doi: 0.1016/j.ress.2018.04.019.

S.-J. Zhang, R. Kang, and Y.-H. Lin. Remaining useful life prediction for degradation with recovery phenomenon based on uncertain process. *Reliability Engineering & System Safety*, 208:107440, 2021. doi: 10.1016/j.ress.2021.107440.

R. Zheng, B. Chen, and L. Gu. Condition-based maintenance with dynamic thresholds for a system using the proportional hazards model. *Reliability Engineering & System Safety*, 204:107123, 2020. doi: 10.1016/j.ress.2020.107123.

X. Zhou, L. Xi, and J. Lee. Reliability-centered predictive maintenance scheduling for a continuously monitored system subject to degradation. *Reliability Engineering & System Safety*, 92(4):530–534, 2007. doi: 10.1016/j.ress.2006.01.006.

E. Zio and M. Compare. Evaluating maintenance policies by quantitative modeling and analysis. *Reliability engineering & system safety*, 109:53–65, 2013. doi: 10.1016/j.ress.2012.08.002.

Appendix A

Appendices

A.1 Supplementary material for Part I

A.1.1 Supplementary tables and figures

In this section we provide summaries of the parameter posterior distributions for the two-state model, the illness-death model and the multi-state model, respectively, for one Monte Carlo cross-validation. In addition, we present the simulation study results comparing the illness-death model to the two-state model and the results comparing the multi-state model to the illness-death model.

Parameter	$\alpha_{0.025}$	$\alpha_{0.5}$	$\alpha_{0.975}$
t_{p_1}	0.5774	1.760	3.452
σ_1	0.5337	0.9616	2.691
t_{p_2}	17.90	18.25	18.57
σ_2	0.2015	0.2141	0.2286
π	0.02760	0.05731	0.1137

Table A.1: Posterior medians and 95% credible intervals for the five two-state model parameters. The quantiles of t_{p_1} and t_{p_2} are in thousands of hours.

Parameter	$\alpha_{0.025}$	$\alpha_{0.5}$	$\alpha_{0.975}$
$t_{p_{01}}$	19.53	19.95	20.36
σ_{01}	0.3777	0.3960	0.4163
$t_{p_{1,02}}$	0.3584	1.563	5.951
$\sigma_{1,02}$	0.5853	1.451	6.251
$t_{p_{2,02}}$	24.61	25.75	27.11
$\sigma_{2,02}$	0.1985	0.2430	0.3034
π	0.01909	0.04935	0.1522
$t_{p_{12}}$	8.598	9.594	10.51
σ_{12}	0.2697	0.2997	0.3344

Table A.2: Posterior medians and 95% credible intervals for the nine illness-death model parameters. The quantiles of $t_{p_{01}}$, $t_{p_{1,02}}$, $t_{p_{2,02}}$ and $t_{p_{12}}$ are in thousands of hours.

Parameter	$\alpha_{0.025}$	$\alpha_{0.5}$	$\alpha_{0.975}$
$t_{p_{01}}$	21.07	21.59	22.15
σ_{01}	0.4050	0.4260	0.4510
$t_{p_{02}}$	14.37	15.24	15.99
σ_{02}	0.2765	0.3087	0.3541
$t_{p_{1,03}}$	0.3719	1.553	6.077
$\sigma_{1,03}$	0.5516	1.450	5.304
$t_{p_{2,03}}$	24.61	25.69	27.05
$\sigma_{2,03}$	0.1962	0.2408	0.3011
π	0.01856	0.04947	0.1285
$t_{p_{12}}$	6.216	7.298	8.337
σ_{12}	0.4196	0.4688	0.5308
$t_{p_{13}}$	8.341	10.70	12.56
σ_{13}	0.2927	0.3673	0.4749
$t_{p_{23}}$	6.282	8.408	10.18
σ_{23}	0.3604	0.4322	0.5330

Table A.3: Posterior medians and 95% credible intervals for the fifteen multi-state model parameters. The quantiles $t_{p_{01}}$, $t_{p_{02}}$, $t_{p_{1,03}}$, $t_{p_{2,03}}$, $t_{p_{12}}$, $t_{p_{13}}$ and $t_{p_{23}}$ are in thousands of hours.

