Reliability Analysis for Systems with Stochastic and Spatial dependence

Bodunrin Nathaniel Brown

A thesis presented in fulfillment of the requirements for the degree of a:

Doctor of Philosophy (PhD)

Department of Management Science
University of Strathclyde, Glasgow

September 28, 2022

Date:.................. Signed:..................

Acknowledgements

I would like to thank the following people for helping me during this process. First of all, I would like to acknowledge and thank my supervisors Professor Matthew Revie, Dr Stuart McIntyre and Dr Bin Liu whose insight and knowledge into the subject matter steered me through this research. Thank you for all your time and support that helped me complete this journey. There is no doubt that you made the process thought provoking and inspiring. Secondly, I would like to say a special thank you to my wife Aisha, thanks for your love and support without which I would have stopped this research a long time ago. You have been amazing. Finally, I want to express my gratitude to my family for all the support you have shown me through this research.

Abstract

This thesis develops models for system reliability prediction considering unobserved heterogeneity and spatial dependence. Three problems are considered in this thesis.

The first problem concerns models to predict system and component level reliability for systems subject to minimal repair and unobserved heterogeneity. Existing models assume that unobserved heterogeneity is constant and methods for component-level prediction had not been considered. However, there are situations where unobserved heterogeneity changes and becomes homogeneous over time. This thesis develops a new reliability model that accounts for the case where unobserved heterogeneity changes and becomes homogeneous over time. We develop a new frailty model using Inverse Gaussian (IG) distribution and develop a method using Empirical Bayes that enable component-level reliability prediction.

The second problem concerns models for reliability assessment for load sharing systems with spatially dependent components and proximity effects. Existing models assume equal-load sharing for systems subject to load-sharing. However, there are systems that operates in a way that the load of a failed component is transferred to the working proximate components. Existing models do not account for the proximity effect. This thesis develops a new reliability model that accounts for load sharing and proximity effect between components. We introduce a function that captures the effect of each load change on the failure rate of a working proximate component. Numerical examples are presented to illustrate the developed model.

The third problem concerns models for reliability assessment and preventive maintenance for load sharing systems with spatially dependent components, proximity effects and shocks. Existing models assume equal-load sharing for systems subject to load-sharing and external shocks. However, models that account for external shocks and proximity effect in the system has not been considered. This thesis develops a function that captures the effect of each load change and shock on the failure rate of a working component. We introduce a Modified failure sequence diagram and an algorithm for system reliability assessment based on Monte Carlo method. In addition, we develop an extension of the age-replacement model for preventive maintenance of the load sharing system. Numerical examples are presented to illustrate the developed models.

Contents

List of Figures

List of Tables

List of Tables

Chapter 1

Introduction

1.1 Introduction to reliability engineering

Reliability was first coined by Samuel T, Coleridge in 1816 in praise of his friend [1, 2]. Oxford dictionary defines reliability as the quality of being trustworthy or of performing consistently well [3]. Broadly, reliability can mean different things to different people. For example, reliability could be interpreted to mean: trustworthiness, accuracy, quality, safety etc. Even across academic disciplines, the use of this concept varies. In psychology, reliability refers to the consistency of a research study, while in statistics, reliability refers to the consistency of a measuring instrument or measures used to describe a test [4]. In the field of reliability engineering, reliability is defined as the probability that a component or system will perform a required function without failure for a given period of time when used under stated operating conditions [5]. According to Bazovsky [6] a clear (unambiguous) definition of a concept is essential for measuring reliability. As a result, the definition used in the field of reliability engineering will be adopted throughout this thesis.

Driven by the need for reliable systems, the birth of reliability engineering took place during the 1950's [1, 2, 7]. Interest in the quality of a system predates the creation of the discipline [1]. However the need for producing reliable systems was first sensed in both commercial and military sectors in early 1950s [8]. The key difference between the quality or reliability of a system rest solely on time. According to Condra

[9] “reliability is quality over time”. Whilst a product’s quality is associated with the workmanship and manufacturing of the product [10], reliability assessment of the product requires testing over a period of time as it may be many months before poor reliability is identified in the system [11].

Reliability is tailored to convey assurance in the successful operation of a system [2], and is believed to be a key component for assessing the performance of a system as “failure of products to function can be costly and lead to unsafe operating conditions” [12]. The desired target for reliability of different systems depends on the consequence of unreliability [8]. Unreliability has a number of unfortunate consequences. In some systems, the ”value placed on reliability is a function of the frustration, annoyance and financial loss incurred when a failure occurs” [8]. In other systems such as medical devices, ”unreliability could result in serious consequences; for example, disability or even death of patients” [13].

The structure of the chapter will take the following form. Section 1.2 introduces the origins and early developments of reliability modelling. Section 1.3 introduces methods for system reliability assessment. Section 1.4 discusses multi-component systems with dependent components. Section 1.5 presents motivation for the thesis with focus on multi-component systems with stochastic and spatial dependence. Section 1.6 outlines the research objectives and section 1.7 outlines the overall structure of the remaining thesis.

1.2 Origins of Reliability Modelling

The vacuum tube is recognized by many as the catalyst for the actual emergence of reliability engineering [1, 2, 14]. At the onset of World War II the vacuum tube initiated the electronic revolution, enabling a series of applications such as the radio, television, radar and others [2]. At the same time, it was also the main cause of equipment failure during World War II [2]. Keeping military electronic equipment working proved to be troublesome and expensive [14]. The vacuum tube was observed to be notably unreliable as replacements were required five times more than all other other electrical equipment [14]. After the war, this experience with the vacuum tubes prompted the US Department of Defense (DoD) to initiate a number of studies to look into these failures [1]. The efforts eventually consolidated in the Advisory Group on Reliability of Electronic Equipment(AGREE) report which gave birth to a new discipline, Reliability Engineering [1]. The AGREE committee was jointly established in 1952 by the DoD and the American Electronics Industry, with the mission of [14]: recommending measures that would result in more reliable equipment; helping to implement reliability programs in government and civilian agencies; disseminating a better education on reliability [2, 14].

The key foundation for the mathematical developments of reliability theory is the theory of probability and statistics (the theory of sampling) [1]. Probability theory was initiated to satisfy the enthusiastic urge for answers to gaming and gambling questions by Blaise Pascal and Pierre de Fermat in the 1600s and later expanded into numerous other practical problems by Laplace in the 1800s [2]. Saleh and Marias [1] state that "probability and statistics are the essential ingredients without which Reliability Engineering as a technical discipline could not have emerged."

Before the World War II, analyses of a system's reliability was intuitive in nature [7] and based on experiences in design and utilization [15]. It was common to make statements such as "the system won't be damaged", "the system is highly reliable", or "it is more reliable than similar products" but these do not represent any measurable and specific meaning at all [15]. Asgari and Lofti [15] stated that "this method was always questionable, improper, and unreliable in terms of engineering judgment". In contrast, after World War II, failure data collection and root cause analyses were launched with the aim of achieving higher reliability in components and devices [2]. Reliability analysis began to be carried out quantitatively and modern reliability modelling began to evolve [8].

Between 1950's and 1960's, there were a couple of notable publications that focused on developing quantitative methods to aid reliability engineers. In 1956 Reliability Analysis Centre (RAC), a major manufacturer of vacuum tubes, published a report on reliability prediction techniques entitled 'Reliability Stress Analysis for Electronic

Equipment'(TR-1100) [2, 7]. The report presented a number of analytical models for estimating component failure rates and was the predecessor of the 1961 military standard MH-217 that is still used today to make reliability predictions [2]. Given specification of quantitative reliability requirements, the published models became building blocks to estimate and predict the reliability of a component before it was built and tested [2].

Following the development of quantitative models, the discipline of reliability engineering proceeded along two paths with some engineers and mathematicians focusing on system reliability improvement (reliability growth modelling) whilst the others focused on maintaining system reliability or availability (asset management). According to Bhamare [7] before 1950s, the focus was either on quality control or on machine maintenance problems. Even now reliability engineering is more established and it aims to address the questions of: why systems fail; how to develop reliable systems; how to measure and test reliability in design, operation and management; how to maintain systems reliability, by maintenance, fault diagnosis and prognosis [2].

1.3 System reliability prediction

Many engineered assets in the industry are repairable systems [16]. The performance of these assets can affect the level of product's quality, production costs, quality of service to customers, and the profit of the company directly [15]. Methods for accurate reliability prediction as well as proper implementation of maintenance policies are, therefore, critical for these assets [17]. Asset management involves the efficient use of assets, maintaining the availability and quality of assets at acceptable performance levels with the lowest cost possible [15]. An engineered asset could be a single-component system (e.g., valve) or multiple component systems (e.g., an automobile). Systems with multiple components are referred to as multi-component systems. Formally, multi-component systems are systems consisting of several subsystems or components organised in a structure [18].

Asset management is concerned with reliability prediction and optimal maintenance policy for assets [15]. Reliability prediction lays a critical foundation for the optimal maintenance policy of the assets [16]. Hence, it is essential to make an accurate reliability prediction for an asset [16]. Assessing the future reliability of a system, in the form of reliability prediction, is important for several reasons. First, reliability prediction allows the forecasting of support costs, spares requirements, warranty costs and marketability [8]. Second, a company can optimise its maintenance strategy based on the prediction of remaining useful life [16]. Third, reliability prediction can be a valuable part of the study and design processes, for comparing options and for highlighting critical reliability features of an asset's design [19]. Lastly, reliability prediction can support decision-making when deciding between different engineering designs [8].

Studies on asset management were first motivated by the military. During the Korean war, maintenance costs were found to be quite significant for some military systems, thus calling for methods of reliability prediction and optimization strategies for asset maintenance and renovation [2]. Many data sets include information about failures for a time interval which is only a small fraction of the actual lifetime (left truncation) [20]. Due to the data quantity problem and poor knowledge of factors that drive component failures, many earlier modelling approaches for reliability prediction consider the grouping of data sets into classes of similar or 'homogeneous' components that, for instance, are made of the same material or have the same diameter and assume component independence [20]. However, heterogeneity or inter-component dependencies could exist in a system which, if overlooked, may lead to inaccurate reliability predictions [21, 22, 23, 24]. Heterogeneity may be due to differences in material, differences in design, differences in location and so on [21, 22, 23, 24] whereas inter-component dependencies could be due to factors that range from the design of the system, maintenance actions conducted on the system's components, and shared environmental conditions to note a few [25]. This thesis is focused on reliability prediction for multi-component systems considering heterogeneity and Inter-component dependencies.

1.4 Load-sharing dependence

Inter-component dependencies between components can be classified into three categories: economic, structural and stochastic [26]. Economic dependence implies that maintenance actions can either save costs (economies of scale) as compared to individual maintenance or become costly in the form of safety requirements, or production losses. Structural dependence occurs if components structurally form a part, such that maintenance of a failed component implies maintenance of working components. Stochastic dependence occurs when the condition of a component influences the lifetime distribution of another component. Load-sharing dependence is a form of stochastic dependence which occurs when multiple components can share the total system load and the load of a failed component is taken up by the remaining working components leading to an increased deterioration in these components [22]. In practice, this applies for example to a set of pumps that are used to distribute gas [27]. Some studies found that the costs from ignoring load-sharing dependence increase significantly with both the number of components and the degree of dependence [22, 28, 29].

Two modelling perspectives have been studied in the literature: failure-based load-sharing and deterioration-based load-sharing. For failure-based load-sharing, upon failure of a component, the remaining components will deteriorate according to the accelerated deterioration process, until it fails, or maintenance is performed [13, 30, 31]. Deterioration-based load-sharing is where component deterioration can also increase the load on the other components [27].

The most important element of the load-sharing model is the rule that governs how the loads on the working components change after some components in the system fail [32]. The equal load-sharing rule which assumes that a constant system load is distributed equally among the working components is the most studied form of load-sharing dependence [32]. Examples of equal load-sharing in practical situations include yarn bundles and untwisted cables that spread the stresses uniformly after individual component failures [32]. For example, Yun et al. [33] considered a consecutive k-out-of-n: F system composed of n identical components with exponential failure distribution.

Do et al. [34] determine optimal inspection intervals for a two-component series system where degradation of one component affects the degradation rate of the other. Liang and Parlikad [35] studied complex systems, where some critical items can be considered as being in parallel, although more than one may be needed for the system to function.

So far, various techniques based on statistical and probabilistic analyses (including the Markov process, Monte Carlo simulations, proportional hazards model and accelerated failure time models, or a combination of them) have been proposed to predict load-sharing system reliability [32]. For example, Amari et al. [36] proposed a closed-form analytical solution to evaluate the reliability of load-sharing *k-out-of-n: G* systems with tampered failure rate (TFR). Amari and Bergman [37], based on the cumulative effect (CE) load-sharing model, presented a method to compute the reliability of k-out-of-n: G load-sharing systems with identical or non-identical components, both following general failure distributions. Zhang et al. [22] characterised the relationship between the failure rate of a component and the load imposed on the component using an accelerated failure time model. Wang et al. [38] studied a load-sharing parallel system with failure dependency using a semi-Markov process. They introduced a dependence function to quantify the failure dependency between components. Zhang et al. [39] developed reliability model of the load-sharing *k-out-of-n: F* system subject to discrete external load in which the occurrence of the external load follows a nonhomogeneous Poisson process (NHPP). Kim and Kvam [40] proposed a maximum likelihood estimation approach for a load-sharing system with equal and monotone load-sharing rule. The work was further extended by considering a parallel system with Weibull distributed components [41].

Most studies on load-sharing systems focus on the determination of reliability or availability function, statistical inference, and maintenance of systems with equal load-sharing, very few are focused on local load-sharing systems. Some recent works have been conducted on systems with equal load-sharing [39, 42, 43, 44, 45, 46, 47, 48, 49]. In contrast, only a handful of recent studies have looked at reliability study on local load-sharing problems [50]. Local load-sharing sharing considers that a load on a failed component is transferred to an adjacent component (that is, due to spatial dependence). Spatial dependence is the propensity for nearby locations to influence each other and to possess similar attributes [51]. Examples of practical local load-sharing systems include cables supporting bridges and other structures, composite materials with bounding matrix joins, and transmission systems [32]. Another example involves a warehousing system in which a bigger warehouse is surrounded by several smaller warehouses such that when the bigger warehouse fails, all the smaller warehouses will take the load. When a smaller warehouse fails, the load will be redistributed among the bigger warehouse and the nearby smaller warehouses [52].

We note that most previous studies that consider the load-sharing in reliability prediction have not distinguished the uniform load distribution associated with equal load-sharing from the non-uniform load distribution associated with local load-sharing. Only a handful of works [50, 52, 53, 54] consider the non-uniform distribution of load for reliability prediction. However, deterministic failure rate values were assumed for each component's load change. Given the presence of non-uniform distribution of loads, the normal failure rate and several accelerated failure rate values will have to be predefined which can be challenging for a complex system. In general, unlike equal load-sharing models that are concerned with capturing the effects of uniform distribution of loads on the failure rate, there is no clear load-failure rate model that captures the impact of non-uniform distribution of loads on the failure rate in a local load-sharing setting. A model that captures the load-failure rate relationship can aid the assessment of factors that influence the system's reliability.

Furthermore, two key characteristics of spatial dependence are not captured in the existing local and equal load-sharing models. First is the topology of the system (i.e., how components in the system are arranged). For instance, Wang et al. [53] noted that for a large intelligent system with spatial dependent components, the spatial pattern among the system's components influences the degree of load-sharing. Second is proximity of components in the system. According to Amari et al. [32], the proportion of the load that the surviving components inherit depends on their distance to the failed component. No methodology has been developed for incorporating these two key characteristics into the load-sharing model for reliability prediction. In addition, none of the studies on local load-sharing have investigated the consequence of ignoring spatial effect if these exists. Previous work has, however, considered the impact of assuming independence when local load-sharing exists in a system (for example see [50, 52, 53, 54]).

1.5 Unobserved Heterogeneity in repairable systems

In practical situations, heterogeneity exist in the inner states of the system and the related working environments [55]. For example, a manufacturing system may produce different products under different workloads [55]. Sometimes systems from the same category may exhibit various failure/degradation processes in the same environment. The latter refers to unit-to-unit heterogeneity and is commonly studied (for example, [56, 57, 58, 59, 60]).

Unit-to-unit heterogeneity may be due to the variability in the inner structures of the considered system, as well as the diversity in their working environment [55, 61]. The variability in the inner structures could be due to material variation, manufacturing variation, process variation, installation variation, operation variation, variation in maintenance procedures and so on [21, 55, 61]. All these variations, cause the unit-to-unit heterogeneity in failure intensity or performance degradation. Furthermore, the factors of influence, also referred to as covariates, which cause the differences can either be observed or unobserved. Covariates describe the system's characteristics or the environment in which the system operates, and they may have varying levels. For example, to describe the reliability of a pump, vibration as a covariate may be, at a high, low, or medium level [61]. Observed covariates are factors whose effects on the failure process are known, and their associated levels are recorded with the failure data [61]. Unobserved covariates are covariates whose effects on the failure process are typically unknown and their associated levels during the operating time or at the time of the failure are not available in the failure database [61].

Unobserved covariates may lead to unobserved heterogeneity [62]. Unobserved heterogeneity refers to the effect of the unknown, unrecorded, or missing covariates [61]. Consider for example, that some pumps may have a soft foot problem in a production process due to a defect in the installation process. The soft foot problem may put the bearing in an over-stressed situation. If there is no information regarding soft foot in the failure database of the bearing then an unobserved covariate could be defined to capture the effect of the soft foot on the reliability of the bearing [61].

Models with random effects are the common technique employed to capture the unit-to-unit heterogeneity when the deterioration process is modelled and to predict the system reliability or remaining useful life [24, 56, 57, 58, 62, 63]. The most typical way to do so is to specify some parameters of the model as random variables governed by distributions with computing convenience, presenting the individuality in deterioration processes from different units, and leaving the rest of the parameters as constants describing the universality in deterioration of systems from the same category or batch [55, 64, 65]. For example, Lin et al. [57] incorporated gamma random effect into a piecewise constant hazard model to explore the impact of a locomotive wheel's position on its service life and to predict its other reliability characteristics. Using a gamma distributed random effect, Lin et al. [58] compared classical reliability test models with Bayesian piecewise constant hazard frailty model and found them to be useful for analysing degradation data.

In the past years, there has been a rising interest in developing models that capture unobserved heterogeneity in the failure process of repairable systems. For example, Lindqvist et al. [63] developed a heterogeneous trend renewal process model, which generalises the homogeneous Poisson process (HPP) and NHPP, to capture unobserved heterogeneity in multiple repairable components. They introduced a gamma-distributed multiplicative factor on the failure intensity. D'Andrea [66] suspected heterogeneity in the failure time data for mining trucks in Brazil. They assumed that the mining trucks were subject to minimal repair and thus modelled the data using NHPP with a gamma-distributed frailty term. Lindqvist and Slimacek [67] extended the basic NHPP to include covariates and unobserved heterogeneity in analysing wind turbine failure data. Yin et al. [68] applied a generalized accelerated failure time frailty model to study systems subject to imperfect preventive maintenance.

We note that most studies on unobserved heterogeneity for repairable systems with the minimal repair assumption have focused on investigating the significance of covariates and the frailty term in the fitted model rather than event prediction for the system and/or individual components (for example see [62, 63, 67]). The few studies that have considered event prediction for point processes with unobserved heterogeneity include: Deep et al. [69] who used a semi-parametric Andersen and Gill model for failure prediction of a new component in a Teleservice system using collected data from old units; and Jahani et al. [70] who developed a multivariate Gaussian convolution process (MGCP) for fleet-based event prediction in which failure prediction for an individual unit is conducted using data collected from other units. However, neither of these studies developed a parametric model considering systems subject to minimal repairs. The ability to predict the occurrence of failure events at system or an individual unit level can aid optimal maintenance decision making for the system or individual components [69].

Furthermore, in a model with unobserved heterogeneity, it is necessary to define the distribution of the unobserved effects [71, 72]. Since the modelled heterogeneity is unobservable, the appropriate choice of distribution of the unobserved effects is not easily discernible [73, 74]. The choice of the distribution of unobserved effects can give interesting general results in terms of the variance of the unobserved effects [74]. For instance, a large variance could indicate deficiencies in the choice of the distribution which may influence the model fit [73, 74]. It is therefore useful to examine the extent to which misspecification of the random effect distribution affects the validity of intensity function estimators [73]. The impact of misspecification of the random effect distribution with the minimal repair assumption has not been investigated.

1.6 Research Aim and Objectives

The aim of this research is to develop models for reliability prediction of systems subject to unobserved heterogeneity or load-sharing with spatial dependence. For systems subject to unobserved heterogeneity and minimal repair, we consider that the components in the system becomes homogeneous over time (see section 2.4.1). For the system subject to load-sharing, we consider two types of systems. First we consider a system with load-sharing, spatial dependence and proximity effect (see section 2.4.2). Second, we consider a system with load-sharing, spatial dependence, proximity effect and external shocks (see section 2.4.3). The following objectives are to be achieved:

1. Develop a new parametric model to capture the unobserved heterogeneity and provide a method for system and individual component event prediction.

    - The first objective of this research is to develop a new model to predict component and system reliability for systems subject to minimal repair and unobserved heterogeneity effects. This new model extends current research by releasing the assumption that treats unobserved heterogeneity as being constant and develops a method to estimate individual frailty values that enable component-level reliability prediction. Most existing methods have only focused on the case of heterogeneity being constant with time and modelled by gamma distribution [24, 58, 62, 75]. The case when unobserved heterogeneity changes and becomes homogeneous over time has not been studied for a system subject to minimal repair. Currently, existing models for reliability prediction of systems subject to minimal repair and unobserved heterogeneity effects cannot make component-level predictions. These models were often applied for statistical fit (for example, [24, 62]) rather than explicit prediction of reliability at the system or component-level. This research addresses these issues.