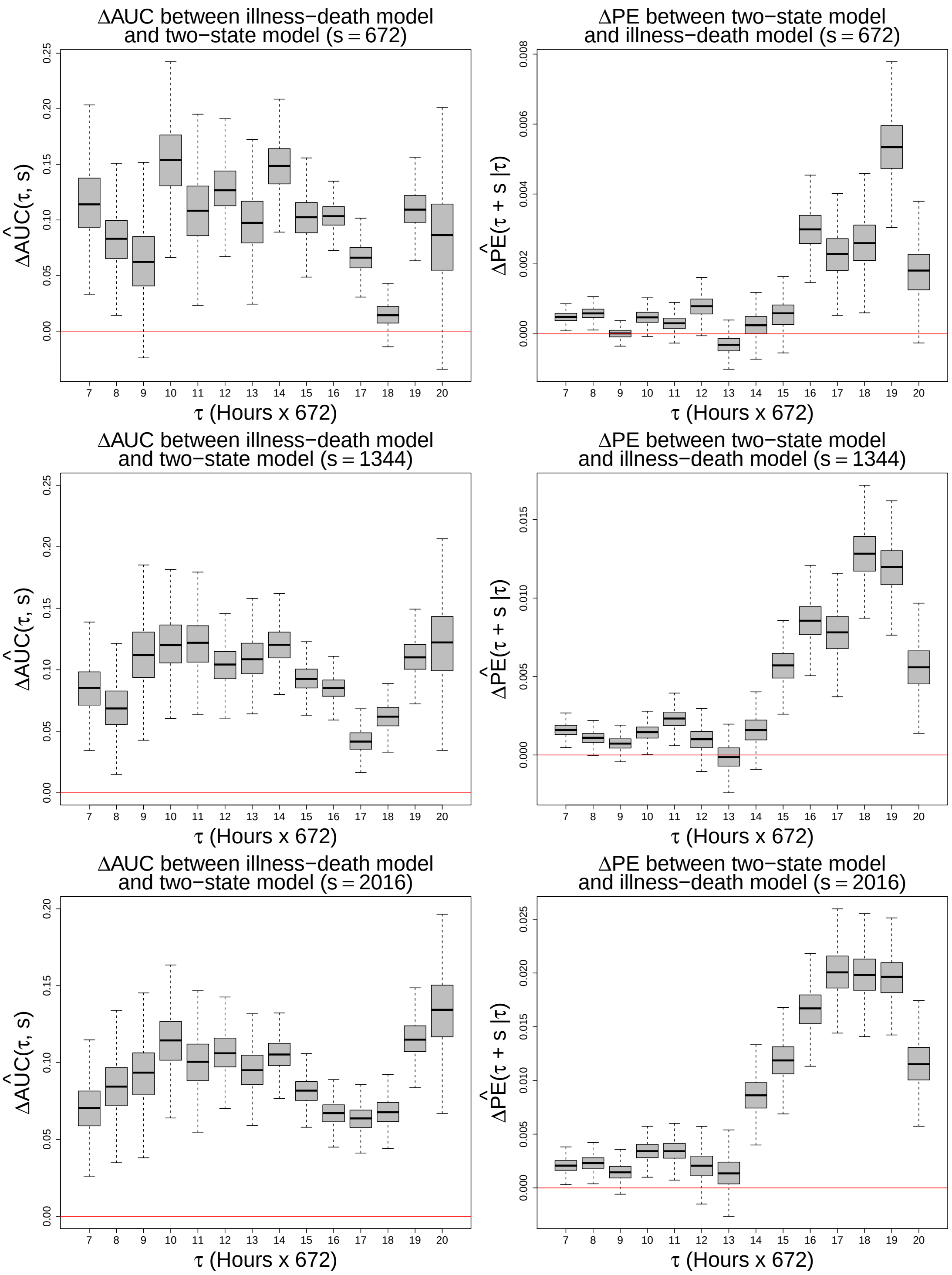

Figure A.1: Top, middle and bottom left panels show the difference in the AUC, at multiple time points, between the illness-death model and the two-state model, for $s = 672, 1344$ and 2016, respectively. Top, middle and bottom right panels show the difference in the PE, at multiple time points, between the two-state model and the illness-death model, for $s = 672, 1344$ and 2016, respectively. The scales of the y-axes vary between plots.