2. Develop a new probabilistic model to capture load-sharing and proximity effects for reliability prediction of a load-sharing system subject to spatial dependence and proximity effects.

    - The second objective of this research is to develop a new model to predict the reliability of a load-sharing system with spatial dependence and proximity effects. This new model extends current research by removing the assumption that the load of a failed component is shared equally by all working components. Industrial experiences have shown that there are a number of situations where the load of a failed component taken up by its spatial neighbour depends on how close they are to each other. Models for reliability of this type of system has not yet been studied. Here, we extend existing literature to address the issue by incorporating proximity effects in failure rates of components with load-sharing and spatial dependence.

3. Develop a new probabilistic model to capture load-sharing, proximity effects, and external shocks for reliability prediction and maintenance optimization of a load-sharing system subject to spatial dependence, proximity effects, and external shocks.

    - The third objective of this research is to develop a new model to predict the reliability of a load-sharing system with spatial dependence, proximity effects and external shocks. This new model extends current research by removing the assumption that the load of a failed component is shared equally by all working components in a load-sharing system subject to external shocks. For load-sharing systems subject to external shocks, industrial experiences have also shown that there are a number of situations where the load of a failed component taken up by its spatial neighbour depends on how close they are to each other. Models for reliability prediction and preventive maintenance of this type of system has not yet been studied. Here, we extend existing literature to address the issue by incorporating proximity effects in failure rates of components with load-sharing, spatial dependence and external shocks and develop a model for preventive maintenance of the system.

1.7 Structure of the thesis

This section provides the structure of this thesis with a brief summary of the content in each chapter. The thesis is structured as follows:

- Chapter 2 provides a review of key literature on multi-component systems with unobserved heterogeneity, stochastic dependence and stochastic dependence modelling. The review, informs the direction of the thesis, the modelling framework and the research objectives.

- Chapters 3, 4 and 5 develop three reliability prediction models.

- Chapter 3 presents the first reliability prediction model for systems subject to unobserved heterogeneity effect in which the component failure time data is in the form of recurring failures and modelled by the power law NHPP model. Random effects introduced in the power law NHPP model are modelled by an Inverse gaussian (IG) distribution. A comparison of gamma and IG distributed random effects is presented and a method for failure prediction for individual component based on empirical Bayes framework is developed.

- Chapter 4 presents the second reliability prediction model for systems subject to load-sharing, spatial dependence and proximity effect. Systems with homogeneous and heterogeneous components are considered. The failure process of the system is modelled by a Markov model. An analytic function which accounts for load-sharing, and proximity effect is developed considering cases when distance information is available or not available.

- Chapter 5, presents the third reliability prediction model for systems subject to load-sharing, spatial dependence and external shocks. The failure process of the system is modelled by Monte Carlo method and Modified failure sequence diagram. External shocks are modelled by HPP and NHPP. An analytic function which accounts for the load-sharing, proximity effect and external shocks is provided. A preventive maintenance policy based on an extension of the classical age-replacement model is developed.

- Chapter 6 brings together the results of the work done in the thesis by summarising this research and the discussions of future research is presented.

Chapter 2

Review of literature on heterogeneity, stochastic dependence effect and modelling for multi-component systems

This chapter reviews previous literature published with respect to models that account for unobserved heterogeneity, stochastic dependence and spatial dependence in reliability assessment and maintenance optimization of multi-component systems. The aim of this chapter is to establish general concepts relevant to subsequent chapters, identify popular research areas within the existing studies and to identify the research gaps.

According to Heping [76], "in order to analyse the reliability of systems, predict the remain useful life or minimize the maintenance cost, the first step is to model the physical properties of components or systems such as how they degrade and when they fail". This implies modelling the failure process of a component or a system through mathematical models. Although the primary concern here is multi-component system with unobserved heterogeneity, stochastic and spatial dependence effect, it is sensible to start the review by introducing general concepts and basic models which underpin the modelling of the failure process of a single component.

The remainder of this chapter is organised as follows. Section 2.1 presents failure modelling for a single component and likelihood inference. When considering failures of multi-component systems, unobserved heterogeneity or dependencies as well as the structure of the system should be taken into account. Therefore section 2.2 presents system configurations, multi-state system reliability evaluation, unobserved heterogeneity and dependencies concepts considered for modelling failures of multi-component systems. Section 2.3 presents review of literature on stochastic dependence modelling between components. Also presents review of literature on stochastic dependence and spatial dependence modelling. Section 2.4 identifies the research gaps in the literature.

2.1 Component failure modelling

The failures of components in a system occur either as recurrent or non-recurrent events. Non-recurrent event refers to a single observation of an event (e.g., a fault or failure) while recurrent events refers to several observations of the same event. In reliability theory, recurrent and non-recurrent events occur in repairable and non repairable components respectively. A non-repairable component is one where after a failure has occurred the component stops functioning and the component cannot be repaired (e.g., light bulb). A repairable component is defined in the literature as a component that is considered repairable if after a failure, its activity can be satisfactorily resumed through repair without the need to replace the component (e.g., water pipe) [77]. Bursts from water pipes represent an example of recurrent failure occurrence while the one-time failure of a light bulb represents an example of non-recurrent failure. In a non-repairable component, interest of a reliability engineer generally lies in the time until failure [78]. In contrast, repairable components can be brought back to working condition after a failure has occurred so that other failures may possibly follow.

Reliability analysis of components are performed using field data or accelerated failure time data. The purpose of collecting such data, is to study the mechanisms that may give rise to the failures by modelling the rate of failures per unit time [78]. The failure of components occur in time according to some underlying physical process which has a defined structure (e.g., corrosion, or natural ageing of components). In addition, the failures occur in time according to the defined structure with some uncertainty or randomness. This uncertainty may occur due to natural variability or external factors affecting the process [19]. Statistical models are appropriate for failure modelling because of their ability to allow for such randomness [78]. For instance, Economou [78] stated that statistical modelling of recurrent events, in pipes, is beneficial both for better understanding of the mechanisms that drive the recurrence and for predicting how these events will behave in the future. In sections 2.1.1, 2.1.2 and 2.1.3, we will introduce the methods used for modelling non-recurrent failures, recurrent failures and Likelihood inference respectively.

2.1.1 Non-recurrent failure modelling

As mentioned in section 2.1, non-repairable components are those that do not get repaired when they fail. Thus, non-repairable components can fail only once. The interest in failure modelling of a non-repairable component, generally lies in the time until failure. Lifetime models provides the distribution of the time to failure of such components [62]. In lifetime models, components are assumed to have only two states (either functioning or failed). Before introducing the lifetime models, some concepts are introduced as follows. Let T be a random variable that represents the lifetime of a component.

Definition 2.1 Lifetime distribution function is the probability that a component fails up to and including t time units. If T is a continuous random variable, it can be expressed as:

$$F(t) = P(T \leq t) = \int_0^t f(x)dx, \tag{2.1}$$

where $f(x)$ is the probability density function of T and t is the length of the period of time (which is assumed to start from time zero).

Definition 2.2 Reliability is the probability that a component performs a given function during t time units as:

$$R(t) = P(T > t) = (1 - F(t)) = \int_t^{\infty} f(x)dx. \tag{2.2}$$

Definition 2.3 The failure (hazard) rate function is defined as:

$$\lambda(t) = f(t)/(1 - F(t)), \tag{2.3}$$

provided that function $F(.)$ is differentiable.

The exponential distribution is one of the first statistical models to be applied to failure data for modelling hazard (failure) rates of components [7]. The exponential distribution is a basic lifetime distribution whose failure rate is constant and so is an excellent model for failures of components that occur by chance (at random) [7]. The exponential distribution has only one parameter which can not fit all types of lifetime data. As a result, distributions like weibull distribution with two parameters (shape and scale) are more flexible in applications [76]. Nelson [79] noted that the exponential distribution is often wrongly used to model failure data that requires the use of more sophisticated distributions. Murphy et al. [80] believes that exponential distributions are not appropriate for modelling the failure behaviour of products that age and are rather suited for failure modelling of only electronic devices because electronic equipment do not age.

Whilst the exponential distribution is only limited to modelling mid-life stages of components, the weibull distribution is more robust statistical model that can account for all phases of a product's operational life [76]. Weibull distribution could either be a two-parameter distribution with only the shape and scale parameters or three–parameter distribution when it includes the location parameter [7]. Unlike the exponential model, the weibull assumes a component's failure frequency could take any of the forms: decreasing, constant or increasing which makes it even more useful for modelling the various aspects: early life, mid-life or wear-out stages of a component's life [7]. When the shape parameter is 1, the weibull's failure rate reduces to that of an exponential distribution with constant failure rate; when the shape parameter is greater than 1, the failure rate is increasing which means that the deterioration speed of a component is accelerated; on the other hand, when the shape parameter is less than 1, the weibull's failure rate is decreasing which means that the deterioration speed is decelerated. Erturk et al. [81] used field data to perform reliability analysis of electronic boards. They combined exponential and weibull distributions to the same data to estimate different hazard rate at various service time intervals.

Other models have also been applied for failure modelling in the literature. For example, the normal distribution is found useful not only for modelling wear - out failures but also for modelling failure modelling failure mechanisms based on stress and strength properties of a product [79]. In contrast, normal distributions do not fit well with data pertaining to failures in the early stages or mid – life of a product [7]. Log-normal distribution is another distribution which is also useful for modelling stress and strength, and for modelling life data of products that show increasing failure rate in their early stages and decreasing failure rates in the later stages [7, 79].

Whilst the aforementioned models require one to make an assumption of a component's failure behaviour, in practice, prior specification of components failures behaviour may not always be possible or obvious [7]. Hence the need for non - parametric lifetime models. Kaplan-Meier estimator is one of the commonly used, non-parametric models for such scenarios. It is used to estimate the proportion out of a number of components that will still be functioning at a certain time. Kaplan-Meier involves the use of probabilities to estimate the proportion of surviving components. In this method, the probability of survival of a component at a certain point in time is calculated based on the cumulative probabilities of each preceding time. For example, for estimating mean time to failure of components in an experiment, Kalaiselvan and Rao [82] compared the mean time to failure (MTTF) from Weibull distribution with those from non-parametric distributions.

Although components in a system deteriorate with usage time, they may also be affected by other factors such as the operating environment, the maintenance effect which might have impact on the deterioration rate [55]. In order to take the factors into consideration, the proportional hazards model was proposed by Cox in 1970s [83]. The proportional hazards model is thereby powerful for modeling lifetime distribution to take account not only the age but also covariates which may influence the lifetime distribution into consideration [76].

The advantage of the proportional hazard model is that, without making any specific assumptions about the form of the baseline hazard function, it is able to analyze experimental data, compute maximum likelihood estimates and use likelihood ratio tests to determine which explanatory variables are highly significant [76]. The weibull and exponential hazard functions are mostly used as baseline hazard function. You and Meng [84] studied preventive maintenance scheduling of systems subject to the effect of imperfect preventive maintenance actions and variable operating conditions. There are other forms of the proportional hazard model (See Gorjian [85] for extensions of the proportional hazard model).

In this section we introduced lifetime models for failure modelling of single failures of components. In practice some components are repairable and as such have recurrent failure occurrences. In section 2.1.2 we will briefly introduce methods for modelling recurrent failures.

2.1.2 Recurrent failures modelling

Failures of a repairable component may occur with some randomness in time. The modelling framework used for recurrent event data which behave stochastically in time is the one involving counting processes [66]. Counting processes belong to the family of stochastic processes which describe the evolution of time dependent random variables according to some specific stochastic structure [78]. Other examples of stochastic processes include the gamma process [75], wiener process [64], and IG process [65]. As mentioned in section 2.1, the mechanisms that may give rise to failures are studied by modelling the rate of failures per unit time. A counting process is characterised through the intensity function $\lambda(t)$. Unlike the hazard function in lifetime models, which can be considered as instantaneous rate of failure, and the intensity function $\lambda(t)$ is a global failure rate in the sense that it accounts for rate of failures over multiple time intervals $(t_a; t_b]$.

Furthermore, in modelling recurrent failures from repairable components using counting process, the degree to which the component is repaired is factored in the model. There are three main classification of repair actions: perfect repair, minimal repair and imperfect repair. A perfect repair means that the component is brought back to the same working condition as if it was new (as good as new). In this case, the times between subsequent failures must necessarily be identically distributed. the times between subsequent failures might be identically distributed. In addition, if these failure times are independent, then the scenario defines a renewal process. One renewal process commonly used in reliability literature is the HPP where the times between failures are exponentially distributed, and its the failure rate is constant in time. The number of failures in any time interval $(t_a; t_b]$ in a renewal process follows a Poisson distribution.

When minimal repair is performed, the component is brought to the condition it was just before the failure occurred (as bad as old). In that case, the times between subsequent failures might not be identically distributed. This defines a NHPP. In contrast to the HPP, NHPP has intensity function $\lambda_0(t)$ which depends on time but still retains the property that the number of failures in an arbitrary time interval follow a Poisson distribution. A third repair assumption used in the literature is an imperfect repair which may bring a component to any condition between as good as new and as bad as old and could be modelled as a superimposition of renewal process and NHPP model [86]. In section 2.1.2.1, nonhomogeneous Poisson process model is introduced.

2.1.2.1 Nonhomogeneous Poisson process

Ascher [87] introduced a model whose reliability is allowed to decrease with time, calling this the "bad-as-old" concept. The idea is that the inter-failure times are dependent in time (but not on the failures themselves). In that case, the times between subsequent failures might not be identically distributed. This defines a NHPP. NHPP has intensity function $\lambda_0(t)$ which depends on time but still retains the property that the number of failures in an arbitrary time interval follow a Poisson distribution.

A typical repairable product that experiences ageing will have the characteristic that during its early life it will fail more frequently due to defects or installation errors

[20]. This period is sometimes called the burn-in or the debugging period. After that, the "useful" life of the pipe begins where the failure rate is approximately constant and smaller than that of the early failures period. Towards the end of its lifetime, ageing has a significant effect on deterioration and the failure rate increases sharply. As a result, the overall shape of the failure rate curve resembles a bathtub [20].

A common method for modelling the failure rate $\lambda_0(t)$ of a NHPP is the power law model which is sometimes referred to as the Weibull model or the Duane model [88, 89]. A NHPP with power law intensity is often referred to as the Power Law Process (PLP). As the name suggests, this model describes $\lambda_0(t)$ by raising time to a power:

$$\lambda_0(t) = \theta\rho t^{\rho-1} \quad t > 0; \rho > 0; \theta > 0, \tag{2.4}$$

where θ is the scale parameter and ρ is the shape parameter that controls the shape of the curve. $\rho = 1$ implies that the rate is constant (HPP); $\rho > 1$ means that the function is increasing and $\rho < 1$ gives a decreasing $\lambda_0(t)$. By definition, $\lambda_0(t) = 0$ at $t = 0$. It is worth mentioning that when using a power law for the NHPP, the distribution of time to the first failure is Weibull and the distributions of inter-failure times are truncated Weibull distributions where the truncation occurs at the time of the immediate previous failure [20]. For that reason, the NHPP with a power law intensity is also referred to as the Weibull process. A paper by Lawless [90] considered the NHPP with an intensity function in which the baseline intensity function is either semi-parametric or parametric. The author also discussed inclusion of random effects in terms of a random intercept in the failure rate. Landers et al. [91] used the semi-parametric proportional intensity model introduced by [92] and compared it with a log linear NHPP. Other examples include Pulcini [93] that used a superposition of two power law functions to capture the complete bathtub shape; Ryan [94] that explored a NHPP with a time dependent mixture intensity function and Krivtsov [95] that presented a family of parametric models to represent $\lambda_0(t)$.

In summary, NHPP model is useful for capturing the uncertainty of component's failures. As this thesis is interested in the data of systems with unobserved heterogeneity where the observed number of failures is in a count form and the system's failure rate changes over the time, NHPP is useful for modelling such data. NHPP model has been frequently applied in many fields, for instance, noise exposure modelling [96], repairable system reliability [62], software reliability analysis [97]. NHPP can deal with failure data which is often considered non-stationary [98]. In addition, power law model is flexible to describe various types of system's phases [99]. Economou [20] note that power law model enables analysing systems where the rate of occurrence of failures may be a decreasing, constant, or increasing function of time, respectively. Such characteristics might be exhibited in systems with unobserved heterogeneity. One limitation of the NHPP model is that its model parameters are fixed and cannot reflect unit-to-unit variation [69]. On the other hand, NHPP can be extended to include covariates and random effect terms to reflect unit-to-unit variation using the cox model framework [24, 67, 69]. The structure of the extended NHPP model is flexible such that various unit-specific factors can be incorporated as covariates [69]. In sections 2.1.1 to 2.1.2 the modelling of the failure process of a single component have been discussed. In section 2.1.3 we will briefly introduce procedures for drawing inferences from data.

2.1.3 Likelihood inference

Likelihood inference is based on the principle that given a statistical model, all information from the data that is relevant to inferences about the value of the model's parameters are contained in the likelihood function [100, 101]. Likelihood inference forms the basis of classical methods like maximum likelihood estimation, and it plays a key role in Bayesian inference [102]. Likelihood inference has been widely used to study heterogeneity and dependence effects (See [24, 40, 41, 62, 63, 74, 103]). Likelihood inference is useful for several reasons. First, likelihood inference provides a simple way to deal with mis-specified models [104]. Second, likelihood inference has attractive asymptotic properties and has good small-sample behaviour even when models are mis-specified [105, 106]. Third, likelihood inference allows one to compare models [101]. In contrast, likelihood inference approach may not be preferable, for example, if we only care about accounting for one dimension of the data, a task that a Method of

Moments can be more suitable [101]. Method of Moments estimates model parameters by matching the theoretical moments from the distribution with the observed moments from the data [107]. Compared to likelihood approaches, Method of Moments may suffer strong biases resulting from using small samples and may not efficiently use all the existing information [101]. For the above reasons, likelihood inference will be adopted in this thesis. In section 2.1.3.1, 2.1.3.2, and 2.1.3.3, Maximum likelihood estimation, Bayesian inference and Empirical Bayesian method which are three established likelihood inference procedures will be discussed.

2.1.3.1 Maximum likelihood estimation

Likelihood inference provides a framework for estimating unknown parameters of a system in a way that the maximum likelihood estimate (MLE) is the parameter value that makes the observed data most probable [108]. MLE is one of the most established parameter estimation methods because it is intuitively appealing and it has many desirable statistical properties [108, 109]. Under some reasonable conditions on the statistical model, it can be shown that MLEs have the following asymptotic properties:

- Unbiasedness: the expected value of the MLE is equal to the true parameter value.
- Efficiency: the MLE has the smallest possible variance among unbiased estimators of the parameter.
- Normality: the sampling distribution of the MLE is Normal (Gaussian).
- Consistency: the MLE becomes arbitrarily close to the true parameter value.

These asymptotic properties hold as the sample size increases to infinity [108]. Of course, no real dataset ever has an infinite number of observations [108]. For finite sample sizes, these properties may not hold; for instance, the MLE may be biased (systematically high or low relative to the true value of the estimate). Errors in the estimates obtained from ML estimation method can be large especially when sample sizes are small [110]. However, it is often the case that these properties are approximately satisfied in practice for reasonable sample sizes [108]. Some examples of the MLE application in reliability literature can be found in [24, 62, 111].

2.1.3.2 Bayesian inference

Like the use of likelihood in maximum likelihood estimation procedure, likelihood functions are also a key component of Bayesian inference. Suppose we have data D and a likelihood model describing the distribution of the data with some unknown parameter θ. Bayesian methods work by firstly placing some prior belief on the parameters' distributions $p(\theta)$, the resulting posterior distribution of the parameters is calculated by the Bayes rule:

$$p(\theta \mid D) \propto p(\theta) \times p(D \mid \theta), \tag{2.5}$$

where $\propto$ means "is proportional to." It is worth noting that the shape of the posterior depends not only on the prior and the observed data, but also on the sample size [108]. If there exists a finite mean for the posterior distribution, then the posterior mean is a method of estimation and is written as:

$$\tilde{\theta} = E[\theta] = \int \theta p(\theta \mid D) d\theta. \tag{2.6}$$

Bayesian inference and MLE are similar in a way that when the sample size is large, Bayesian inference often provides results for parametric models that are very similar to the results produced by MLE [112]. However, compared with Bayesian inference, MLE requires a large data set (or sample size) to achieve statistically significant results [7]. Furthermore, under the MLE, model parameters are interpreted as fixed whereas the Bayesian approach interprets the parameters as random variables [113]. Bayesian methods provide a framework for dealing with unobserved heterogeneity problems due to some reasons. First, given that unobserved covariate is commonly modelled by introducing a random effect term [55], the uncertainties in the unobserved covariate can be modelled using Bayesian method. Gelman et al. [114] note that Bayesian method is useful for summarizing uncertainty, making estimates, and predictions using probability statements which are conditional on observed data and an assumed model. Second, the uncertainty about the unobserved covariate can be reduced as more information is obtained. Bayesian methods provide a way of combining prior information with data such that past information about a parameter can be incorporated to form a prior distribution for future analysis [108]. In contrast, there are some limitations with the use of Bayesian methods. First, there is no correct way to choose a prior. Bayesian inferences require skills to translate subjective prior beliefs into a mathematically formulated prior and require considerable problem structuring to assess uncertainties with all possible events [110]. Furthermore, the quantification of an expert's uncertainty can result in a significant cognitive burden and so needs to be managed with care [110]. Some examples of applications of Bayesian inference to heterogeneity problems can be found in [115, 116, 117].