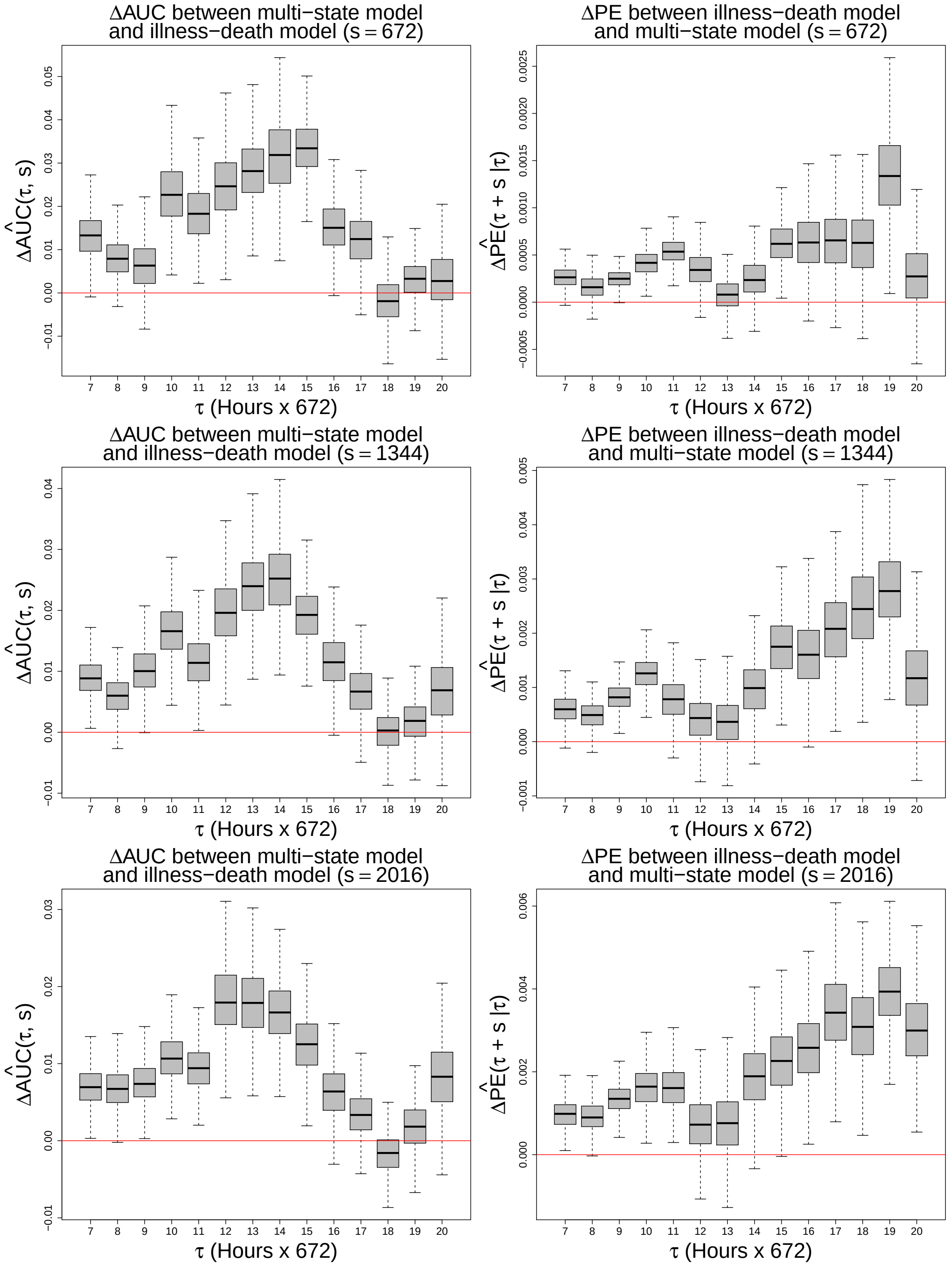

Figure A.2: Top, middle and bottom left panels show the difference in the AUC, at multiple time points, between the multi-state model and the illness-death model, for $s = 672, 1344$ and 2016, respectively. Top, middle and bottom right panels show the difference in the PE, at multiple time points, between the illness-death model and the multi-state model, for $s = 672, 1344$ and 2016, respectively. The scales of the y-axes vary between plots.

A.2 Supplementary material for Part II

A.2.1 Simulation Algorithm to Evaluate CBM Policy

In this section we provide the simulation algorithms used to obtain the action to be performed at a maintenance opportunity for the parallel system with three identical components considered in Section 8.2.

Algorithm A.1 Simulation algorithm to obtain the action to be performed at a maintenance opportunity.

Input: c, S, t_k (current time), t_1', t_2', t_3' (most recent failure times), t_1^*, t_2^*, t_3^* (most recent replacement times), $\boldsymbol{x}_i(t_i^*, t_k) = \{x_i(s) : t_i^* < s \leq t_k\}$, $\boldsymbol{\pi}_i(t_i^*, t_k) = \{\pi_i(s) : t_i^* < s \leq t_k\}$ for $i = 1, 2, 3$, $\boldsymbol{y}(t_k) = (y_1(t_k), y_2(t_k), y_3(t_k))$, $\xi = 6$, $\boldsymbol{\theta}^b$, for $b = 1, \ldots, B$, where $\boldsymbol{\theta}^b$ are parameter samples from the most recently updated posterior distribution, where $\boldsymbol{\theta}^b = (\boldsymbol{\theta}_D^b, \boldsymbol{\theta}_L^b, \boldsymbol{\theta}_X^b)$, where $\boldsymbol{\theta}_D^b = \rho^b$, $\boldsymbol{\theta}_L^b = (\mu_L^b, \sigma_L^b, \alpha_L^b, \beta_L^b)$, and $\boldsymbol{\theta}_X^b = (\mu_X^b, \sigma_X^b, \alpha_X^b, \beta_X^b)$, and $I \in \mathbb{N}_1$, where I is chosen large enough to provide sufficient precision.

for $b \in \{1, \ldots, B\}$ **do**
    Set $\boldsymbol{\theta} = (\rho, \mu_L, \sigma_L, \alpha_L, \beta_L, \mu_X, \sigma_X, \alpha_X, \beta_X) = \boldsymbol{\theta}^b$.
    Obtain $\boldsymbol{D}(t_k)$ using Equations (8.2.2) and (8.2.3).
    Simulate $\boldsymbol{X}_{\text{sys}}(t_k, \infty; \boldsymbol{\theta}_X) = \{X_{\text{sys}}(s; \boldsymbol{\theta}_X) : t_k < s < \infty\}$ using Equation (8.2.1).
    Obtain $\bar{c}$ using Algorithm A.2 with $t = t_k$, and $D(t) = 0$.
    Obtain $\boldsymbol{z}$ using Algorithm A.3.
    Sample ξ random uniform observations from $\text{U}(0, 1)$ and store them in a vector, $\boldsymbol{p}$.
    for $u_{t_k} \in U_{t_k}$ **do**
        if $|u_{t_k}| > \xi$ **then**
            Exit iteration.
        Obtain $\boldsymbol{y}^+(t_k)$, $\boldsymbol{D}^+(t_k)$, and $\pi^+(t_k)$ using Equations (7.2.6), (7.2.7), and (7.2.8), respectively.
        Obtain τ_i^-, for $i = 1, 2, 3$, using Algorithm A.4 with $t = t_k$, $\boldsymbol{Y}(t) = \boldsymbol{y}(t_k)$, $\Pi(t) = 3$, $\boldsymbol{D}(t) = \boldsymbol{D}(t_k)$, and $T = t_k$.
        Update $\boldsymbol{z}$ and $\boldsymbol{p}$ using Algorithm A.5 with $u = u_{t_k}$.
        Obtain τ_i^+, for $i = 1, 2, 3$, using Algorithm A.4 with $t = t_k$, $\boldsymbol{Y}(t) = \boldsymbol{y}^+(t_k)$, $\Pi(t) = 3$, $\boldsymbol{D}(t) = \boldsymbol{D}^+(t_k)$, and $T = t_k$.
        Obtain T_{k+1}, $\boldsymbol{Y}(T_{k+1})$, and $\boldsymbol{D}(T_{k+1})$, using Algorithm A.4 with $t = t_k$, $\boldsymbol{Y}(t) = \boldsymbol{y}^+(t_k)$, $\Pi(t) = \pi^+(t_k)$, $\boldsymbol{D}(t) = \boldsymbol{D}^+(t_k)$, and $T = \infty$.
        Obtain $\Pi(T_{k+1})$ using $\boldsymbol{Y}(T_{k+1})$.
        Obtain $\tau_i^{+,L}$, for $i = 1, 2, 3$, using Algorithm A.4 with $t = t_k$, $\boldsymbol{Y}(t) = \boldsymbol{y}^+(t_k)$, $\Pi(t) = \pi^+(t_k)$, $\boldsymbol{D}(t) = \boldsymbol{D}^+(t_k)$, and $T = T_{k+1}$.
        Obtain $\Lambda(u_{t_k} \mid T_{k+1}, \boldsymbol{X}_{\text{sys}}(t_k, \infty), \boldsymbol{\theta})$ using τ_i^-, τ_i^+, T_{k+1}, and $\tau_i^{+,L}$ and Equation (7.2.9).
        if $|u_{t_k}| = \xi$ **then**
            Store $\Lambda(u_{t_k})$ in memory, $\Gamma_{u_{t_k}}$, and exit iteration.