2.1.3.3 Empirical Bayes method

Empirical Bayes method is like the Bayesian method described in section 2.1.3.2 in the sense that it combines the likelihood function and a prior distribution for statistical inference. However, while Bayesian priors about the probability of an event are usually constructed from subjective beliefs [118], empirical Bayes provides a means of pooling observed data to form an empirical prior [110]. In that regard, empirical Bayes can be classical statistics [119]. Empirical Bayes methods can be categorized as either parametric empirical Bayes or nonparametric empirical Bayes. In parametric empirical Bayes, one assumes that the prior distribution of the parameters is in some parametric class. In nonparametric empirical Bayes, one assumes that the parameters are independent and identically distributed [119]. Some examples of applications of empirical Bayes include Green and Strawderman [120] that used empirical Bayes method to estimate individual tree volume equation development. Brown [121] used empirical Bayes method to predict the batting averages of players in Major League Baseball. Vesely et al. [122] discussed application of empirical Bayes method to emergency diesel generators for binary data. Vaurio [123] used empirical Bayes method for estimating the rate of common cause failures.

In summary, empirical Bayes method is useful for dealing with unobserved heterogeneity problems for the following reasons. First, empirical Bayes avoids the limitation of traditional Bayes in the sense that unlike the traditional Bayes method which relies on subjective opinions, empirical Bayes relies on observed data [119]. Based on this principle, empirical Bayes could allow the quantity of interest to be described by a probability distribution and provide estimates by pooling empirical data from multiple similar objects [124]. This can be very useful for estimating the random effect term for components whose data is not available. Second, empirical Bayes method allows the use of the observed data to estimate some features of the prior distribution [119]. To estimate the prior parameters, a prior predictive distribution can be constructed which represents the distribution of the observed data. Then, based on the Bayes theorem, the prior distribution can be updated to obtain a posterior distribution as new observations become available [125]. Lastly, empirical Bayes can be combined with other parameter estimation methods such as Maximum Likelihood Estimation and Method of Moments for estimating the prior parameters [125]. Based on the above, the estimation of the random effects at the component level is one of the problems in studies on unobserved heterogeneity and empirical Bayes could be useful for estimating the random effects terms. As noted by Slimacek and Lindvist [74], the ability to estimate random effects at the component level is important for predictions and can provide important additional information about the modelled system, however, it is often overlooked. An application of empirical Bayes method to unobserved heterogeneity problem can be found in [69].

So far, we have introduced general concepts and basic models which underpin parameter estimation and the modelling of the failure process of a single component. One might also be interested in modelling system level failure behaviour. In this case, unobserved heterogeneity in component failure processes, component dependence mechanisms as well as the structure of the system are considered. In section 2.2 we will briefly introduce multi-component systems failure modelling by considering: system configurations in section 2.2.1, multi-state system reliability evaluation 2.2.2, the approach for modelling unobserved heterogeneity in section 2.2.3 and dependence concepts in section

2.2.4.

2.2 System configurations, multi-state system reliability evaluation, component heterogeneity and dependencies

2.2.1 System configurations

The reliability requirement for products is increasingly set very high in order to meet a predefined performance level. As such, maintaining a high reliability for an entire system, requires that the individual components which make up the system have extremely high reliability [126]. In addition, the functional arrangement of the components in the system can have consequences for the system's reliability [76]. Some of the main configurations for modelling system reliability in literature are classified as series, parallel, series-parallel, and k-out-of-n.

A series system configuration is one that is arrange such that the system functions if all components in the system are functioning. In a series system, each component is essential for the proper functioning of the system. The implication of a series configuration for a maintenance policy can be in two ways. On one hand, the failure of a single component can lead to high unavailability costs, so maintenance actions are performed at a relatively early stage to prevent downtime [27]. On the other hand, the series structure can imply that the complete system may have to be stopped to perform maintenance on a certain component which offers an opportunity to perform maintenance actions on the other components and possibly save cost [27]. The implications for reliability of the system is that a series system is not sensitive to stochastic dependence because the system fails if one component fails [76].

A parallel system functions properly if at least one component in this system is functioning. In a parallel configuration, only one component needs to function, while any other component failure is typically assumed to have no impact on the system performance. Parallel configuration of a system is a way to prevent a system's unavailability through the use of redundancy [27]. The redundancy occurs in a way that only a few components may be required for the system to work, thus the remaining components could be placed as back up [127]. The implication of a parallel system configuration for maintenance policies is that maintenance actions on the system can be performed at a relatively late stage [27]. Unlike a series system, the implications of a parallel configuration for a system's reliability is that, the system become less reliable when stochastic dependence effect is present on the components of the system [76].

Other system configurations include: series-parallel system configuration which is composed of several series subsystems in parallel; a parallel-series system which is composed of several parallel subsystems in series [128]; and k-out-of-n system which is a system that functions if at least k number of components in this system are functioning. Both the series and parallel systems are special cases of the k-out-of-n system, and can be obtained by setting k equal to n or 1, respectively [76].

Sometimes reliability model for a system may not be based on any of the above mentioned configurations [129]. Reliability model of complex systems are an example of such situation. A complex system is composed of a number of components which cannot be reduced to a series-parallel system [129]. Hwang et al. [129] provides a review of some complex system structures.

2.2.2 Multi-state system reliability evaluation

Every engineered system is designed to perform an intended task. Some systems can perform their tasks with various distinguished levels of efficiency usually referred to as performance rates [130]. A system that can have a finite number of performance rates is called a multi-state system (MSS) [130]. The performance rates exist because engineering systems consist of different components whose availability have a cumulative effect on the entire system's performance. As such different numbers of available units can provide different levels of the task performance as well as the states of the system as completely working, total failure or partial failure [130, 131]. In contrast, a binary system is the simplest case of a MSS having two distinguished states (perfect functioning and complete failure). Reliability models suitable for reliability analysis of binary systems have been introduced in sections 2.1.1 and 2.1.2. The number of system states that needs to be considered as well as the increasingly high requirements for accurate reliability evaluation and optimal design make it difficult to use traditional binary reliability techniques [130]. Reliability analysis considering multiple possible states is known as multi-state (MS) reliability analysis. Multi-state system reliability models allow both the system and its components to assume more than two levels of performance. From reviewing reliability literature, there are three common methods that are often used for MSS reliability assessment: Monte Carlo methods, stochastic process approach and universal generating function approach. In section 2.2.2.1 to section 2.2.2.3, a review of methods for multi-state system reliability is presented.

2.2.2.1 Monte Carlo methods

Monte Carlo methods, also known as Monte Carlo simulation, are often used in simulation of physical and mathematical systems [132]. The idea of Monte Carlo method is the generation of random events in a computer-based model, the generation of random events are repeated many times and the number of a specific event is counted [132]. With increasing computer speed, Monte Carlo methods have been widely applied to studies on heterogeneity and dependence. For example, Zio et al. [133] proposed a Monte Carlo simulation approach that allows modelling the complex dynamics of multi-state components subject to operational dependencies with the system overall state. Asfaw and Lindqvist [62] study the consequence of ignoring unobserved heterogeneity in system reliability prediction. Zio and Podofillini [134] used Monte Carlo simulation to estimate all the importance measures of the components at a given performance level in a multi-state series-parallel system. Taghipour and Kassaei [13] proposed an optimization model based on simulation for periodic inspection of a k-out-of-n load-sharing system over its lifecycle. Monte Carlo methods can be used to solve the reliability prediction problem of almost every real world system that cannot be solved analytically [16, 131]. However, efficient Monte Carlo algorithms are often difficult to develop. Another issue of concern is the time and expenses involved in the development and execution of the method [130].

2.2.2.2 The universal generating function

Universal generating function (UGF) method is an analytical modeling approach for reliability assessment of multi-state systems [135]. UGF is a method to describe multi-state components and construct the overall model of complex multi-state systems. UGF uses the system operating states,its corresponding probabilities, as well as system configuration to obtain the system reliability analysis [136]. For application of UGF, Wang et al. [137] and Ding et al. [138] used UGF to model wind generators. Levitin et al. [139] presented an algorithm for evaluating performance distribution of complex series–parallel multi-state systems with propagated failures and imperfect protections. Li and Zio [135] used UGF for reliability assessment of a distributed generation system.

2.2.2.3 Stochastic process models

Stochastic process models are commonly used for the MSS reliability analysis. Stochastic process models, as applied for MSS, are of two forms - discrete or continuous. A continuous-time stochastic process is a stochastic process for which the index variable t takes a continuous set of values, as contrasted with a discrete-time process for which the index variable t takes only distinct values [140]. Continuous-time stochastic process have been used for MSS reliability analysis to describe the degradation of systems. In the case of degradation modelling, stochastic process models describe the evolution of one or more degradation parameter by gradual, stochastic increments over time, and the failure of a system occurs when the degradation parameter value reaches a predefined threshold [141]. The degradation process in a system is usually traced by measuring one index or multiple ones [140]. The degradation measures can be visibly observed or not. For example, the crack growth in civil infrastructures can be visually recorded, while the reduction in a voltage output of an electronic device cannot be observed and could be recorded by monitoring sensors [140]. A degrading system is assumed to have more than two states. The states include perfectly functioning, several partial failed states and complete failure. The partial failed states are intermediate states which are unobservable or hard to be measured [76].

Some examples of works on degradation modelling include: Welte et al. [142] that model the degradation process of hydro power plant by a Markov chain with four finite states where the sojourn time of the process is modeled by gamma distribution. Lin et al. [143] studied reliability assessment of systems subject to dependent degradation processes and random shocks using a combination of multi-state models (MSM) and physics-based models (PBM). The MSM and PBM were used to describe the degradation process in a discrete way considering limited degradation data. Kharoufeh and Cox [144] estimate the residual life distribution for a single-unit system subject to Markovian environment-based degradation with finite states. They consider that the environment follows a time-homogeneous Markov chain with finite states with different degradation. Continuous-time stochastic process models such as gamma process [75], wiener process [64], and IG process [65] have also been used for degradation modelling.

Another stochastic process for modelling MSS is a Markov chain. A Markov chain is a stochastic process with a discrete state space and discrete time space [140]. Markov chains with continuous-time are referred to as Markov processes [140, 145]. The model assumes that the conditional distribution of a future state of the system is determined only by the present state and not by previous state or the time at which it reached the present state [146]. Thus, the sojourn time of each state is exponentially distributed and the transition probability to each state is independent of the process history. The Markov model evaluates the probability of jumping from one known state into the next logical state. The process continues until the system being considered has reached the final failed state or until a particular mission time is achieved. Natvig and Streller [147] first applied the Markov model to MS system reliability evaluation. Xue and Yang [148] proposed combining the Markov chains and coherent structure function. Sometimes a Markov chain cannot describe a systems deterioration very well, hence a semi-Markov process is used to model the reliability of the system [16]. A semi-Markov model is an extension of the Markov chain with discrete states and continuous time [149]. For instance, Wang et al. [38] studied a load-sharing parallel system with failure dependency using a semi-Markov model. They introduced a dependence function to quantify the failure dependency between components.

Markov models are useful for reliability prediction of systems with dependencies for several reasons. First, the assumption that a system has a finite state space and a series of possible transitions between these states suits many systems. For instance, Markov chains are one of the first MSS models that found application for the reliability assessment of multi-state power systems and some types of communication systems even before MSS was theoretically defined [130]. Second, using a Markov model, various failure modes, changing failure rates, standby and maintenance activities all can be described as different states and all corresponding transition probabilities can be derived [16]. For instance, Shao and Lamberson [150] examined the reliability and availability of shared load repairable k-out-of-n: G system with imperfect switching. Sharifi et al. [151] studied reliability of a k-out-of-n: G system with increasing failure rates. Liang and Parlikad [35] developed a condition-based maintenance strategy for an asset with critical components and non-critical components. Broek et al. [152] investigated joint condition-based production and maintenance policies for two-unit systems with economic dependency and whose units have adjustable production rates. Third, Markov models are flexible for describing different evolution of systems, for instance, when the system evolves in time according to the same transition rate as from the beginning [153] or different transition rates [154]. However, sometimes, it is not easy to find all transition probabilities as they may be numerous for large systems [16, 130]. Furthermore, the Markov equations are difficult to solve analytically for some systems [16]. In section 2.2.2 we have introduced the MSS models used for system reliability evaluation. In section 2.2.3 we will introduce methods for modelling unobserved heterogeneity effect for system reliability prediction and review key literature.

2.2.3 Modelling unobserved heterogeneity effect in component failures

Traditional reliability models such as Weibull and lognormal models deal with the simplest case of independent and identically distributed data [61, 62]. The models assume that the study population is homogeneous. When studying the problem of predicting the behaviour of a system based on failure data from several similar systems, there may be unobserved heterogeneity among the systems which, if overlooked, may lead to inaccurate prediction of a system's failure behaviour [62, 69, 155]. Unobserved heterogeneity among systems may be due to differences in material, differences in design, differences in location and so on [21, 24]. An example of such a case is a manufacturing system which may produce different products under different workloads [55].

If covariates are known, they can be included in the analysis, but it is often impossible to include all important risk factors, either because one has little or no information at the component level [61]. Sometimes, we may not know the relevance of the risk factor or even that the factor exists [156]. In other cases, it may be impossible to measure the risk factor without great financial cost or time effort [156]. As a result, there could be specific individual factors that result in unobserved heterogeneity of the lifetimes which cannot be captured by observed covariates [61]. In section 2.2.3.1, 2.2.3.2 and 2.2.3.3, mixture distributions, frailty models and frailty distributions commonly used for modelling unobserved heterogeneity will be discussed. A summary is provided in section 2.2.3.4.

2.2.3.1 Mixture distributions

A mixture distribution is the probability distribution of a random variable that is derived from a collection of other underlying random variables each selected with a given probability [157]. The mixture distribution is a multivariate distribution. The distributions of the underlying random variables that are combined to form the mixture distribution are called the mixture components, and the probabilities (also referred to as weights) associated with each component are called the mixture weights [157]. Let T_i be a positive-valued random variable. The distribution of T_i is defined as a mixture of lifetime distributions if its density function can be represented as:

$$f(t_i \mid \psi, \theta) = \int_n f(t_i \mid \psi, w_i) dP(w_i \mid \theta), \tag{2.7}$$

where $f(t_i \mid \psi, w_i)$ represents the density function of a lifetime distribution parameterized in terms of ψ and w_i. w_i is a realized value of a random variable W_i which has distribution function $P_{W_i}(w_i \mid \theta)$.

The mixture distribution has the property that the reliability function retains the same structure as in Eq.(2.7) but is not valid for the hazard function. A similar representation for the hazard function exists, however, it involves a different mixing distribution. A study of the distributions generated by Eq.(2.7), and its properties is presented in [158]. This approach intuitively leads to flexible distributions based on a known distribution by mixing over a parameter. Mixture models are flexible in the sense that varying the underlying model, generates a wide class of lifetime distributions [159]. A wide variety of shapes and tails can be generated by Eq.(2.7) which can accommodate unobserved heterogeneity [156, 159]. Some applications of Mixture models to unobserved heterogeneous can be found in [160, 161, 162]. Furthermore, mixture modeling can be interpreted as the introduction of a random effect on the reliability distribution. In the literature, the random variable W_i is referred to as frailties, a term that was originally introduced by Vaupel et al. [163]. In this context, the model in Eq.(2.7) is usually called univariate frailty model [159]. Some examples of frailty modelling can be found in [62, 164, 165].

Mixture model has occasionally been used to describe a class of models that are referred to as compound probability distributions. A compound probability distribution is the probability distribution that results from assuming that a random variable is distributed according to some parametric distribution whose parameters themselves are random variables (latent variables) [162]. The compound distribution is the result of marginalizing (integrating) over the latent random variable representing the parameter of the parametric distribution. For example, a Poisson-Gamma mixture derived from a Poisson distribution whose rate parameter is assumed to be a random variable following a Gamma distribution. The resulting compound probability distribution (Poisson-Gamma mixture) can be shown to follow a negative binomial distribution [162].

Compound probability distributions have been widely applied to study data that are often characterized by overdispersion [162, 166]. For instance, Zou et al. [167] proposed the Sichel generalized additive models to handle severe dispersion in crash data. The Sichel (SI) distribution is a mixture of Poisson distribution and generalized inverse Gaussian distribution. Wu et al. [162] proposed the negative binomial (NB) model (also called Poisson-gamma) as a good alternative to the Poisson model for analyzing crash data where the variance is greater than the mean. [166] proposed the Poisson inverse Gaussian (PIG) model to the same data set.

There is a relationship between the compound probability distribution and the notion of posterior predictive distributions in Bayesian models. For instance, consider data where some parts of the data are observed, and the others are missing. If we want to make predictions about the missing data points given the latent variable and the observed data, a posterior predictive distribution like the Bayesian posterior predictive distribution is derived for the missing data points [168]. The posterior predictive distribution for a missing data $\tilde{x}$ is given by:

$$p(\tilde{x} \mid x) = \sum_{z} p(z \mid x) p(\tilde{x} \mid z; x), \tag{2.8}$$

where $p(z \mid x) = \frac{p(z|x)p(z)}{\sum_z p(z|x)p(z)}$ is the posterior distribution of z given the observed data x.

2.2.3.2 Frailty model

A frailty model is a random effects model for failure data, where the random effect (frailty) has a multiplicative effect on the baseline hazard function [156]. It can be used for univariate (independent) lifetimes, that is, to describe the influence of unobserved covariates in a proportional hazards model (heterogeneity) [156]. Frailty model assumes that the hazard of an individual depends on an unobservable, age-independent random variable W, which acts multiplicatively on the baseline hazard function:

$$\lambda(t; W) = W \lambda_0(t), \tag{2.9}$$

where W is considered as a random mixture variable, varying across the population. This frailty model is commonly referred to as a univariate frailty model [164].

The term frailty itself was introduced by Vaupel et al. [163]. In its simplest form, frailty is an unobserved random proportionality factor that modifies the hazard function of an individual, or of related individuals [164]. The key idea of frailty models is that individuals have different frailties, and that the frailest will die earlier than the lesser frail. Consequently, systematic selection of robust individuals takes place, which distorts what is observed [156]. This may be useful where changes in the hazard function is of interest. An example could be death rates of cancer patients where the longer the patient lives beyond a certain time, the better his or her chances of survival are. In this case, the population intensity starts to decline simply because the high-risk individuals have already died. In contrast, the intensity of a given individual might continue to increase [156]. Note that frailty modelling is random effect modelling in the sense that a model parameter is a random variable. However, unlike traditional random effect modelling where the random variables are allowed to be positive or negative, in frailty models the random variables are only allowed to be positive random variables [169].

Whilst random effects models are similar to Bayesian methods because they assume a distribution for the parameters, random effect models including frailty models are not Bayesian [170]. To be Bayesian, a distribution will be specified for the random effects, a second set of distributions will be specified for each parameter that defines the random effect distribution, and distributions will be specified for all the fixed effects parameters in the model [170]. In order words, all parameters are random effects under the Bayesian paradigm [170]. In contrast, if we want to predict a random effect for a component, one will use the component's data and resort to Bayes rule. Here the random effects distribution works like a prior and the method can be described as empirical Bayes [171]. Some examples of Bayesian frailty modelling can be seen in [116, 172, 173].

Some extensions of the univariate frailty model are Shared frailty models, Correlated frailty models and Cure rate models. Shared frailty models are one of the most popular extensions of the univariate frailty model. It aims to account for correlation between clustered observations [164, 174]. These models are used for grouped datasets where, conditional on the observed covariates, survival times are assumed to have the same distribution within each cluster [164]. In such a case, the frailty terms take a common value for all individuals belonging to the same group [175]. As a result, the frailty term introduces intra-cluster dependencies. A survey about this subject can be found in

Duchateau and Janssen [175]. Correlated frailty models assume that assigning the same frailty value to all observations within a cluster is not always appropriate and attempt to model intra-cluster variations as well [165]. Hence, correlated frailty models assign a joint distribution to the mixing parameters associated to each group (for example [117, 175, 176]). Cure rate models are suitable for situations where there is a proportion of individuals who will never experience the event of interest and another that will experience the event of interest [160]. Following a medical nomenclature, individuals that will not experience the event of interest are commonly labelled as cured units [156, 160]. Frailty models accommodate these type of event data by using a mixing distribution that assigns a positive probability to not observing the event of interest. In this case, the hazard function is equal to zero [156]. For example, Price and Manatunga [161] combined the cure model with a frailty distribution and showed that the extended cure model can handle both the heterogeneity with the frailty distribution but also the cured fraction as well.

The study of heterogeneity is widely applied in various fields like politics [177], and medical research [178]. In survival analysis, the effects of unobserved heterogeneity have been studied in several papers including but not limited to [163, 165, 172, 174]. Apart from unobserved heterogeneity across individuals, the dependence of the events experienced by an individual have also been studied using correlated frailty models [179]. This is because the recurrent events of an individual are possibly correlated because of underlying characteristics of the individual [180, 181]. In survival analysis, the papers that have studied correlated events are not limited to [103, 117, 172, 182].

For non-repairable systems, some examples include [56, 57, 58] that used the frailty model to analyse the lifetime of locomotive wheels. Lin and Asplund [56] considered an integrated data approach to reliability assessment of locomotive wheels by considering both degradation data and re-profiling data. They studied the log linear life stress model and weibull baseline failure rate. Lin et al. [57] incorporated gamma shared frailty model into a piecewise constant hazard model to explore the impact of a locomotive wheel's position on its service lifetime and to predict its other reliability characteristics. Lin et al. [58] compared classical reliability test models with bayesian piecewise constant hazard frailty model and found them to be useful for analysing degradation data.