for $u_{T_{k+1}} \in U_{T_{k+1}}$ **do**
  if $|u_{t_k}| + |u_{T_{k+1}}| > \xi$ **then**
    Exit iteration.
  $\vdots$
  for $u_{T_{k+\mathcal{K}-1}} \in U_{T_{k+\mathcal{K}-1}}$ **do**
    if $|u_{t_k}| + |u_{T_{k+1}}| + \cdots + |u_{T_{k+\mathcal{K}-1}}| > \xi$ **then**
      Exit iteration.
    $\vdots$
  $\vdots$
Obtain $\Lambda^b(u_{t_k}) = \min\{\Gamma_{u_{t_k}}\}$
Obtain

$$\bar{\Lambda}(u_{t_k}) = \frac{1}{B}\sum_{b=1}^{B} \Lambda^b(u_{t_k}). \tag{A.2.1}$$

The action to be performed at t_k is

$$u_{t_k} = \underset{u_{t_k} \in U_{t_k}}{\operatorname{argmin}}\{\bar{\Lambda}(u_{t_k})\}. \tag{A.2.2}$$

return u_{t_k}.

Algorithm A.2 Simulation algorithm to obtain the expected component cost per unit time, $\bar{c}$.

Input: $D(t)$, $\boldsymbol{X}_{\text{sys}}(t, \infty; \boldsymbol{\theta}_X)$, ρ, $\boldsymbol{\theta}_L = (\mu_L, \sigma_L, \alpha_L, \beta_L)$, I, and c.
Obtain $\boldsymbol{D}(t, \infty) = \{D(s) : t < s < \infty\}$, where

$$D(s) = \frac{1}{3^\rho}\int_t^s X_{\text{sys}}(v; \boldsymbol{\theta}_X) dv. \tag{A.2.3}$$

Obtain $F_L(D(s); \mu_L, \sigma_L, \alpha_L, \beta_L)$, for $t \leq s < \infty$, where $F_L(\cdot)$ is the cdf of the truncated normal distribution.
for $l \in \{1, \ldots, I\}$ **do**
  Sample w_l from $W \sim \text{U}(0, 1)$.
  Obtain $D(\tau_l) = F_L^{-1}(w_l)$.
  Obtain $\tau_l = D^{-1}(D(\tau_l))$.
Obtain

$$\bar{\tau} = \frac{1}{I}\sum_{l=1}^{I} \tau_l, \tag{A.2.4}$$

and

$$\bar{c} = \frac{c}{\bar{\tau}}. \tag{A.2.5}$$

return $\bar{c}$.