For repairable systems, unobserved heterogeneity has been studied with the minimal repair assumption. Early works include [183, 184]. Lindqvist et al. [63] developed a heterogeneous trend renewal process model, which generalises the HPP and NHPP, to capture unobserved heterogeneity in multiple repairable components. They introduced a gamma distributed multiplicative factor on the failure intensity. D'Andrea [66] suspected heterogeneity in the failure time data for mining trucks in Brazil. They assumed that the mining trucks were subject to minimal repair and thus modelled the data using NHPP with a gamma distributed frailty term.

Asfaw and Lindqvist [62] investigated heterogeneous population composed of independent NHPP using gamma-distributed frailty. Lindqvist and Slimacek [67] extended the basic NHPP to include covariates and unobserved heterogeneity in analysing wind turbine failure data. Lindqvist and Slimacek [74] developed the method for parameter estimations in heterogeneous NHPP population when the distribution of frailty is unspecified. The NHPP model was extended to include covariates in [24]. Most research on minimal repair have parametrically modeled heterogeneity using the gamma frailty model in which unobserved effects are assumed to be gamma distributed.

For studies on repairable systems that consider perfect or imperfect repair assumption, Lindqvist et al. [185] developed a trend renewal process model which is a general case of NHPP and renewal process. They considered gamma distribution and power law model as baseline. The TRP model was illustrated on Tractor engine and Air conditioner data. Yin et al. [68] applied a generalized accelerated failure time frailty model to study systems subject to imperfect preventive maintenance. Liu et al. [186] investigated the effect of heterogeneity on the failures of repairable systems that undergo imperfect repairs. The basic model they considered is the Kijima type II virtual age process with constant repair efficiency and a Weibull baseline distribution. The heterogeneity between the systems were assumed to be gamma-distributed.

Unobserved heterogeneity effect has also been studied on degradation data when components observed over time are assumed to degrade at different rates, even though there is no apparent difference in treatment or their environments (e.g., unobserved heterogeneity in degradation data modelled by: gamma process [75], wiener process [64], IG process [65]. A review on degradation modelling for systems with unobserved heterogeneity can be found in [55].

Another thing worth noting is that frailty modelling approach is also useful for studying spatial dependence with respect to spatial correlated events. Spatial correlated events occurs when it is suspected that there is spatial correlation between observations in different geographical locations [187, 188]. In biostatistics literature the approach has been used for studying spatially correlated survival data and for analysing patterns in health and ecological related data (see for example [187, 188, 189, 190, 191, 192]). When frailty model is applied for studying spatial dependence effect, log normal distribution is used. Log normal frailty models are useful in modelling dependence structures in multivariate frailty models [193]. The frailty model makes use of a single parameter to capture the spatial dependence.

2.2.3.3 Frailty distributions

The choice of the frailty distribution is very important in the area of frailty modelling. A suitable distribution for heterogeneity modelling is one with a positive random variable. The shape of the distribution also plays an important role in frailty modelling. The tails of the distributions can determine the type of dependence a frailty model describes. Distributions with a large right tail such as positive stable distribution led to strong early dependence, whereas distributions with a large left tail such as gamma, and Weibull distributions lead to strong late dependence [181].

There are various frailty distributions that have been suggested in the literature, however, the gamma and IG distributions are the two commonly used distributions. According to Gachau [169], some of the other suggested distributions are not used in practice due to software limitations and the lack of sound estimation procedures for more complex frailty models. The one-parameter gamma distribution, which was first proposed by Clayton [194], is the most popular frailty distribution since it is very tractable. Hougaard [195] also proposed the gamma, and the IG distributions for frailty model. Oakes [196] introduced the IG and log-normal models for the distribution of frailty. We will focus on the IG and gamma distributions in this thesis.

The gamma distribution takes variety of shapes as the shape parameter varies. When the shape parameter equals 1, the distribution is identical to the exponential distribution. The gamma distribution has simple density function. Although gamma frailty models does not have closed form expressions for reliability and hazard functions, from a computational view, it fits well to frailty data and it is easy to derive the closed form expressions for unconditional reliability and hazard functions. For this reason, this distribution is used often in most applications [197]. When a gamma distribution is applied to model frailties, the frailty terms in the conditional likelihood function can be integrated to give a simple expression for the marginal likelihood. Thus, it is easy to obtain parameter estimates of a gamma frailty model by maximizing the marginal likelihood. The likelihood construction is discussed in more detail in chapter 3.

Early applications of gamma frailty model can be found in several disciplines. For instance, Lancaster [198] used this model for the duration of unemployment. Vaupel [163] used the gamma distribution in their studies on population mortality data from Sweden. Aalen [199] studied the expulsion of intrauterine contraceptive devices. Ellermann et al. [200] studied recidivism among criminals using gamma-Weibull model. Andersen et al. [201] used the gamma frailty model to check the proportional hazards assumptions in his study of malignant melanoma. The gamma distribution has two advantages as a frailty distribution. The frailty distribution of the survivors at any given age is again a gamma distribution, with the same parameter and a different scale parameter [197]. The second advantage is that the frailty distribution among the persons dying at any age is also a gamma distribution, with the same shape parameter plus one, and a scale parameter as a function of the age at death [197].

As an alternative to the gamma distribution, Hougaard [202] introduced the IG as a frailty distribution. The IG distribution has a history dating back to 1915 when Schrodinger and Smoluchowski presented independent derivations of the density of the first passage time distribution of Brownian motion with positive drift. The IG distribution has a unimodal density and is a member of the exponential family. Its shape resembles that of other skewed density functions, such as lognormal and gamma distribution. Gacula et al. [203] analysed shelf life of several products using the IG law and found the distribution to be a good fit. Vilmann and Svend [204] have studied the histomorphometrical analysis of the influence of soft diet on masticatory muscle development in the muscular dystrophic mouse. The muscle fibre size distributions were fitted by an IG. Barndorff [205] considers a finite tree whose edges are endowed with random resistances, and shows that, subject to suitable restrictions on the parameters, if the resistances are either IG or reciprocal IG random variables, then the overall resistance of the tree follows a reciprocal IG law. For more real-life applications (see [206]).

Chhikara et al. [207] studied the IG distribution and found that there are many striking similarities between the statistics derived from this distribution and those of the normal distribution. These properties make the IG potentially attractive for modeling purposes with survival data [165]. The IG distribution has some advantages as a frailty distribution. It provides much flexibility in modelling, when early occurrences of failures are dominant in a lifetime distribution and its failure rate is expected to be non-monotonic [165]. In such situations, the IG distribution might provide a suitable choice for the lifetime model. Also, IG is almost an increasing failure rate distribution when it is slightly skewed and hence is also applicable to describe lifetime distribution which is not dominated by early failures [165]. Secondly, Hougaard [202] remarked that survival models with gamma and IG frailties behave very differently, noting that the relative frailty distribution among survivors is independent of age for the gamma, but becomes more homogeneous with time for the IG. The conclusion was derived by observing the coefficient of variation of the two frailty distributions of survivors. For the gamma distribution, this is a constant. For the IG distribution, the coefficient of variation is a decreasing function of time. The IG distribution has a shape that resembles the other skewed density functions, such as log-normal and gamma. Duchateau and Janssen [175] fit the IG frailty model with Weibull hazard to the udder quarter infection data.

2.2.3.4 Summary on unobserved heterogeneity

Traditional reliability methods in reliability analysis assume that populations are homogeneous, meaning all individuals have the same risk of failure. As mentioned above, it is often important to consider the population as heterogeneous, that is, a mixture of components with different failure rate.

Mixtures of life distributions is useful for modelling unobserved heterogeneity problems, particularly when traditional models are not able to capture this feature of the data for the following reasons. First, a wide variety of shapes and tails can be generated by Eq.(2.7) which can accommodate unobserved heterogeneity [156, 159]. Second, if an (underlying) distribution is underpinned by theoretical or practical reasons, the same reasons hold for the mixture model in the presence of unobserved heterogeneity [159]. Conditional on the mixing parameters, lifetimes are distributed as in the underlying model but with a different value w_i for each individual component. Lastly, the extent of unobserved heterogeneity is controlled by the spread of the mixing distribution. If n is a finite set of values, the distribution of T_i is a finite mixture of life distributions whereas if n contains a single value, the mixture recovers the original underlying distribution (no unobserved heterogeneity) [159]. Discrete mixtures of lifetime distributions are explored in [208, 209, 210] among others. For the case in which w_i in Eq.(2.7) is a continuous random variable, some studies can be found in [163, 174, 175].

Mixture modeling can be interpreted as the introduction of a random effect on the reliability distribution and the random variable W_i can be referred to as frailties. Thus, the mixture model in Eq.(2.7) can be called univariate frailty model [159]. Frailty models which are based on hazard functions have been widely used for studying unobserved heterogeneity for certain reasons. One of the reasons this model is so popular is because of the ease with which technical difficulties such as censoring, and truncation are handled by hazard-based models [156]. This is due to the interpretation of the hazard function as a risk that changes over time [156]. In addition, the model allows for the entering of covariates in order to describe their influence and to model different levels of risk for different subgroups. However, in general it is impossible to include all relevant risk factors, perhaps because we have no information on individual values, which is often the case in many studies (for example in demography [156] and in reliability [61]). Lastly, beyond representing unobserved heterogeneity between specific individuals, frailty models can also accommodate more complex data structures [159]. For the above reasons this thesis will focus on modelling unobserved heterogeneity using frailty models. So far, we have presented methods for modelling unobserved heterogeneity effect for system reliability prediction and key literature. In section 2.2.4 we will introduce dependence concepts used for system reliability prediction.

2.2.4 Dependence concepts

Dependencies occur due to several factors which range from the design of the system, maintenance actions on the system's components, and shared environmental conditions to name a few [25]. Thomas [26] was first to identify the forms of dependencies (interactions) that occurs between components in a system. These were identified as; economic, structural or stochastic.

Stochastic dependence involves the influence of a state change between component. The state change might be caused by degradation factors like age or failure rate [211]. Economic dependence occurs when components are joined by the cost of their collective spare parts. Here it is assumed that it cost less to buy replacement parts for a group of components [211]. Structural dependence is mostly observed in coupled systems where a working component has to be removed in other to repair the failed component hereby resulting in the downtime of the working component [211]. Stochastic dependence focuses on interaction between the failure mechanisms of components [27]. Economic dependence focuses on the maintenance activities performed on the components while structural dependence is concerned with the placement of components in the system [27].

Keizer et al. [27] introduced a new form of dependence as resource dependence. According to Keizer et al. [27], resource dependence occurs when multiple components are connected through, e.g. shared spares, tools, or maintenance workers, budget and so on. See Keizer et al. [27] for more on resource dependence. In section 2.2.4.1

and 2.2.4.2, structural and economic dependencies will be briefly discussed respectively while stochastic dependence will have a deeper review in section 2.3 because we are interested in spatial dependence with respect to stochastic dependence between components.

2.2.4.1 Structural dependence

Structural dependence occurs if components structurally form a part, such that maintenance of a failed component implies shutting down other working component [26]. Structural dependence concerns the structural, and static relationships between different components. Of all the main classifications of dependencies studied in the literature, structural dependence is the least studied without considering other dependence mechanism [27]. Structural dependence is mostly observed in coupled systems where a working component has to be removed in other to repair the failed component hereby resulting in the downtime of the working component [211]. The reason for this is that structural dependence involves explaining how one component's performance, maintenance or failure affects another because of the way the system is coupled [27]. Camci [212] studied a system in which a working component is shutdown by the maintenance personnel due to failure of or maintenance action on another component.

Keizer et al. [27] classified structural dependence as technical dependence or performance dependence. Technical dependence involves restrictions that one might encounter in the usage of or maintenance of components because of a system's highly technical configuration. Performance dependence occurs when the performance of a system depends on both the performance of its components and their configuration within the system. Performance dependence also referred to as functional relationships of series, parallel, k-out-of-n and so on which were introduced earlier in section 2.2.1.

From the definition of technical dependence, one may notice two forms of restrictions emerging which are based on maintenance and usage. Usage restrictions occurs when the failure of or maintenance on a component can have negative impacts on other components. One example concerns the processing of milk, where different processes have to take place shortly after each other, and all related components (such as dryers and mixers) are coupled via pipelines [27]. Maintenance restrictions occurs in a situation where in order to perform maintenance actions on a certain component one might need to either perform maintenance on other components or prohibit maintenance on the others. An example concerns the tires of an airplane, which are required to have the same thickness and are thus replaced jointly [27]. Dismantling working components to fix a faulty component, because they block access to it, could make them damaged as well and require maintenance [27].

2.2.4.2 Economic dependence

Economic dependence pertains to whether maintenance actions can either save costs (economies of scale) or become costly in form of safety requirements, or production losses [26]. The dependence occurs when components are joined by the cost of their collective spare parts. Here it is assumed that it cost less to buy replacement parts for a group of components [211]. The two forms of economic dependence identified in the literature are negative and positive economic dependencies.

Positive economic dependence occurs when maintaining several components simultaneously is cheaper than maintaining them separately [213]. This occurs, for instance, when high costs are involved in travelling to a location where maintenance activities have to be executed on windmills at sea [27]. In contrast, a system is subject to negative economic dependence when maintaining several components simultaneously leads to higher costs than maintaining them separately. According to Nicolai and Dekker [214], such negative economic dependence can be present in systems with manpower restrictions, safety requirements, or production losses, which are usually incorporated via restrictions in the maintenance model. However, such restrictions are examples of structural dependence or resource dependence [213].

Given that structural and stochastic dependencies can occur between components of a multi-components system, how to schedule an effective maintenance action on the components may be the next concern to keep the system functional [27]. Maintenance actions on a system incurs costs and the costs related to a certain maintenance policy can be influenced by the degree of economic dependence. As a result, a maintenance manager may have to decide whether it is better to perform group maintenance or individual maintenance given the economic dependence [27].

Now that we have introduced component failure modelling, system configuration, unobserved heterogeneity, and dependence concepts; in section 2.3 we will review key literature pertaining to the forms of stochastic dependence effect studied for multi-component systems. We will also review existing modelling approaches for studying stochastic dependence effect.

2.3 Stochastic dependence and modelling

Stochastic dependence occurs if the state of a component influences the lifetime distribution of other components, or if there are causes outside the system which bring about simultaneous failures and hence correlate the lifetimes [211]. Stochastic dependence means that the failure of one component can affect or modify one or more of the remaining components of the system. Stochastic dependence is also referred to as failure dependence or failure interaction in reliability literature.

Murthy and Nguyen [215], and [216] introduced the two forms of failure interaction as: Type I, and Type II. Type I failure interaction occurs when the failure of one component can either induce simultaneous failure of the other components or have no effect on them. Type II failure interaction occurs when the failure of a component only induces an increase in the failure rate of the other components. Nicolai and Dekker [217] put the Types I and II failure interactions together into one definition that the failure of a component could either affect the failure rate of other components or cause their immediate failure. This definition is sometimes referred to as Type III.

Whilst failure interaction addresses the dependence between components when failure occurs, degradation interaction can also happen. Degradation interaction occurs when the state of one component influences the state of other components without failure of the influencing component [27] (see [27], and [218] for details on degradation interactions for multi-component systems).

The review of stochastic dependence literature will be done in two parts. The studies that involve stochastic dependence will be classified either under failure interaction, or load-sharing. Then a review of the models used for studying each stochastic dependence effect will be presented. Failure interaction occurs when the failure of one component can cause damage to other components which can lead to instantaneous failure or increase in the deterioration level of these components. Load-sharing is when the load of a failed component is taken up by the remaining working components leading to an increased deterioration in these components. The review of failure interaction and load-sharing dependence will be presented in section 2.3.1 and 2.3.2 respectively. Lastly we will present the research gaps in section 2.4.

2.3.1 Failure interaction

Failure interaction occurs when the failure of one component can cause a major, one-time, damage to other components leading to an increase in their deterioration level or even an immediate failure of these components [27]. The concept can be seen in the definition of stochastic dependence by Dekker [211] as the situation where the state of a component influences the lifetime distribution of other components. For example, a propeller can come off of an airplane and pierce the fuselage, causing tremendous additional damage and safety risks [27]. Another example can be observed in a gearbox, where defects in a bearing will cause it to vibrate. The deterioration of the subsystem that includes related shaft and several gears can accelerate due to the excessive vibration caused by the bearing [219]. This forms of stochastic dependence will be referred to as failure interaction in this thesis.

Failure interaction interactions assumes that components in the system can be split into two groups: the influencing components and the victim components where the influencing components affects the victim component by some factors such as vibrations which acts as shocks (e.g., see Sun et al. [219]). Whether a system is well-designed or not, it is assumed that it is merely impossible to perfectly prevent an interaction with external factors [220]. According to Sung et al. [220], the effect of external shocks is particularly important for mechanical systems that often have a protective external component subject to the state of the surrounding environment. Nakagawa and Murthy [221] derived the optimal number of failures to minimize the expected cost per unit of a two-component system with shock damage interaction for an infinite time case. Satow and Osaki [222] extended the work of Nakagawa and Murthy [221] and proposed a two-parameter (T, k) replacement model for a two-component system subject to shock damage interaction. Where the first component is repaired whenever it fails. Failure of the first component occurs according to a NHPP and causes damage to the second component. The damage is accumulated in the second component. Lai and Chen [223], and Lai [224] proposed an economic periodic replacement model for a two-component system with failure rate interaction and with or without external shocks. The system is assumed to be completely replaced upon failure, or preventively replaced at age T, whichever occurred first. Sung et al. [220] derived a long-term replacement policy for a two-component system by combining the concepts of type II failure interaction and external shocks.

There are also studies that consider influence of shocks in multi-component systems. Jhang and Sheu [225] considered Type I interaction in a multi-component system. They compare maintenance strategies for a system with n components where each component i can be subject to a minor failure with a probability $(1-p)_i$ or cause a major failure to all other components that could stop the system with a probability p_i. Lai [226] extended their work on Type II interactions from two components to multi-component systems. They considered a system consisting of n components with one dominant (non-repairable) component and $(n-1)$ secondary (repairable) components. Li et al. [227] consider reliability assessment of a voting system with n components where one of them is dominant and non repairable and the others are secondary and repairable. They assume the failure of the any of the $(n-1)$ secondary components causes an increase in the failure rate of primary component.

In practice industrial systems tend to be more and more complex and multiple components can share a critical function or have more than one critical component [228]. Considering Type I interactions, Liu et al. [23] represent the interaction between components by a probability matrix P_{ij}. Thus, a component i was assumed to cause an instant failure to a component j with a probability p_{ij} and has no effect with a probability $(1-p_{ij})$. They identify some issues in the early stages of a system's design that could affect the predetermined warranty given interaction between component. In addition to reliability implications of failures on a system, they showed that failure interaction would in the long run have an extensive and accelerating effect on the decrease of a system's economic value. They define w as a system's period of warranty, $p_i(w)$ the probability of failure for a component i during w and $F_i(w)$ is the probability that a component i failure will cause the system failure, and $F_s(w)$ is the system's probability distribution. The model was applied to adjust the warranty cost of series and parallel structures considering the number of failures occurrences.

Although, there are systems that function in a way that the shock interaction is unidirectional, there are other systems where the shock interaction can be bidirectional. Sun and co-authors [229] and [219] studied a multi-component system in which the shock interaction could be bidirectional. They identify the interactions with an immediate effect and the interactions with a gradual effect. They assume the hypothesis accurately represents physical systems.

There are also considerations of failure mode interactions on a component. In practice, some system may be subject to diverse failure modes such that the modes can influence each other, mutually or not [228]. Zequeira and Berenguer [230] studied inspection policies for a two-component parallel standby system with failure interaction. They classified failure modes as maintainable and non-maintainable modes. The modes are distinguished based their repairability in the occurrence of a failure. Preventive maintenance corrects the deterioration due to maintainable modes while Non-maintainable modes can only be corrected by a complete overhaul of the system. Minimal repairs are considered in the case of failures. Castro [231] assumed that the occurrence of maintainable failures could be correlated with the number of non-maintainable failures denoted by $N_2(t)$ following the installation of a system. Fan et al. [232] a system whose component are subject to two failure modes that have bidirectional stochastic dependence. They assumed that the failure rate of mode i depends on the number of the failure in the other mode and vice versa.

One of the few research on failure interaction considering spatial interaction was presented by Levitin [233] who studied reliability analysis of a series-parallel multi-state system in which the failure of a component can cause an immediate failure of nearby component (known as Type I stochastic dependence). They assume local failure interaction only occurs between nearby components. They refer to the selective form of failure interaction as propagated failures. The study was motivated by the failure interaction between the pumps and reactors in a production module where the fires from a failed pump unit affects nearby reactors. They developed a model for reliability assessment based on universal generating function.

2.3.1.1 Failure interaction models

In the literature on failure interaction, application of analytical models of component failure rates is very common. In this case, the failure rates of several components or failure modes are linked by an analytical function. In general, the analytical function then has an incremental effect on the system's overall failure rate. Whatever structural and/or economic dependencies that exists within the system also influence the development of the function [228]. As a result, these models address configuration such as series, parallel, and so on. The analytical models are built from failure times in maintenance logs and do not explicitly address physical degradation processes [228]. Thus the accuracy of this models rely solely on the accuracy of the maintenance logs [228]. In contrast, the main point of this modelling approach is to introduce a certain proportionality in failure rate interaction that is intuitively assumed. The potential for chain reactions and the shared environment contribute to an acceleration of the failure rate.

Some papers applied analytical modelling approach for studying systems with two components referred to as two-component systems. Lai et al. [226], Lai and Chen [223], [234] and Lai [224] base their work on Type II failure interaction to determine a replacement policy with an optimal number of minimal repairs for a maintenance cycle of optimal duration T. They developed a model that considers one repairable (component 2) and one non-repairable (component 1). The failure rate of component 2 follows a NHPP of intensity $h_2(t)$. Every failure of component 2 increases the failure rate of component 1. Inversely, if component 1 fails, component 2 instantly fails. The expected failure rate of component 1 $h_1(t)$ is given by

$$h_1(t) = \sum_{0}^{\infty} h_1(t \mid N_2(t) = j) \times P(N_2(t) = j), \tag{2.10}$$

where $N_2(t)$ is the number of failures of component 2. $P(N_2(t) = j)$ is the probability of j number of failures of component 2. $h_1(t \mid N_2(t) = j)$ is the conditional failure rate of component 1 given that component 2 had j number of failures.

Golmakani and Moakedi [235] extended the model in Eq.(2.14). They considered a two-component system with Type II failure interaction. They assume the components are subject to two kinds of failures. In their work, they assumed that Component 1 is subject to soft failures following a NHPP while component 2 is subject to hard failures following a HPP. Each hard failure has an instant detectable effect and requires immediate intervention when they occur. The soft failures can only be detected by a scheduled inspection because they do not stop the system but decrease its performance. A component's hard failure can be the root cause of another component's soft failure if this component serves a secondary function (e.g., protective apparel) for a more critical component. In their work, a coefficient p was introduced to depict the percentage of increase of the failure rate of component 1 due to hard failures from component 2. The developed model is given by

$$h_1(t) = \sum_{0}^{\infty} (1 + \frac{p}{100})^j h_1^0(t) \times P(N_2(t) = j), \tag{2.11}$$

where $h_1^0(t)$ is the initial failure rate of component 1. They assumed that the estimation of the coefficient p would rely on experimental data or the availability of a physical model. However, assuming a constant incremental effect (p), suggests that the interaction have little to no variability. A limitation of the two models is that they can be restrictive since no retroactive effect is accounted for.

Sung et al. [220] derived a long-term replacement policy for a two-component system by combining the concepts of Type II failure interaction and external shocks. They assumed that the effect of external shocks is critical for mechanical systems that often have a protective external component whose state is subject to the state of the surrounding environment. They assume that the external shocks occur following a NHPP with intensity $r(t)$ and could cause minor failures with a probability p or catastrophic failures with a probability $(1-p)$ to a component (component 2). They assume that the failures of component 2 act as internal shocks as well and increase the failure rate of another component (component 1). The entire system fails if component 1 fails. The number of component 2 failures $N_2(t)$ follows a NHPP with intensity $h_2(t) = h_2^0(t) + pr(t)$. Where $h_2(t)$ is the baseline failure rate which is independent of the external shocks. Thus the failure rate of the system depends on the failure rate of component 1 and the number of external shocks endured by the other component. The expected failure rate of component 1 is

$$h_1(t) = \sum_{0}^{\infty} h_1(t \mid N_2(t) = j) \times P(N_2(t) = j). \tag{2.12}$$

The analytic model was also extended for studying multi-component systems. Li et al. [227] studied a series-parallel (voting) system with n components. They assumed that one of the components is dominant and non repairable while the others are secondary and repairable. The secondary components are mutually independent and follow exponential distribution. Unlike the study of Lai [226] the set of components interacting and failing can vary. The failures of secondary components increase the failure rate of the dominant component denoted by

$$h_p^c(t) = P(t \mid k_1, k_2 \ldots, k_{n-1}). \tag{2.13}$$

where k_j is the number of failures of the j^{th} secondary component. Where $h_p^c(t)$ is the failure rate of the dominant component and is assumed to depend on the conditional probability of failures of the secondary component. Thus the expected failure rate of the primary component is given as

$$h_p(t) = \sum_{k_1} \sum_{k_2} \ldots \sum_{k_{n-1}} h_p^c(t) P(N_1(t) = k_1) P(N_2(t) = k_2) \ldots P(N_{n-1}(t) = k_{n-1})). \tag{2.14}$$

Likewise Lai [226] extended their work on Type II interactions from two components to multi-component systems using the same modelling approach. They considered a system consisting of N components with one dominant (non-repairable) component and $(N-1)$ secondary (repairable) components. They assumed that the secondary components are mutually independent and follow NHPP. The failure of a secondary component increases the failure rate of the primary component. They proposed a replacement policy following the system's age and assumed that the secondary failures are corrected by minimal repairs.

Although, there are systems that function in a way that a single component is dominant, leading to a unidirectional interaction. In practice some systems tend to have more complex structure with components having bidirectional relationships [219] [228]. Sun et al. [219], and [229] studied systems where failure interaction between components is bidirectional. They classify the interaction as one that either occurs with an immediate effect or with a gradual effect. They assume that the interactions that occur with a gradual effect accurately represents physical systems. They considered a system with N identical components that has an independent failure rate $h_i^0(t)$ and an interactive failure rate $h_i(t)$, which depends on the failure rate of other components in the system. They distinguish an independent failure rate denoted by

$$h_i(t) = h_i^0(t) + \sum_{j=1}^{n-1} \theta_{ij} h_j^0(t). \tag{2.15}$$

where θ_{ij} is an interactive coefficient that captures how much influence the j^{th} component's failure rae will have on the interactive failure rate of component i. The derived analytical expression provides an idea of an update in a failure rate from independent to dependent by a linear combination with coefficients denoted by θ_{ij}.

Zhao et al. [236] extended the model by Sun et al. [219] and applied it to study a gyroscope. They demonstrate that the concept of interactive failures are indeed applicable and have a significant influence on the overall reliability of a system subject to multiple failure modes. Wang and Li [237] reprise the analytic model of Sun et al. [219], and [229] and propose an allocation method of redundancies in the design of a system. They demonstrate that the effect of interactive failures can be minimized if they are properly modeled in the system design process.

2.3.2 Load-sharing dependence

Load-sharing dependence occurs when the load imposed on a system is shared by multiple components in the system [22]. Systems with this kind of dependence mechanism are called load-sharing systems. According to Zhang et al. [42], Type II failure interaction between components is common in the load-sharing systems. In a load-sharing system, if a component fails the remaining components share its load and the increased load on the remaining components induces higher failure rate in them causing their reliability to be affected due to the increased load in form of added stress, current, and so on [22]. Improper treatment of such dependency can lead to misleading reliability assessment [22].

An example of a load-sharing system is seen in the operating process of cables in a suspension bridge [37]. Another example can be observed in a system consisting of several pipelines used to transport water, oil, or gas; whenever a pipeline segment fails, its load is transferred to other segments which are still operational [13]. Therefore, the components of a load-sharing system are stochastically dependent on each other due to the components sharing a specific amount of load [13].

A common feature in a load-sharing system, is that all components in the system comply by a certain set of load-sharing rule by which the load of a failed component is automatically re-distributed to the remaining components [238]. Some common load-sharing rules include equal load-sharing, monotone load-sharing [32], and local load-sharing. Equal load-sharing assumes the same workload is shared by all the remaining components [239]. Monotone load-sharing rule indicates that the load on the surviving components is non-decreasing even after the repair of the failed component while Local load-sharing rule implies that the load on a failed component is transferred to adjacent components [37].

The literature on load-sharing can also be viewed from the two types of load-sharing mechanism studied i.e., static load-sharing (constant total load) and dynamic load-sharing (time-varying total load). Studies with static load-sharing models assumes that the total load imposed on a system is fixed while dynamic load-models assumes that the load imposed on the system can change during the systems usage time [37]. According to Amari [37] majority of published works are based on the assumption of static load-sharing. Some works that consider static load-sharing problem include [240, 241, 242].

Due to the ease of application, a lots of studies of load-sharing systems assumes that components are identically distributed. Scheuer [243] obtained the reliability of k-out-of-n system when the component failure induces higher failure rates in survivors by assuming iid components with constant failure rates. Shao and Lamberson [150] examined the reliability and availability of shared load repairable k-out-of-n: G system with imperfect switching comprising n independent and identically distributed (iid) components with constant failure rates. Gupta [244] discussed load-sharing effects on reliability of k-out-of-n: G system. Tang and Wang [245] studied a load-sharing repairable parallel system with time varying failure rates. Sharifi et al. [151] studied a k-out-of-n: G system having n identical components with increasing failure rates. Yun et al. [33] considered a consecutive k-out of-n: F system composed of n identical components with exponential failure distribution.

In practice, components of a load-sharing system may not be identically distributed. As a result, there are studies that have considered systems with non identically distributed components. Altiol and Baykai-Gursoy [246] considered the load-sharing system with two types of load affecting properties under different failure rates. Liu [247] developed a model to assess the reliability of a load-sharing k-out-of-n: G system which consists of non-identical components with arbitrary failure distributions. Yinghui and Jing [248] studied a k-out-of-n load-sharing system with different components. The components were assumed to be non-repairable with exponentially distributed lifetimes. Huang and Xu [249] estimate the reliability of a load-sharing system with general life distributions. However, they make the assumption that the components are non-repairable. Jain [250] studied the reliability of a load-sharing system with common cause failures. Maatouk et al. [251] studied a series-parallel multi-states system with propagated failures and load-sharing occurs between the system's components. They developed an approach for analysing the system's reliability based on the hybridization of markov process and universal generating function.

Some works consider impact of load on degradation rate, Ye et al. [252] used the cumulative exposure model to describe the degradation process of components in the system. They assumed that degradation is the main reason for the system failure, and the relationship between the degradation of components and components cumulative shared workload follows the inversely Gaussian distribution. Yang et al. [253] considered a load-sharing system subject to degradation. They assumed that the value k in the k-out-of-n system changes as the degradation rate changes, and this variable k is derived by a performance degradation model. Zhao et al. [43] used a link function to describes the relationship between the work load and the degradation rate for a load-sharing system with degrading components.

Some researchers have also focused on statistical inference for load-sharing systems. For example, Kim et al. [40] proposed classical maximum-likelihood estimation for parameter estimation of a load-sharing system with equal and monotone load-sharing rule. They assumed that all the components have identical exponential distribution. Park et al. [41] extended the model of Kim et al. [40] by considering a parallel system with Weibull distributed components. Closed-form Maximum Likelihood Estimator and conditional Best Unbiased Estimator of the Weibull parameters were derived. Singh et al. [254] considered classical and Bayesian estimation of the parameters a load-sharing parallel system whose components in which some of the components are non-identical with respect to their failure-rates. They assumed that some of the components have constant failure rates while the others have linearly increasing failure rates. Singh et al. [255] also considered classical and Bayesian estimation of the parameters a load-sharing system whose components are arranged in parallel and each component's lifetime follows Lindley distribution.

Ibnabdeljalil and Curtin [256] studied the reliability and strength of fiber-reinforced composites under a local load-sharing condition in which they assumed that stress from broken fibers is transferred predominantly to the nearby unbroken fibers. Durham et al.

[257] and Ibnabdeljalil and Curtin [256] both used monte carlo simulation for system reliability analysis. Although the studies by Ibnabdeljalil and Curtin [256], and Durham et al. [257] were first to consider local load-sharing system in which their models were developed for specific application to fibers. Wang and co-authors [52, 53, 54, 258], proposed markov models to study the availability of markov repairable local load-sharing systems and considered circular, star, and lattice system structures. They assumed that if a component fails, its left and right neighbours will detect the change and take up its load. Guo et al. [259] developed a reliability model for a consecutive $(2, k)$-out-of-$(2, n)$ systems. They assumed that the system is subject to local load-sharing. A consecutive system assumes that the system fails if k consecutive adjacent component fails.

2.3.2.1 Load-sharing systems with shocks

In section 2.3.2, the studies reviewed only account for the workload shared by each surviving component. Sometimes, the components in a load-sharing system may not only bear the workload but also experience the external load, such as temporary shocks [50]. In reality, systems can be exposed to random shock processes, which are caused by technical conditions (e.g., overvoltage, overheating, mechanical stresses), environmental threats (e.g., extreme temperatures, wind gusts, floods) or malicious attacks [260]. The workload usually leads to the component internal degradation, while the shock load caused by temporary shocks results in the abrupt degradation. According to Che et al. [261] it is a common phenomenon that most of load-sharing systems are subject to degradation processes and random shocks simultaneously. As a result, competing failures may occur and any of them can cause the failure of the system [261]. Thus, shock load generated either during the intended operation of the systems or from external sources, is an important mechanism accounting for load-sharing systems failure [261].

For the load-sharing system with shocks, Ye et al. [252] used the cumulative exposure model to describe the degradation process of components in the system. They assumed that degradation is the main reason for the system failure, and the relationship between the degradation of components and components cumulative shared workload follows the inversely Gaussian distribution. Che et al. [261] propose a numerical reliability model to describe the dependent degradation and shock processes of the components in a global load-sharing system. They considered that once a component in the system fails, it would equally increase the work and shock load of the surviving components.

Only one paper was found which has considered local load-sharing, and shocks. Guo et al. [50] developed a reliability model for a consecutive k-out-of-n: F system with local load-sharing subject to dependent degradation and external shocks. They assumed that the components degradation processes may be different because their degradation rate is dependent on work load and abrupt degradation is dependent on shocks. They modelled the continuous degradation of each component as a linear degradation model and the abrupt degradation caused by the shocks by normal distribution and poisson process. They assumed that the workload and shock load on failed components will be equally shared by their adjacent neighbours. In sections 2.3.2 and 2.3.2.1 we reviewed research on reliability prediction, availability prediction and statistical inference of load-sharing systems. In section 2.3.2.2, we will review research on maintenance optimization of load-sharing systems.

2.3.2.2 Maintenance optimization of the load-sharing system

Most of the studies on maintenance policies for load-sharing can be largely grouped into inspection policies and condition-based maintenance policies. For systems whose deterioration are constantly monitored via sensors, condition-based maintenance has been studied. For example, Liang and Parlikad [35] developed a two-tiered approach to model and optimize the condition-based maintenance strategy for an asset with critical components and non-critical components that undergoes fault propagation and load-sharing between components. The optimization model was based on a continuous-time markov model. Keizer et al. [29] explored condition based maintenance for a parallel system that is subject to both failure dependence through load-sharing and economic dependence. The system's states were modelled using a Markov decision process. Broek et al. [152] investigated joint condition-based production and maintenance policies for two-unit systems with economic dependency and whose units have adjustable production rates. The system states were formulated as a Markov decision process in order to determine optimal policies.

For systems whose deterioration are not monitored via sensors but by physical inspection of the system, inspection policies for load-sharing systems have also been considered. For example, Taghipour and Kassaei [13] proposed an optimization model for periodic inspection of a k-out-of-n load-sharing system over its lifecycle. They considered a finite planning horizon in which the optimization model finds the optimal inspection interval for the system. They assume that the failures follow a NHPP. Liu et al. [239] developed constant and cumulative load models for assessing the reliability of load-sharing systems with continuously degrading components. The developed models were used to formulate preventive maintenance policies with periodic inspection. To evaluate the performance of the maintenance policy, they adopt a long-run cost rate model, where the preventive maintenance threshold and the periodic inspection interval are the two decision variables. Ahmadi [262] explored the joint determination of optimal inspection and threshold-type replacement policy for a repairable load-sharing k-out-of-n system whose components state is hidden and detected only by inspections.

For other maintenance policies, Zhang et al. [42] proposed system and component level maintenance policies for a two-component load-sharing system. The system is assumed to be subject to imperfect preventive maintenance. Jamali and Pham [263] developed an opportunistic maintenance model for load-sharing k-out-of-n systems. The proposed maintenance model considers two intervals where minimal repairs are sufficient for the failed components in the first interval. In the second interval, corrective Maintenance of the failed components together with opportunistic preventive Maintenance of the survived ones are performed after m failures.

Some studies consider optimal design problems, for instance, Shekhar et al. [264] studied the optimal design of fault-tolerant machining system with various types of machining hindrance. They considered a K-out-of-M+Y+S: F redundant repairable machining system with the support of mixed standbys facing independent failure, switching failure, state-dependent failure, and common cause failure. The repair facility (rates) was considered as decision variables. Endharta and Ko [154] considered the design problem for a circular k-out-of-n: G balanced system with equal load-sharing units. Two maintenance polices were considered: corrective maintenance, which was only triggered by a failure event, and CBM, which was triggered when the number of failures reached a certain number or at system failure time, whichever occurred first.

2.3.2.3 Load-sharing modelling

In sections 2.3.2, 2.3.2.1 and 2.3.2.2 we provided an overview of research for multi-component systems with load-sharing dependence. In this section, we will discuss the models used for modelling load-sharing in the literature. Research on load-sharing systems can be placed under the following methods: markov-processes, covariate models (proportional hazards model (PHM) and accelerated failure time models (AFTM)), simulation and statistical distributions.

The first method for modelling load-sharing dependence was introduced by Freund et al. [265] in 1960s. Freund et al. [265] presented a bivariate exponential distribution to model a two component dynamic load-sharing system. The model considers a 1-out-of-2 system where components A and B are assumed to have individual constant failure rates λ_1 and λ_2. After one component has failed, the surviving component gets a modified (higher) failure rate ($\lambda_1^{'}$ or $\lambda_2^{'}$). The combination of the components' failure density functions to the system density yields a bivariate density function. However because the failure behaviour of one component depends on the failure behaviour of the other, it is then assumed that the marginal distribution of the bivariate distribution can not be exponential [266]. Whilst it is possible to expand the Freund model for systems with n components. It has some limitations. First the number of possible failure combinations m will increase tremendously ($m = n!$) if components sizes increases to n [22].In addition, it is limited to parallel systems [22]. Lin et al. [267] later extended the Freund's model to a three-component model and a special n-component model by using Markov model to derive the system reliability function.

Markov process is another method used for analysing load-sharing systems. Using markov model, the transition rate of a working component is modified to include the effect of an extra load. The markov model can be applied to systems with more than two components with individual load steps and for deriving the systems reliability function [22]. Shao and Lamberson [150] present a markov model to analyze the reliability and availability of a k-out-of-n load-sharing system with repairable components. Wang et al. [38] studied a load-sharing parallel system with failure dependency using a semi-markov process. They introduced a dependence function to quantify the failure dependency between components.

A widely applied approach for modelling load-sharing mechanism is a capacity flow model [266]. The model assumes that given the total system load L, total number of components n and an equal load-sharing rule, then each working component will carry a load of L/n at time t. If one component fails, the remaining components will obtain a higher load of $L/(n-1)$. The failure behaviour of the components is statistically dependent. The failure rate for each working component is given by

$$\lambda_i = \left(\frac{n}{n-x}\right)^{\gamma}\lambda_0, \tag{2.16}$$

where x = number of failed components and γ is the load factor. The load factor describes the influence of the increased load to the surviving components ($\gamma = 0$: no influence of the increased load). According to [266], the system reliability function can be obtained using a general erlang distribution with n steps.

Accelerated failure time model (AFTM) have been used to characterize the relationship between the failure rate of a component and the load imposed on the component [22]. It is assumed that the failure rate of a component increases if a higher load is imposed on the component. Hence, the failure rate of component is formulated as

$$\lambda(t; z) = \phi(z).\lambda_0(t.\phi(z)), \tag{2.17}$$

where $\lambda_0(.)$. the baseline failure rate. The function $\phi(.)$ is modelled by either the power law (i.e. $\phi(z) = z^{\gamma}$) or the exponential law (i.e. $\phi(z) = e^{z^{\gamma}}$). The model parameter can be obtained from accelerated life testing analysis [268].

Combining the concept of capacity flow model and AFTM models, Amari et al. [36] proposed a closed-form analytical solution to evaluate the reliability of load-sharing k-out-of-n: G systems with tampered failure rate (TFR). A tampered failure rate model assumes that the fluctuations in load does not affect the age of the remaining components, and only change their hazard rates [13]. Amari and Bergman [37], based on cumulative effect (CE) load-sharing model, presented a method to compute the reliability of k-out-of-n: G load-sharing systems with identical or non-identical components, both following general failure distributions. Cumulative effect model assumes that the load changes affect not just the failure rate but also the age of the working components [13].

Proportional hazard model (PHM) have also been used to characterize the relationship between the failure rate of a component and the load imposed on the component [22]. Similar to the Accelerated failure time model, it is assumed that the failure rate of a component increases if a higher load is imposed on the component. The failure rate of a component in the proportional hazard model is written as

$$\lambda(t; z) = \lambda_0(t) f(z, \beta), \tag{2.18}$$

where $z = z_1, z_2, \ldots, z_n$ is a set of loads imposed on the component, and $\lambda_0(t)$ is the baseline failure rate (also called hazard rate) function and β is the load factor. The term $f(z\beta)$ is the function that characterizes the relation between loads and the baseline failure rate.

Mohammad et al. [269] applied proportional hazards method to model the load-dependent time varying failure rate of a component in a load-sharing system. The effects of load on a component was assumed to be multiplicative in nature. As such, the failure rate of the component was modelled as a product of both a baseline hazard rate, which can be a function of time t, and a multiplicative factor which is function of the current load on the component. Zhang et al. [22] applied proportional hazards method to model the load-dependent time varying failure rate of a component in a multi-state load-sharing system.

Lastly, Pozsgai et al. [266] proposed the use of simulation techniques for modelling load-sharing dependence. According to Pozsgai et al. [266] the failure rates of components of a mechanical system are not constant due to aging and wear-out mechanisms as such a more adequate method for modelling load-sharing in the case of mechanical systems is the application of simulation techniques.

2.4 Gaps in the literature

In sections 2.3.1 and 2.3.2, we presented current research and modelling approaches for stochastic dependence effect in the form of failure interaction and load-sharing respectively. In section 2.2.3, we presented current research and modelling approaches for unobserved heterogeneity effect. Sections 2.4.1, 2.4.2 and 2.4.3 presents the gaps in the literature that will be addressed in chapter 3, chapter 4 and chapter 5 respectively. Furthermore, with regards to failure interaction and spatial dependence effect, only two studies have considered spatial dependence in the form of propagated failure on multi-state system [233, 251]. Existing studies have considered Type I stochastic dependence where the failure of a component can cause an immediate failure of nearby component. However, there is no other work yet that has considered Type II or Type III failure interaction together with spatial dependence for reliability assessment or maintenance optimization of a multi-component system. An example of such systems can be found in a water distribution system consisting of several pipe components in which the failure of a pipe can trigger failures in neighboring pipes due to sudden pressure changes in the system [270, 271].