Algorithm A.3 Simulation algorithm to obtain $\boldsymbol{z}$, the probabilities used to generate component failure times.

Input: $\boldsymbol{D}(t_k) = (D_1(t_k), D_2(t_k), D_3(t_k))$, and $\boldsymbol{\theta}_L = (\mu_L, \sigma_L, \alpha_L, \beta_L)$.

for $i \in \{1, 2, 3\}$ **do**
  Obtain $F_L(D_i(t_k); \mu_L, \sigma_L, \alpha_L, \beta_L)$, where $F_L(\cdot)$ is the cdf of the truncated normal distribution.
  Sample z_i from $Z \sim \mathrm{U}(F_L(D_i(t_k)), 1)$.
return $\boldsymbol{z} = (z_1, z_2, z_3)$.

Algorithm A.4 Simulation algorithm to obtain the subsequent maintenance opportunities, $\mathcal{T}$, the failure states at subsequent maintenance opportunities, $\boldsymbol{Y}(\mathcal{T})$, the degradation levels at subsequent maintenance opportunities, $\boldsymbol{D}(\mathcal{T})$, and τ_i^-, τ_i^+, and $\tau_i^{+,L}$, for $i = 1, 2, 3$.

Input: t, t'_1, t'_2, t'_3, $\boldsymbol{Y}(t)$, $\Pi(t)$, $\boldsymbol{D}(t)$, $\boldsymbol{X}_{\mathrm{sys}}(t, \infty; \boldsymbol{\theta}_X)$, ρ, $\boldsymbol{\theta}_L = (\mu_L, \sigma_L, \alpha_L, \beta_L)$, T, and $\boldsymbol{z} = (z_1, z_2, z_3)$.

for $i \in \{1, 2, 3\}$ **do**
  if $Y_i(t) = 0$ **then**
    Set $\tau_i = t'_i \leq t$.
    Set $D_i(\mathcal{T}) = D_i(t)$ for $\mathcal{T} > t$.
  else
    Obtain $\boldsymbol{D}_i(t, \infty) = \{D_i(s) : t < s < \infty\}$, where

$$D_i(s) = D_i(t) + \int_t^s \frac{X_{\mathrm{sys}}(v; \boldsymbol{\theta}_X)}{\mathcal{N}(v)^\rho} dv, \tag{A.2.6}$$

    where

$$\mathcal{N}(v) = \begin{cases} \Pi_i(t), & \text{for } t < v \leq T, \\ 3, & \text{for } v > T. \end{cases} \tag{A.2.7}$$

    Obtain $F_L(D_i(s); \mu_L, \sigma_L, \alpha_L, \beta_L)$, for $t \leq s < \infty$, where $F_L(\cdot)$ is the cdf of the truncated normal distribution.
    Obtain $D_i(\tau_i) = F_L^{-1}(z_i)$.
    Obtain $\tau_i = D_i^{-1}(D_i(\tau_i))$.
Obtain

$$\mathcal{T} = \min\{\tau_i : \tau_i > t, i = 1, 2, 3\}, \tag{A.2.8}$$

and

$$\mathcal{I} = \min_i\{\tau_i : \tau_i > t, i = 1, 2, 3\}. \tag{A.2.9}$$

Set $\boldsymbol{Y}(\mathcal{T}) = \boldsymbol{Y}(t)$.
Set $Y_{\mathcal{I}}(\mathcal{T}) = 0$.
return τ_i, for $i = 1, 2, 3$, $\mathcal{T}$, $\boldsymbol{Y}(\mathcal{T}) = (Y_1(\mathcal{T}), Y_2(\mathcal{T}), Y_3(\mathcal{T}))$, and $\boldsymbol{D}(\mathcal{T}) = (D_1(\mathcal{T}), D_2(\mathcal{T}), D_3(\mathcal{T}))$.

Algorithm A.5 Simulation algorithm to update $\boldsymbol{z}$ and $\boldsymbol{p}$, the probabilities used to generate component failure times.

Input: u, $\boldsymbol{z}$, and $\boldsymbol{p}$.

```
for i ∈ {1, 2, 3} do
    if Component_i ∈ u then
        Set z_i equal to the first element of p.
        Remove the first element of p.
return z, and p.
```