2.4.1 Gap to be addressed in chapter 3: unobserved heterogeneity

In section 2.2.3, we reviewed literature on unobserved heterogeneity. For repairable systems subject to minimal repair, the following research gaps have been identified. First, most studies on unobserved heterogeneity have predominantly focused on investigating the significance of covariates and the frailty term in the fitted model rather than prediction for the system and/or individual components. The few works that have considered event prediction for point processes with unobserved heterogeneity includes Deep et al. [69] that used a semi-parametric Andersen and Gill model for failure prediction of a new component in a Teleservice system using collected data from old units; and Jahani et al. [70] who developed a multivariate Gaussian convolution process (MGCP) for fleet-based event prediction in which failure prediction for an individual unit is conducted using data collected from other units. The ability to predict the occurrence of failure events at an individual unit level can aid optimal maintenance decision making for individual components [69]. The first objective of chapter 3 is to develop a parametric method for predicting the occurrence of failure events at the component level considering a repairable system with minimal repair. To accomplish this objective, an empirical Bayes framework will be adopted to update the frailty term and the chapter will concentrate on the modeling of repairable systems using the power law NHPP. NHPP provides flexibility to model the failure time occurrences such as whether a system is improving or deteriorating [74]. In addition, power law model is flexible to describe various types of system's phases [99]. Lastly, NHPP can be extended to include random effect terms to reflect unit-to-unit variation using the Cox model framework [24, 67, 69].

Second, no other distribution has been investigated for modelling unobserved heterogeneity effect apart from gamma distribution because it is mathematically tractable. As an alternative to gamma, Hougaard [202] introduced the inverse Gaussian (IG) distribution for modeling unobserved effects. The IG distribution has a unimodal density and is a member of the exponential family. Its shape resembles that of other skewed density functions, such as the lognormal and gamma distributions. Chikara and Folks [272] studied the IG distribution and found that there are many striking similarities between the statistics derived from this distribution and those of the normal distribution. These properties make the IG potentially attractive for modeling the failure data [165]. The IG distribution has some advantages as a frailty distribution. It provides flexibility in modelling, when early occurrences of failures are dominant in a lifetime distribution and its failure rate is expected to be non-monotonic [165]. Also, IG is almost an increasing failure rate distribution when it is slightly skewed and hence is also applicable to describe lifetime distribution which is not dominated by early failures

[165]. Hougaard [202] noted that survival models with gamma and IG frailties behave very differently, specifically that the relative frailty distribution among survivors is independent of age for the gamma but becomes more homogeneous with time for the IG. The conclusion was derived by observing the coefficient of variation of the two frailty distributions of survivors. For the gamma distribution, the coefficient of variation is a constant. However, for the IG distribution, the coefficient of variation is a decreasing function of time. The likelihood function for the IG frailty model can be easily obtained and has closed-form representation, which indicates fast and simple estimation of parameters [165, 273]. In fact, a few studies in areas such as medicine and epidemiology have suggested the IG frailty model as an alternative to the gamma frailty model for modeling unobserved effects (for example, [172, 178, 274, 275, 276]). IG frailty model have also been considered for degradation analysis (see [277]). In the paper, IG process was extended to include gamma and IG distributed frailties. When compared on real data, they found that the two models presented similar results.

Despite these desirable properties, IG distribution has not been investigated for modelling unobserved heterogeneity in repairable systems failure data. Hence, another objective of chapter 3 is to evaluate the application of the NHPP model with IG distributed frailties for analysing failure data from repairable systems and to compare its results with the gamma distributed frailties. To achieve this objective, we shall consider a system in its burn-in phase. A system in its burn-in phase has increasing inter-failure times. Also, within a mission time, components in the wear-out phase (components with decreasing inter-failure times) are replaced with new components whereas components in the burn-in phase (relatively new components) are subject to minimal repair upon failure. As more components are replaced, the relative frailty among the components in the system become homogeneous over time.

We present one example to illustrate this situation - water distribution systems. Many pipes behave heterogeneously in terms of failures mainly due to the age of the pipes and the type of materials used in water delivery to the end user [20, 278]. Compared with other material types, asbestos cement pipes fail more frequently and is no longer allowed to be used in some countries (for example, Sweden [279], and Mexico

[280]). However, due to limited financial resources, asbestos-cement pipes are either refurbished by reinforcement or replaced by polyethylene pipes [279, 280]. As a result, more homogeneous failure behaviour might be observed in the pipes failure data over time.

Lastly, the impact of misspecification of the frailty distribution has not been investigated. In a model with unobserved heterogeneity, it is necessary to define the distribution of the unobserved effects [71, 72]. Since the modeled heterogeneity is unobservable, the appropriate choice of distribution of the unobserved effects is not easily discernible [73, 74]. Furthermore, the choice of the distribution of unobserved effects can give interesting general results in terms of the variance of the unobserved effects [74]. For instance, a large variance could indicate deficiencies in the choice of the distribution which may influence the model fit [73, 74]. It is therefore useful to examine the extent to which misspecification of the frailty distribution affects the validity of intensity function estimators [73]. The third objective of chapter 3 is to examine the impact of wrongly specifying the frailty distribution in a NHPP model. To accomplish these objectives, NHPP model with IG distributed frailties will be developed and compared with a gamma distributed frailty model through a simulation study and analysis of a real dataset. Statistical fit and prediction performance of the two models will be compared over different heterogeneity levels, sample sizes and component failure behaviours.

Although chapter 3 of this thesis is like Morita et al. [274] and Kheiri et al. [172] in the use of IG distributed frailties to model heterogeneity effects, it is different from the work done by Morita et al. [274] and Kheiri et al. [172] in three ways. First, the underlying process studied differs from the two studies. Chapter 3 of this thesis considers recurrent failures using Power law NHPP model, whereas Kheiri et al. [172] and Morita et al. [274] considered non-recurrent failures and degradation using piece constant hazard model and IG process respectively. Second, using simulation study, Morita et al. [274] studied the consequences of ignoring unobserved heterogeneity in the degradation rate of a systems by comparing two IG process models – one IG process model with IG or gamma distributed frailties and another ordinary IG process models.

Kheiri et al. [172], fit three models – univariate frailty model, IG shared frailty model and IG correlated frailty model to a bilateral corneal graft rejection dataset to identify the best model for that data. In contrast, this chapter studies the impact of misspecification of the frailty distribution by comparing two power law models - one power law model with gamma distributed frailties with another power law with IG distributed frailties. The study was conducted using both simulation study and an application to real data. Finally, this chapter considers prediction of a component's mean residual life and prediction of the expected number of failures at the component and system levels. In contrast, Morita et al. [274] focused on investigating the significance of the frailty term in the fitted model rather than prediction for the system and/or individual components. Kheiri et al. [172] focused on fitting three frailty models to a bilateral corneal graft rejection dataset to identify the best model for that data.

To summarise, we have established that most research focused on investigating the significance of covariates and the frailty term in the fitted model and a parametric method for system and component event prediction has not been developed. We also established that unobserved heterogeneity is commonly modelled by gamma distribution which assumes constant heterogeneity over time and systems whose components becomes homogeneous over time have not been studied. Lastly, we established that the impact of misspecification of the random effect distribution has not been investigated for systems with minimal repair assumption. The contribution of chapter 3 is threefold. First, methods for predicting the occurrence of failure events at system and component levels will be developed. Second, a model for systems whose components becomes homogeneous over time will be developed. Lastly, the impact of wrongly specifying the random effect distribution in a NHPP model will be investigated.

One of the significant questions that this thesis will explore in chapter 3 is what is the impact of wrongly assuming a random effect distribution for a repairable system subject to minimal repair.

2.4.2 Gap to be addressed in chapter 4: load-sharing and spatial dependence

In section 2.3.2, we reviewed literature on load-sharing systems. The following research gaps have been identified. First, no research has considered models for reliability prediction for load-sharing system with spatial dependence and proximity effect. Most of the recent studies have focused on the determination of reliability or availability function, statistical inference, and maintenance of systems with equal load-sharing. Very few are focused on local load-sharing, however, these studies focused on local load-sharing without considering proximity effect. In chapter 4, we will develop a novel model for reliability prediction of load-sharing systems with spatial dependence and proximity effect. In such systems, if a component fails, its neighbours will take up extra load if they are close enough. Each working component's performance depends on its spatial neighbour's state, proximity, and the spatial pattern among components. Unlike equal load-sharing structure, where the load of a failed component is taken up by all the remaining working components, in a system with spatial dependence, the load of a failed component is taken up either by its nearest neighbouring components (while non-neighbours operate at their normal rate) [52] or there may exist a proximity effect that measures the degree to which components in the system affect each other depending on how close they are [32].

We present two examples to illustrate this situation. First, example involves a warehousing system in which a bigger warehouse is surrounded by several smaller warehouses such that when the bigger warehouse fails, all the smaller warehouses will take the load depending on how close they are to the big warehouse. When a smaller warehouse fails, the load will be redistributed among the bigger warehouse and the smaller warehouses based on the distance between the warehouses [52]. Second, a water utility has a network of pipes that collectively serve the purpose of delivering water from a treatment site to houses and businesses. When a leak occurs, there will be a drop in pressure but when the maintenance action takes place, the network is shut off, and so an increase in pressure will be observed higher up the network as water is backed up. An individual pipe failure causes a drop in pressure but does not cause the system to fail. If sufficient failures occur, a sufficient drop in pressure results in failure and an increased load on neighbouring pipes. The distribution of that load is not uniform across the entire network. Instead, neighbouring pipes will pick up that additional pressure, and hence have an increased failure rate.

Second, the importance and significance of the spatial effect in a load-sharing system has not been studied. None of the studies on local load-sharing have not considered studying the consequence of ignoring spatial effect if it exists. The contribution of chapter 4 is twofold. First a generic model for estimating the reliability of a load-sharing system with spatial dependent components will be developed. Second, using the developed model, the consequence of ignoring spatial effect if it exists will be studied. An extension of the capacity flow model described in section 2.3.2.1 will be used to characterize the relationship between the failure rate of a component and the load imposed on the component. A Markov model will be used to characterize the deterioration process of the entire system. Markov model evaluates the probability of jumping from one known state into the next logical state. The process continues until the system being considered has reached the final failed state or until a particular mission time is achieved [145]. Markov model have been found to be useful and suitable for modelling the distribution of time to failure of load-sharing systems in various research [47, 49, 52, 53, 54, 281].

To summarise, we established in this section that no research has considered models for reliability prediction for load-sharing system with spatial dependence and proximity effect. We also established that the works that focused on local load-sharing failed to consider proximity effect and the consequence of ignoring spatial effect if it exists. The contribution of chapter 4 is twofold. First a generic model for estimating the reliability of a load-sharing system with spatial dependent components will be developed. Second, using the developed model, the importance and significance of the spatial effect will be studied.

Another significant question that this thesis will explore in chapter 4 is what is the impact of ignoring spatial dependence if it exists in a load-sharing system.

2.4.3 Gap to be addressed in chapter 5: load-sharing, spatial dependence and shocks

In section 2.3.2.1 and 2.3.2.2, we reviewed literature on reliability and maintenance optimization of load-sharing systems with external shocks. The following research gaps have been identified. First, that no research has considered models for reliability prediction for load-sharing system with spatial dependence, proximity effect and external shocks. Most of the recent studies on load-sharing systems with shocks have focused on equal load-sharing systems with external shocks. Only one paper (Guo et al. [50]) focused on reliability assessment with respect to local load-sharing and shocks. Thus, chapter 5 of this thesis will develop a novel model for reliability prediction of load-sharing system with spatial dependence, proximity effect and shocks. In the system, if a component fails, its neighbours will take up extra workload if they are close enough. Each working component's performance depends on its spatial neighbour's state, proximity, and the spatial pattern among components. The components in the system may not only bear the workload but also experience the external load, such as temporary shocks. The failure rate of a component is affected by the extra workload taken from failed neighbours and random shock processes.

We present an example to illustrate this situation. One example is a cable structure system in suspension bridge where component failures usually only affect proximate components instead of all working components. When one boom in the system fails, all the working proximate booms are subjected to direct force transmission and they suffer the most power of the force, so the proximate booms take an increased force. However, the booms far from the failed one are only subjected to the indirect force and the transfer effect on them less compared with those on the proximate booms. As a result, the load added to the booms far from the failed one is very small. The working booms constantly suffer from the continuous shock processes of random winds, which cause abrupt wear debris of the booms.

Second, preventive maintenance optimization of load-sharing system with spatial dependence, proximity effect and external shocks has not been studied so far. Most of the research on maintenance optimization of load-sharing systems have focused on equal load-sharing. No study was found that focused on maintenance optimization with respect to local load-sharing and load-sharing system with spatial dependence, proximity effect and external shocks.

Lastly, the impact of ignoring spatial dependence effect if it exists in a load-sharing system subject to spatial dependence and external shocks have not been considered and its implication for maintenance decisions have not been investigated. Thus, the contribution of this chapter is twofold. First, a generic model for estimating the reliability of a load-sharing system with spatial dependent components and random shock processes will be studied. Second, preventive maintenance optimization of load-sharing system with spatial dependence, proximity effect and external shocks will be developed. Lastly, using the developed model, we will investigate the impact of ignoring spatial dependence effect if it exists in a load-sharing system subject to external shocks.

To summarise this section, we have established that no research has considered reliability prediction for load-sharing system with spatial dependence, proximity effect and external shocks. We also established that preventive maintenance optimization of load-sharing system with spatial dependence, proximity effect and external shocks have not been developed. Lastly, we established that the impact of ignoring spatial dependence if it exists in a load-sharing system subject to spatial dependence and external shocks has not been investigated. The contribution of this chapter is threefold. First, a generic model for estimating the reliability of a load-sharing system with spatial dependent components and random shock processes will be developed. Second, preventive maintenance of load-sharing system with spatial dependence, proximity effect and external shocks will be studied. Lastly, using the developed model, we will investigate the impact of ignoring spatial dependence effect if it exists in a load-sharing system subject to external shocks.

The third significant question that this thesis will explore in chapter 5 is what is the impact of ignoring spatial dependence if it exists in a load-sharing system with external shocks.

Chapter 3

Reliability evaluation of a repairable multi-component system considering unit heterogeneity using frailty model

In chapter 2, we reviewed the literature on multi-component systems with unobserved heterogeneity, stochastic dependence and spatial dependence. In section 2.4.1 of chapter 2, we established the following: First, for systems with minimal repair assumption, most research focused on investigating the significance of covariates and the frailty term in the fitted model and a parametric method for system and component event prediction has not been developed. Second, unobserved heterogeneity is commonly modelled by gamma distribution which assumes constant heterogeneity over time and systems whose components becomes homogeneous over time have not been studied. Lastly, the impact of misspecification of the random effect distribution has not been investigated for systems with minimal repair assumption.

Based on the established gaps in section 2.4.1 of chapter 2, our contribution in this chapter is threefold. First, we will develop an IG frailty model and the parameter estimators for repairable systems whose components become homogeneous over time.

Using a simulation study, we will examine the performance of the estimators. Second, the impact of wrongly assuming the random effect distribution for a repairable system subject to minimal repair will be examined by comparing the robustness of the IG and gamma frailty models on simulated data and a classic dataset. We will develop the two frailty models as an extension of the NHPP with power law (PL) intensity function. Finally, using empirical Bayes framework, we will develop a method for prediction of a component's mean residual life and prediction of the expected number of failures at the component and system levels.

The remainder of this chapter is outlined as follows: Section 3.1 describes a general system and its failure process. Section 3.2 presents the IG frailty model and develops the procedure of the maximum likelihood estimation. Section 3.3 presents the design and results of examining the performance of the IG model estimators and the robustness of the IG and gamma frailty models in a simulation study. The method for event prediction and application to a classic dataset is presented in section 3.4. Section 3.5 presents the conclusion.

3.1 System description

Consider a system subject to minimal repairs upon failure. When a minimal repair is performed upon failure, the times between subsequent failures may not be identically distributed, which constitutes an NHPP [77, 86].

A system in its burn-in phase has increasing inter-failure times. Also, within a mission time, components in the wear-out phase (components with decreasing inter-failure times) are replaced with new components whereas components in the burn-in phase (relatively new components) are subject to minimal repair upon failure. As more components are replaced, the relative frailty among the components in the system become homogeneous over time.

A common method for modelling the intensity function of a NHPP is the PL model (see section 2.1.2.1). In this chapter, the PL model will be used to describe the intensity function of the system. The PL model is flexible to describe various types of system phases [20, 99]. Furthermore, PL model is effective in representing a system which is experiencing reliability improvement (i.e., inter-failure times are increasing) [282]. The PL model is given as:

$$\psi_0(t) = \omega \rho t^{\rho - 1}, \tag{3.1}$$

where $\psi_0(t)$ is the intensity function at time t, ω is the scale parameter and ρ is the shape parameter that controls the shape of the curve. The parameter ρ in the power law model gives the following information about the system: if $\rho > 1$, then the system is deteriorating; if $0 < \rho < 1$, then the system is improving and if $\rho = 1$ the NHPP model reduces to an HPP.

Consider that a system has m independent repairable components. Each component is observed from the start of operation to time τ. Let n_j be the recorded number of failures for the j^{th} component and let t_{ij} be the age at the i^{th} occurrence of failure for component j, where $i = 1, 2, \ldots, n_j$, and $j = 1, 2, \ldots, m$. We assume that the number of failures vary across the components and there is no available covariate recorded for each component (however this information could be easily incorporated to the modelling). A positive random variable z_j, drawn from a distribution $f(z_j; \theta)$, is included in the NHPP model to account for component unobserved effects, where θ is the variance parameter. z_j acts multiplicatively on the intensity function [62], making the model a NHPP with random effect (also referred to as a frailty model [66]). The variance parameter θ is the parameter of interest in which small values of θ reflects homogeneity in the failure pattern of a group of components and large values of θ reflects high heterogeneity in the failure pattern of a group of the components. To make the baseline intensity function identifiable, a restriction is placed on $f(z_j, \theta)$ such that the frailties (random effects) are assumed to have an expected value $E(z_j) = 1$ and $Var(z_j) = \theta$. Thus components with $z_j > 1$ will fail more often than components with $z_j < 1$ [163].

Conditional on the frailty term z_j, the intensity function for component j can be expressed as:

$$\psi(t \mid z_j) = \omega_j \rho t_{ij}^{\rho - 1} = z_j \lambda \rho t_{ij}^{\rho - 1}, \tag{3.2}$$

where $\omega_j = z_j \lambda$ is the scale parameter and ρ is the shape parameter. We assume that components in system have the same ageing behaviour (i.e., same shape parameter ρ)

but different magnitudes (i.e scale parameter). As a result, the scale parameter ω_j will be made to be component specific by making λ fixed then we will introduce a random effect z_j which is modelled by a parametric distribution. We will assume that λ has a value of one to reduce the effect λ might have on each z_j.

3.2 IG frailty model

This section presents an IG frailty model based on the NHPP with IG distributed random effects. IG distribution has a simple Laplace transform which is useful for deriving the reliability function [283, 284]. In addition, gamma and IG distributions both have uni-modal density functions [284]. IG distribution is described by two characteristics, namely, a mean parameter $\mu > 0$ and precision parameter $\delta > 0$. The two-parameter IG distribution is given as:

$$f(z) = \sqrt{\frac{\delta}{2\pi z^3}} exp^{\left(\frac{-\delta(z-\mu)^2}{2z\mu^2}\right)}. \tag{3.3}$$

As mentioned above, the frailty model poses restrictions on the mean and variance on the distribution. Let $E(z) = 1 = \mu$ and $Var(z) = \theta = \frac{\mu^3}{\delta}$. The one-parameter IG distribution is:

$$f(z) = \frac{1}{\sqrt{2\pi\theta}} z^{\frac{-3}{2}} exp\left(\frac{-(z-1)^2}{2z\theta}\right). \tag{3.4}$$

3.2.1 Maximum likelihood for IG frailty model

Consider that the random effect z_j of the conditional intensity Eq.(3.2) is drawn from an IG distribution Eq.(3.4), then the conditional likelihood function for data from component j with random effect z_j is given below as:

$$L_j(\lambda_0(t_{ij}) \mid z_j) = \left(\prod_{i=1}^{n_j} z_j \lambda_0(t_{ij})\right) exp^{-z_j \Lambda_0(\tau)}, \tag{3.5}$$

where $\lambda_0(t_{ij}) = \lambda \rho t_{ij}^{\rho-1}$ and $\Lambda_0(\tau) = \lambda \tau^\rho$. τ is the observation length of component j. Here, j is a fixed identification number for a component. t_{ij} represents each failure time t_i observed over the entire life span of component j. z_j is a fixed value generated for component j and it influences each failure time t_i of component j for the rest of its life.

Since z_j is unobservable from the data, the contribution of the j^{th} component to the full likelihood is obtained when the marginal likelihood of the j^{th} component is derived. The marginal likelihood of the j^{th} component is derived after integrating out the random effects from the conditional likelihood function. The conditional likelihood function of the j^{th} component is given as:

$$L_j(\theta \mid \lambda_0(t_{ij})) = \left(\prod_{i=1}^{n_j} \lambda_0(t_{ij})\right) \int_0^\infty z_j^{n_j} exp^{(-z_j \Lambda_0(\tau))} f(z_j;\theta) dz_j, \tag{3.6}$$

where $f(z_j;\theta)$ is the IG density.

$$L_j(\theta \mid \lambda_0(t_{ij})) = \left(\prod_{i=1}^{n_j} \lambda_0(t_{ij})\right) \int_0^\infty \frac{1}{\sqrt{2\pi\theta}} z_j^{n_j-\frac{3}{2}} exp^{\left(-z_j\Lambda_0(\tau)-\frac{(z_j-1)^2}{2z_j\theta}\right)} dz_j. \tag{3.7}$$

The marginal likelihood of the j^{th} component is given as:

$$\begin{aligned} L_{j(marg)} &= \frac{\left(\prod_{i=1}^{n_j} \lambda_0(t_{ij})\right) 2exp^{\frac{1}{\theta}}}{\sqrt{2\pi\theta}} (2\theta\Lambda_0(\tau)+1)^{-\frac{n_j-1/2}{2}} \dots \\ &\times k_{(n_j-1/2)}\left[\frac{\sqrt{(1+2\theta\Lambda_0(\tau))}}{\theta}\right], \end{aligned} \tag{3.8}$$

where $k_{(n_j-1/2)}[.]$ is a modified bessel function of the second kind and $\int_0^\infty x^{s-1}exp^{(ax^h-bx^{-h})}dx = \frac{2}{h}(\frac{b}{a})^{(\frac{s}{2h})}k_{(\frac{s}{h})}[2\sqrt{ab}]$ (see Shoukri et al. [285]).

The total likelihood function for all the components is given by:

$$
\begin{aligned}
L &= \prod_{j=1}^{m} L_{j(marg)} \\
&= \prod_{j=1}^{m} \frac{\left(\prod_{i=1}^{n_j} \lambda_0(t_{ij})\right) 2exp^{\frac{1}{\theta}}}{\sqrt{2\pi\theta}} \left(2\theta\Lambda_0(\tau)+1\right)^{-\frac{n_j-1/2}{2}} \dots \\
&\times k_{(n_j-1/2)} \left[\frac{\sqrt{(1+2\theta\Lambda_0(\tau))}}{\theta}\right].
\end{aligned}
\tag{3.9}
$$

Taking the logarithm of the likelihood function of Eq.(3.9) results in:

$$
\begin{aligned}
l = &\sum_{j=1}^{m}\left(\sum_{i=1}^{n_j} \log(\lambda_0(t_{ij}))\right) + \sum_{j=1}^{m} log\left(2exp^{\frac{1}{\theta}} \frac{1}{\sqrt{2\pi\theta}}(1 \dots\right. \\
&\left. + 2\theta\Lambda_0(\tau))^{-\frac{n_j-1/2}{2}} k_{\left(\frac{n_j-1/2}{2}\right)} \left[\frac{\sqrt{(1+2\theta\Lambda_0(\tau))}}{\theta}\right]\right.,
\end{aligned}
\tag{3.10}
$$

$$
\begin{aligned}
l = &\, n\log\lambda + n\log\rho + (\rho-1)\sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij} + mlog(2)\dots \\
&+ \frac{m}{\theta} - \frac{m}{2}log(2\pi\theta) - \sum_{j=1}^{m}\frac{(n_j-1/2)}{2}log(1+2\theta\lambda\tau^{\rho})\dots \\
&+ \sum_{j=1}^{m} logk_{\frac{n_j-1/2}{2}}\left[\frac{(1+2\theta\lambda\tau^{\rho})}{\theta}\right],
\end{aligned}
\tag{3.11}
$$

where $k_{\left(\frac{n_j-1/2}{2}\right)}[.]$ is a modified bessel function of the second kind. $n = \sum_{j=1}^{m} n_j$. Taking the derivative of the log likelihood function with respect to λ, ρ and θ:

$$
\frac{\delta l}{\delta\lambda} = \frac{n}{\lambda} - \frac{\sum_{j=1}^{m}\tau^{\rho}k_{\frac{n_j+1/2}{2}}\left[\frac{(\sqrt{r_o})}{\theta}\right]}{\sqrt{r_o}k_{\frac{n_j-1/2}{2}}\left[\frac{(\sqrt{r_o})}{\theta}\right]},
\tag{3.12}
$$

where $r_o = 2\theta\lambda\tau^\rho + 1$.

$$\frac{\delta l}{\delta \rho} = \frac{n}{\rho} + \sum_{j=1}^{m}\sum_{i=1}^{n_j} \log t_{ij} - m\lambda\tau^\rho \log\tau. \tag{3.13}$$

$$\begin{aligned}\frac{\delta l}{\delta \theta} = & -\frac{n}{\theta^2} - \frac{n}{2\theta} + \sum_{j=1}^{m} \frac{n_j - 1/2}{\theta} + \sum_{j=1}^{m} \frac{\sqrt{r_o}k_{\frac{n_j+1/2}{2}}\left[\frac{(\sqrt{r_o})}{\theta}\right]}{\theta^2 k_{\frac{n_j-1/2}{2}}\left[\frac{(\sqrt{r_o})}{\theta}\right]} \cdots \\ & - \sum_{j=1}^{m} \frac{\lambda\tau^\rho k_{\frac{n_j+1/2}{2}}\left[\frac{(\sqrt{r_o})}{\theta}\right]}{\theta\sqrt{r_o}k_{\frac{n_j-1/2}{2}}\left[\frac{(\sqrt{r_o})}{\theta}\right]}.\end{aligned} \tag{3.14}$$

Estimates of λ, ρ and θ can be obtained by setting the derivatives to zero. The expression for the estimators of λ, ρ, and the heterogeneity parameter θ cannot be derived analytically. One way to deal with this problem is to use a numerical method such as Newton-Raphson's method which run a gradient descent algorithm to optimise the likelihood function [62]. Gradient descent approaches assume that the likelihood function to be maximised is smooth and concave. The limit of the algorithm is that it may provide a sub-optimal (local optimal) solution if the objective function is not concave. For illustrative purpose, we present the maximum likelihood estimation of gamma frailty model in the appendix A.

3.3 Simulation Study

In this section, we conduct a simulation study. This section is structured as follows. In Section 3.3.1 we describe the simulation design and data simulated. In Section 3.3.2 we assess the performance of the parameter estimators of the IG frailty model. The performance of the IG estimators will be examined with respect to bias in the estimates of the scale λ, shape ρ, and heterogeneity parameter θ given some known input parameter values. In Section 3.3.3 we assess the robustness of the IG and gamma frailty model in a misspecification study.

3.3.1 Simulation design

Throughout the simulation study, we consider that the underlying failure process for the components in a system follows a NHPP with basic rate of occurrence of failures of a PL process $\lambda\rho t_{ij}^{\rho-1}$ and is conditional on a frailty term z_j whose distribution $f(z_j, \theta)$ is known. When the IG frailty model estimators are examined, the frailty term z_j is IG distributed. However, when the effectiveness of the IG and gamma frailty models are examined, the frailty term z_j is either gamma or IG distributed. For each process, input values of λ are fixed to 1 while ρ and θ are varied. We use 19 sample sizes, $n = 10$ to 100 with 5 step size increments, to assess the impact of increasing the sample size on parameter estimates. Sample size in this chapter refers to the number of components in the system. The mean of 1000 parameter estimates are assessed for ρ and θ. We assume equal observation length of $\tau = 0$ to τ=10 for each system. Whilst it is, in principle, straight forward to generalize the likelihood function to the case of different observation lengths, doing so complicates the computation of estimates [62]. Thus, in order to simplify the estimation, we decided to keep the observation lengths equal.

To perform the simulation study, we generate data. The following algorithm is used to generate failure time data from each frailty model. If we assume that failure times of a component are independent given the frailty term and that the sequence of failure times forms a counting process, we can then generate samples of an NHPP from an HPP. Let us consider an HPP with intensity function $\lambda(t)$, an interval $(t_{(i-1)}, t_i)$ for component j, and mean function:

$$\Lambda(t_{(i-1)}, t_i) = \Lambda(t_i) - \Lambda(t_{(i-1)}),$$

where the random variables $\Lambda(t_{(i-1)}, t_i)$ are independent and identically exponential distributed with mean 1. Let us suppose that there is no failure at installation time ($t = 0$), then by inverse probability method we can express the reliability function as:

$$exp^{\left(-\Lambda(t_{(i-1)})\right)} = U_{(i-1)},$$

and

$$\Lambda(t_{(i-1)}) = -\log(U_{(i-1)}),$$

where $U_{(i-1)}$ is drawn from a uniform distribution with an interval between 0 and 1. Suppose there are covariates then the mean function conditional on the frailty term becomes:

$$\Lambda(t_{(i-1)}) = \Lambda_0(t_{(i-1)}) z_j exp(\beta \mathbf{X}),$$

where β is a vector of regression coefficients. The covariates are depicted by X such that when $X = x(t)$ the covariates are time varying and $X = x$ implies time invariant covariates. If we assume that the intensity function is based on the PL model and each z_j is drawn from either a gamma or an IG distribution then we have:

$$\Lambda(t_{(i-1)}) = \lambda {t_{(i-1)}}^{\rho} z_j exp(\beta \mathbf{X}). \tag{3.15}$$

Solving for a failure time $t_{(i-1)}$ from Eq.(3.15) leads to the expression given by:

$$t_{(i-1)} = \left[\frac{\Lambda(t_{(i-1)})}{z_j \lambda exp(\beta \mathbf{X})} \right]^{1/\rho},$$

and

$$t_{(i-1)} = \left[\frac{-\log(U_{(i-1)})}{z_j \lambda exp(\beta \mathbf{X})} \right]^{1/\rho}. \tag{3.16}$$

The NHPP process of component j observed up to the i^{th} failure time t_i can be generated by the following recurrence formula:

$$t_i = \left[\frac{-\log(U_i)}{z_j \lambda exp(\beta \mathbf{X})} + {t_{(i-1)}}^{\rho} \right]^{1/\rho}. \tag{3.17}$$

If we assume that there are no available covariates i.e., $X = 0$, Eq.(3.16) becomes:

$$t_{(i-1)} = \left[\frac{-\log(U_{(i-1)})}{z_j \lambda} \right]^{1/\rho}, \tag{3.18}$$

and the recurrence formula becomes:

$$t_i = \left[\frac{-\log(U_i)}{z_j \lambda} + (t_{(i-1)})^\rho \right]^{1/\rho}. \tag{3.19}$$

To illustrate how to generate data when covariates are unavailable, we will simulate a single NHPP process for component j by the recurrence formula in Eq.(3.19) using the algorithm below. For clarity, we present the simulation flowchart in Figure 3.1.

1. Set parameter values for λ, ρ and θ
2. Set a value for τ
3. Draw a random value z_j from the known frailty distribution $f(z_j, \theta)$
4. Set $t_0 = 0$
5. Set $i = 1$
6. Draw U_i from a uniform distribution $\sim U(0, 1)$
7. Generate the i^{th} failure time by

$$t_i = \left[\frac{-\log(U_i)}{z_j \lambda} + (t_{(i-1)})^\rho \right]^{1/\rho}$$

8. If $t_i > \tau$ stop simulation otherwise $i = i + 1$
9. Repeat steps 6 to 8 until $t_i > \tau$.

3.3.2 Evaluation of IG Estimator

To assess the performance of the IG frailty model's estimator developed in Section 3.2.1, the Bias % of the mean from each estimated parameter will be determined. Given the input value of an arbitrary parameter p with its estimator $\hat{p}$. The Bias of the estimates after 1000 simulations will be calculated using

$$Bias(p) = \frac{\sum_{n=1}^{1000} \hat{p_n}}{1000} - p. \tag{3.20}$$

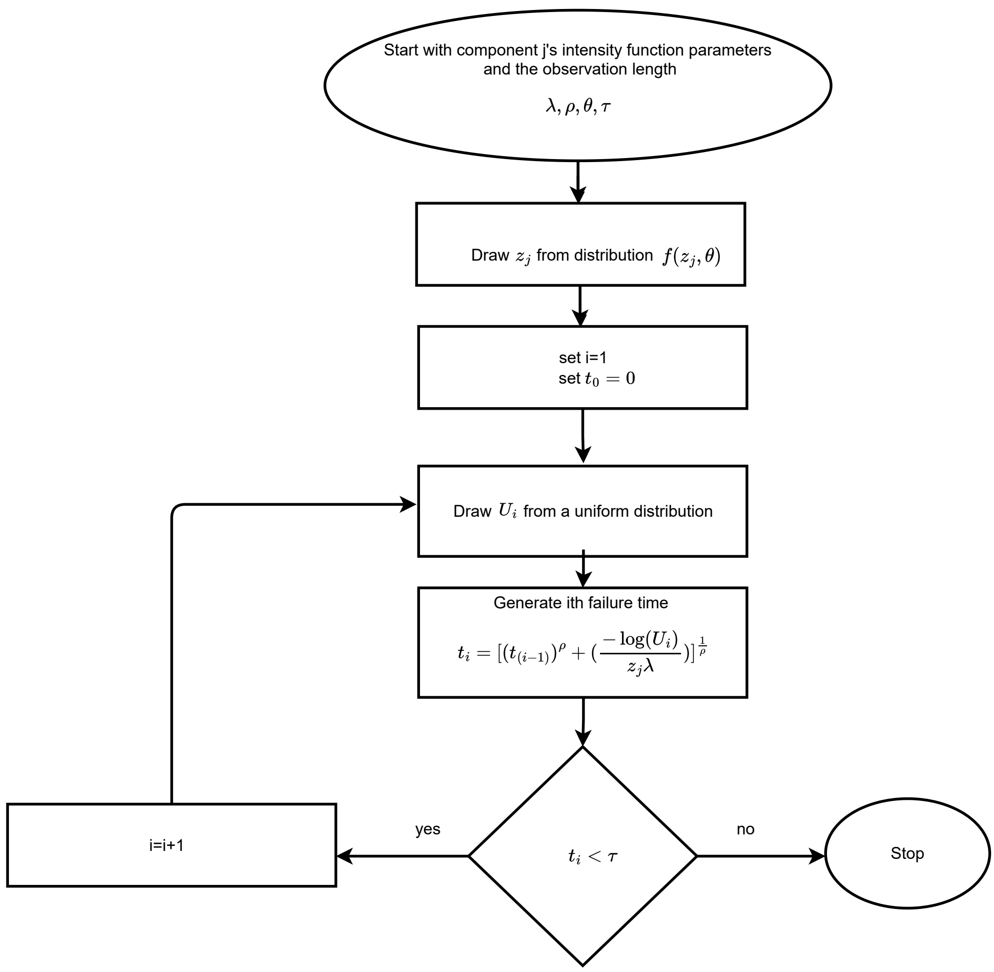
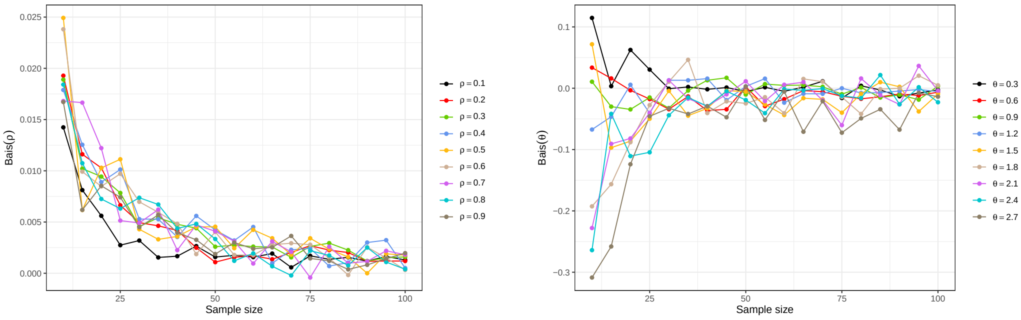


Figure 3.1: Simulation flowchart.



(a) Bias on ρ input values (b) Bias on θ input values

Figure 3.2: Plot of the Bias on estimates from IG frailty model estimators.

We assess the IG estimators developed in section 3.2.1 by observing the $Bias(p)$ from the mean of 1000 estimates of ρ and θ. The results are presented in Figure 3.2. The input values for λ, ρ, and θ are 1, $(0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$ and $(0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.4, 2.7)$ respectively.

From Figure 3.2a we observe that the bias in the estimated ρ values was 2.5 % for sample size 10 however the bias in the estimated ρ values reduced as the sample size increased. From Figure 3.2b, the bias in the estimated θ values ranged between -0.3 and 0.1 when the sample size is set to 10 but the estimated values rapidly improve and converge to the input value as the sample size is increased. We see that the estimator is robust to different choices of ρ and θ, and beyond a sample size of 10, the estimator performs well.

3.3.3 Misspecification study of gamma and IG frailty models

In this section, we assess the performance of the gamma and IG estimators to misspecification of the frailty distribution. When one of the models is known as the correct frailty model which the data is generated from, the other frailty model will be wrongly specified. The performance of the wrongly specified model will be assessed in terms of the model's goodness of fit to generated data and prediction accuracy when predicting the expected number of component failures. Since model selection using only maximum likelihood could be misleading due to variation in data, the sample sizes will be chosen to be sufficiently large in order to reduce the probability of selecting the wrong model

[286].

Analysis of the robustness of each wrongly specified frailty model (gamma or IG) is conducted by an assessment of the proportion of selections of each wrongly specified model using Akaike information criterion (AIC). AIC is used for goodness-off fit (GOF) test because AIC is one of the commonly used information-based criteria for model selection [287] and the AIC has been widely used either on its own or together with other tests (for example, [288] and [60]). AIC is similar to the Bayesian information criterion (BIC) in the sense that they both compare maximum likelihood values to select the appropriate model. Compared to AIC, BIC penalizes model complexity more heavily meaning that more complex models will perform worse and will, in turn, be less likely to be selected [289]. In this paper, the Gamma and IG frailty models considered have the same number of parameters and similar level of complexity. Thus, model selections from BIC will not differ from AIC because the penalty term in the BIC will have the same effect on the two model's values. AIC has been widely used to compare frailty models in the literature some examples include [290, 291]. The expression of AIC is given by [287]:

$$AIC = -2log(L) + 2p, \tag{3.21}$$

where L is the maximized likelihood value and p is the number of parameters in the model. When two models are compared, AIC considers the model with the smallest AIC value to have a better fit to the data. Further information on the AIC can be found in [287]. We will generate 1000 simulated data for fixed input values of λ, ρ and θ and then we will count the number of times each of fitted model is chosen by the AIC. Out of the 1000 simulated data, we will note the number of times the wrongly specified model is selected as the better model.

Furthermore, we examine the robustness of each wrongly specified frailty model in terms of their prediction of the expected number of component failures in the simulated data. The expected number of component failures in the system in a specific interval is given as:

$$E[N_S(t_1,t_2)] = \sum_{j=1}^{m} E[N_j(t_1,t_2)], \tag{3.22}$$

where $E[N_j(t_1,t_2)]$ is the expected number of the failures of the j^{th} component between times t_1 and t_2. $E[N_j(t_1,t_2)] = \frac{1}{\theta}log(\theta\Lambda_0(t_1,t_2)+1)^{-\frac{1}{\theta}}$ and $E[N_j(t_1,t_2)] = \frac{1}{\theta}(\sqrt{1+2\theta\Lambda_0(t_1,t_2)}-1)$ for gamma and IG frailty models respectively. $E[N_j(t_1,t_2)]$ is derived from the marginal reliability function for each component. We present the derivation of $E[N_j(t_1,t_2)]$ and the marginal reliability function in the appendix B and C.

We compare the predicted results with those from the correct model using root mean squared error (RMSE). We then assess the proportion of the wrong model selected as the better model by the RMSE. The expression of RMSE is given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n_I}(f_i - o_i)^2}{n_I}}, \tag{3.23}$$

where $f_i = E[N_S(t_i,t_{i+1})]$ is the forecasts (expected number of component failures between time (t_i,t_{i+1}) and o_i depicts the observed number of component failures between time (t_i,t_{i+1}) from the simulated data. n_I is the number of time intervals of equal length that the observed data is grouped into and $t_{n_I} = \tau$. To compute RMSE, the difference between the forecasts and observed number of component failures are squared and averaged over the number of time intervals to get the mean squared error (MSE); RMSE is the square root of MSE [292]. The model with the smallest RMSE value has the best prediction accuracy.

The performance of each specified model is assessed in four cases involving fixed ρ, and θ values.

- Case one involves setting θ to 0.3 to reflect low heterogeneity between components and modifying ρ.
- Case two involves setting θ to 3 to reflect high heterogeneity between components and modifying ρ.
- Case three involves setting ρ to 0.3 to reflect components' early life behaviour

with frequent failures and modifying θ.

- Case four involves setting ρ to 0.9 to reflect components with similar behaviour as those in the mid-life phase and modifying θ.

In each case, when one of the two parameters is fixed, the other is varied. For fixed settings of ρ, the θ values analysed are $(0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.4, 2.7, 3)$. For fixed settings of θ, the ρ values analysed are $(0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$. The result for each case is presented in Figures 3.3, 3.4, 3.5 and 3.6.

3.3.3.1 Case One - low heterogeneity

In case one, we set $\theta = 0.3$ to reflect low heterogeneity between components. The results are presented in Figure 3.3. For low heterogeneity, the likelihood of mis-specifying the IG or the gamma model is zero for ρ values from 0.4 to 0.9. In contrast, when $\rho = 0.3$ the proportion of assumed model selection was 1.5 % for IG and 1 % for gamma for sample size less than 25. The probability of mis-specifying the models was zero for sample sizes greater than 25. Thus, for data with low heterogeneity when one wrongly specify either the IG or gamma frailty model, either in terms of model fit or prediction, the chances of the wrong model being selected is very small.

3.3.3.2 Case Two - high heterogeneity

For case two, we set θ to 3 to reflect high heterogeneity between components. The results are presented in Figure 3.4. From Figure 3.4a and 3.4b we see that in terms of model fit, the proportion of the wrong model selected when the sample size is less than 50 ranged between 3% to 11% and 1% to 10% for IG and gamma model respectively. However, as the sample size is increased, the proportion reduces to zero for ρ values from 0.3 to 0.8. For $\rho = 0.9$ the proportion of wrong models selected can be seen to slowly decline. Because $\rho = 0.9$ mimics the behaviour of components in the mid-life stages, failure observations in the data are few. The few failure occurrence together with high heterogeneity between components increases the possibility of selecting the wrong model. In contrast, the proportion of the wrong model selected can be seen to reduce as sample size increased.

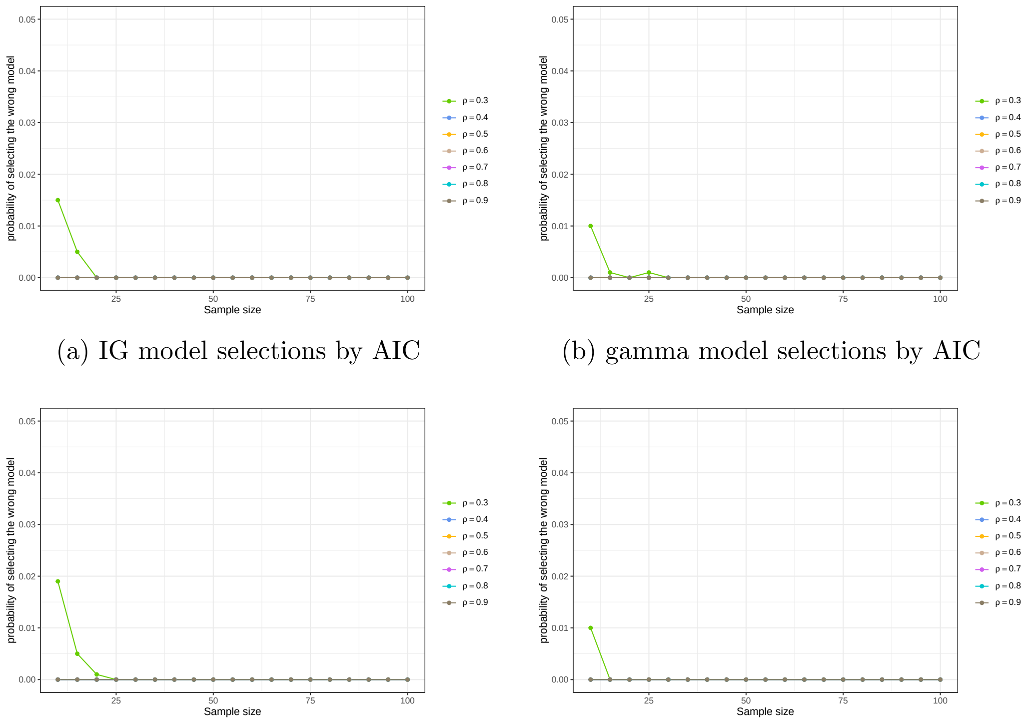
The results for determining the appropriate model when using prediction as the measurement are presented in Figure 3.4c and 3.4d. We observe a similar reduction in incorrectly selected models when gamma is wrongly specified. When IG is wrongly specified, the proportion of IG model selection is zero for ρ values from 0.4 to 0.9. In contrast, when $\rho = 0.3$ the proportion of IG model selection was up to 6% for sample size less than 50 and zero otherwise. The 6% selection of the IG model may be due to the sensitivity of the RMSE to outliers in the predicted number of component failures. For data with high heterogeneity, the probability of selecting the wrong model whether for model fit or for prediction is low and only happens when the sample size is small.

3.3.3.3 Case Three - component early life behaviour

In case three, we set ρ to 0.3 to reflect components' early life behaviour in which failures occur due to component defects or installation issues. The results for case three are presented in Figure 3.5. From the four plots 3.5a, 3.5b, 3.5c and 3.5d we see that in terms of prediction or model fit, the proportion of wrong model selections (whether IG or gamma) could be as high as 10% when the sample size is less than 50 and heterogeneity is high, i.e. close to 3. However, as the sample size is increased from 50, the likelihood of selecting the wrong model reduces to zero. More broadly, we see that the proportion of wrong models selected increases as θ increases from 0.3 to 3. This supports our findings when we compare Sections 3.3.3.1 and 3.3.3.2.

3.3.3.4 Case Four - component mid-life behaviour

In case four, we set ρ to 0.9 to reflect components nearing the mid-life stages where time between component failures are almost constant. The results are presented in Figure 3.6. From Figure 3.6a and 3.6b we see that in terms of model fit, the proportion of selections of the wrong model increased as heterogeneity levels increased for IG and gamma model reaching 12% and 10% respectively. In contrast, as the sample size is increased from 10 to 100 the proportion of wrong model selections reduced for all θ values from 0.3 to 3.



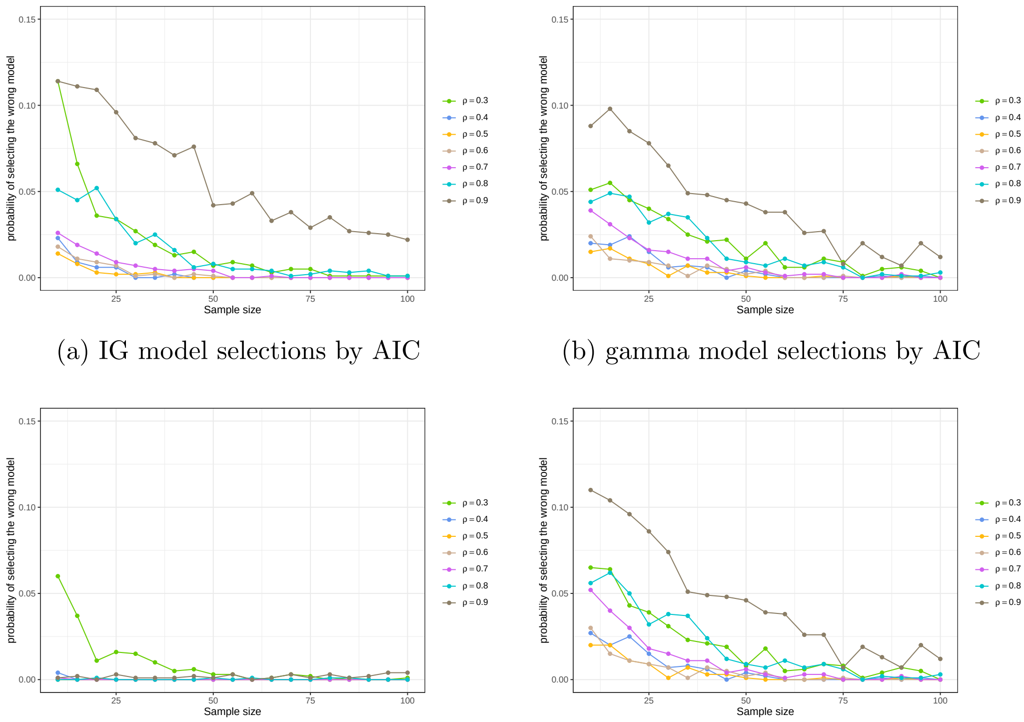
(a) IG model selections by AIC

(b) gamma model selections by AIC

(c) IG model selections by RMSE

(d) gamma model selections by RMSE

Figure 3.3: Plot of probability of selecting the wrong model when heterogeneity is low $\theta = 0.3$.



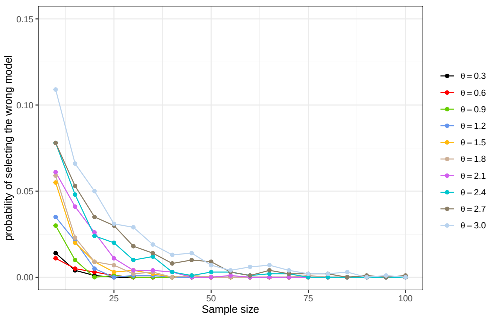
(a) IG model selections by AIC

(b) gamma model selections by AIC

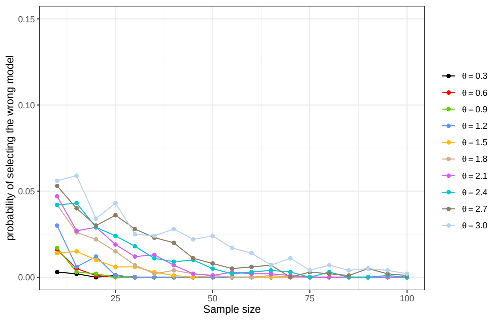
(c) IG model selections by RMSE

(d) gamma model selections by RMSE

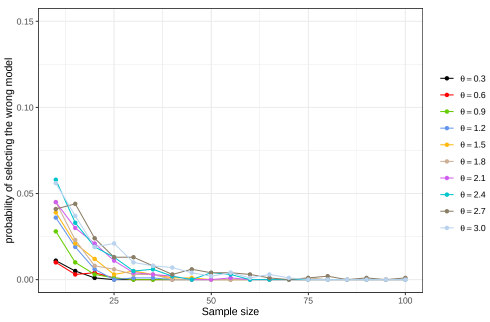
Figure 3.4: Plot of probability of selecting the wrong model when heterogeneity is high $\theta = 3$.



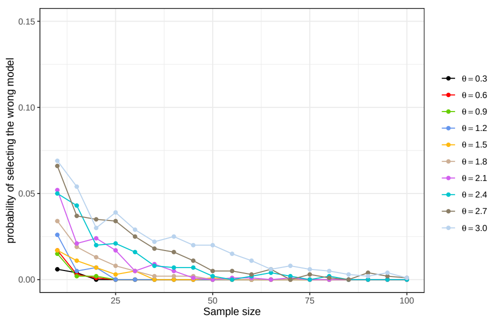
(a) IG model selections by AIC



(b) gamma model selections by AIC

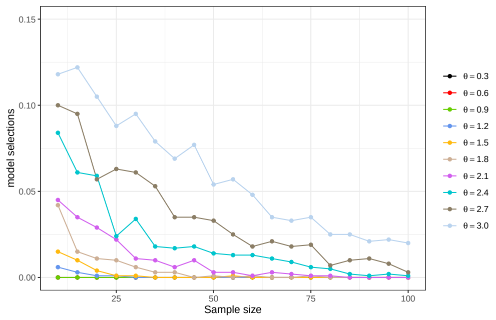


(c) IG model selections by RMSE

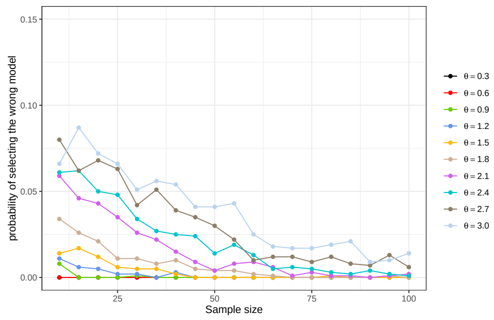


(d) gamma model selections by RMSE

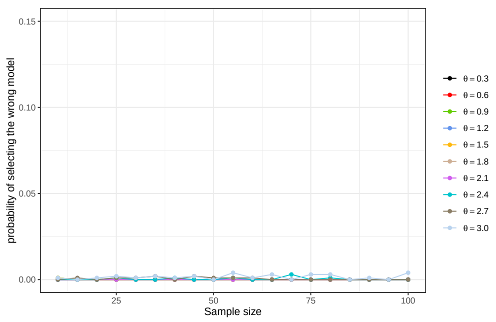
Figure 3.5: Plot of probability of selecting the wrong model when early life failures are considered $\rho = 0.3$.



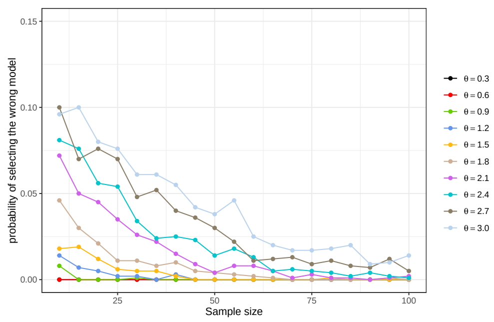
(a) IG model selections by AIC



(b) gamma model selections by AIC



(c) IG model selections by RMSE



(d) gamma model selections by RMSE

Figure 3.6: Plot of probability of selecting the wrong model when component failures are similar to the mid-life phase $\rho = 0.9$.

3.3.3.5 Summary of Analysis

Reflecting on our study, some conclusions can be drawn from our analysis. First, when the sample sizes increase above 50, the likelihood of selecting the wrong model significantly reduces for almost all choice of parameters. This provides confidence that when we have many components, we are likely to select the correct model. However, there will be many cases where the sample size will be much smaller. Second, the effect of heterogeneity for low sample sizes is noticeable. As we increase the heterogeneity from $\theta = 0.3$ to $\theta = 3$, the probability of incorrectly selecting the model increases. Third, we see little difference between the ability of the AIC or the RMSE to select the correct model. When the distribution is known to be gamma, there is little difference between the two methods for selecting the correct distribution. In terms of model selection with respect to the predictive accuracy of the models for the number of observed failures, there were two cases when the underlying distribution is from an IG, that the correct model was identified by RMSE regardless of sample size. In contrast, in terms of statistical fit, the correct model was not easily identified by AIC for small sample sizes, but the selection of the correct model improved as sample size increased. One reason why the RMSE could have identified the correct model regardless of sample size could be because it is a direct measure of the fit of the models to the observed data compared with the AIC that depends on the value of the likelihood function. Another reason why the RMSE could have identified the correct model regardless of the sample size could be because its result is not based on the maximum likelihood whose asymptotic performance depends on the sample size.

3.4 Application to classic dataset

We compare the two frailty models and a PL model using a dataset from the literature, i.e. failure times of air conditioner components from a set of airplanes studied by [293]. This data was used as [293] compared a Heterogenous trend renewal process, an NHPP model with gamma distributed random effect, and a HPP with gamma distributed random effect. [293] found that an NHPP model with gamma distributed random effect provides a better fit than Ordinary NHPP model, identifying heterogeneity among systems in the data. Here we compare the IG and gamma frailty models using the successive failure times before truncation as given in Table 3.1 to see if the IG provides a better fit.

The results of fitting IG frailty model, gamma frailty model and PL model are summarised in Table 3.3. Parameter estimates for the gamma frailty model were obtained as $\lambda = 3.353 \times 10^{-3}$, $\rho = 1.1424$, and $\theta = 0.1334$ while the parameter estimates of the IG frailty model were obtained as $\lambda = 5.861 \times 10^{-3}$, $\rho = 1.1186$, and $\theta = 0.9805$. Parameter estimates of the PL model were obtained as $\lambda = 5.502 \times 10^{-3}$, and $\rho = 1.0720$. The values of θ by the two frailty models, show a low level of heterogeneity between the components of the system. Whilst Lindqvist et al. [293] found the gamma frailty model as the better model for the data when compared to Ordinary NHPP model, based on AIC values in Table 3.4, the IG frailty model has a better fit compared with the PL model and gamma frailty model.

Variance of the estimates in Table 3.3 were derived using bootstrap technique. Bootstrapping is sampling with replacement from observed data to estimate the variability in a statistic of interest. In other words, bootstrap is typically used to estimate quantities associated with the sampling distribution of estimators and test statistics [294]. By resampling from the pseudo-population, we can compute any quantity of interest using the bootstrap technique [284]. From Table 3.3 we observe that the variance of the estimates of λ and ρ of the data are low for all the models. In contrast, the variance of the estimate of θ from IG is higher than those from gamma indicating that the value of the estimate of θ from IG will differ a bit for each sample.

Furthermore, we compared predictions of the expected number of failures for the system from the three models with observed data using RMSE (see Table 3.2 and column "RMSE" in Table 3.4 for the predicted number of failures and RMSE values respectively). The expected number of failures from IG and gamma frailty models were calculated from Eq.(3.22) while the expected number of failures for the PL model was calculated using $E[N_S(t_1, t_2)] = \Lambda_0(t_1, t_2)$. We found that there is little difference in the performance of the models when observed in terms of their predictions of the observed number of system failures.

Next, we compared the models in term of the cumulative number of failures, the cumulative number of failures were derived using the values in Table 3.2. Column "RMSE (cumulative failures)" in Table 3.4 presents the RMSE values for the three models. We found that there is little difference in the predictions of the cumulative observed number of system failures from the PL, gamma and IG frailty models when time is less than 500 days (see Table 3.5 and Figure 3.7 for the cumulative number of failures and the plot of the cumulative number of failures respectively). However, in line with Hougaard [202] findings that the relative frailty distribution among survivors from an IG model becomes more homogeneous with time, we found that once we extend beyond 500 days the survivors become more homogeneous and the IG model substantially outperforms the gamma but it's performance was similar to the PL. Based on the selection values of AIC and "RMSE (cumulative failures)", in terms of model fit and prediction, one can infer that the IG frailty model is marginally better in terms of model fit and better than gamma frailty model in terms of predictions of the expected number of component failures for the air conditioner data when number of days is greater than 500.

Equation (3.22) is useful for predicting the expected number of failures at the system level. For component level prediction, the predicted number of failures for all the components will always be the same, even if there are clear differences in the pattern of failure occurrences. The reason is that the value of the variance parameter θ in the equation for the expected number of failures Eq.(3.22) is the same for each component. Thus, predictions of the expected number of failures for individual components will be computed by using an empirical Bayes approach to account for variability using the mean frailty estimates (see Carlin et al. [295] for more on empirical Bayes).

Using an empirical Bayes approach, we developed a method for individual component prediction of the mean residual life (MRL) and expected number of failures conditional on the expected frailty value. The developed method is then applied to make predictions for components in the Airconditioner data. The developed method uses Bayes theorem to update the frailty distribution for each j component based on the observed data from the j^{th} component. With the updated frailty distribution, we get a point estimate (in this case the mean of the distribution) of the value of the frailty term for the j^{th} component. Then, we use the new frailty term to compute the intensity, survival function, and make individual event prediction of the MRL and expected number of failures for each j^{th} component.

Using Bayes theorem, we can have the posterior distribution of the frailty term z_j^* of the j^{th} component as:

$$h_j(z_j^* \mid \theta) = \frac{L_j(\lambda_0(t_{ij}) \mid z_j) f(z_j;\theta)}{\int_0^\infty L_j(\lambda_0(t_{ij}) \mid z_j) f(z_j;\theta) dz_j}, \tag{3.24}$$

where $h_j(z_j^* \mid \theta)$ is the posterior distribution of Z_j^*, $L_j(\lambda_0(t_{ij}) \mid z_j)$ is the data likelihood for component j, and $f(z_j;\theta)$ is the prior distribution (i.e., the gamma or IG distributions). $L_{j(marg)} = \int_0^\infty L_j(\lambda_0(t_{ij}) \mid z_j) f(z_j;\theta) dz_j$ is the marginal likelihood for component j. We present the derivations of $h_j(z_j^* \mid \theta)$, and the point estimate of the value of the frailty term for IG and gamma in appendix B and C respectively. We used the maximum likelihood methods presented in the third section to estimate θ, and parameters of $\lambda_0(t_{ij})$ then we used the Bayesian method to estimate the updated frailty term. The use of frequentist to estimate parameters and Bayesian method to estimate the value of the random effects is common in mixed-effects models (see for example [69, 296]).

The expected number of failures of the j^{th} component conditional on the expected frailty value between time t_1 and $t_2$is given as:

$$E[N_j(t_1, t_2 \mid \overline{z_j})] = \int_{t_1}^{t_2} \overline{z_j^*} \lambda_0(t) dt. \tag{3.25}$$

System level prediction of expected number of component failures is $E[N_S(t_1, t_2)] = \sum_{j=1}^{m} E[N_j(t_1, t_2)]$. The expression of the MRL is given as:

$$mrl_j(\tau) = \int_\tau^\infty R_j(x \mid \tau) dx = \frac{\int_\tau^\infty R_j(x) dx}{R_j(\tau)}, \tag{3.26}$$

where τ is the end of observation length and $R_j(.) = exp^{-\overline{z_j^*}\Lambda_0(.)}$.

For each of the components (Airplanes) in the Air conditioner data, we predicted the mean residual life and compared predictions of the expected number of failures from the IG and Gamma models with observed data using RMSE. The mean frailty estimate for each component $\overline{z_j}$, mean residual life prediction, and RMSE values for the expected number of failures for each Airplane is presented in Table 3.6. The RMSE and RMSE (cumulative no. of failures) values in Table 3.6 shows that out of 13 components, the IG frailty model is better at predicting the expected number of failures for atleast 9 components compared to the gamma frailty model. For illustration we present the plots of the predictions of the expected number of failures for four Airplanes (Airplane 7909, Airplane 7914, Airplane 7915, and Airplane 7917) in Figures 3.8a, 3.8b, 3.8c, and 3.8d. From the Figures, we see a difference in the predictions of the observed number of failures for each of the Airplanes from the gamma and IG frailty models when time is more than 250 days.

In terms of mean residual life predictions for each component Airplane, we apply Eq.(3.26) for IG and gamma frailty models and found that there are differences in the MRL predictions of the Airplanes by the two models. For some Airplanes, for example Airplanes 7910, 7912, and 8044, the difference was as little as one or two days. In other Airplanes such as: 7909, 7911, 7913, 7914, 7915, 7916 and 7917, the differences ranged from eleven to seventy seven days. However, given the selection of the IG frailty by the RMSE values of a lot of the Airplanes, the MRL predicted values by the IG frailty model may be chosen for optimal maintenance decision making for most of the Airplanes.

As illustrated in Figure 3.7, using only the PL model, the expected number of failures and MRL at the system level can be estimated. In contrast, predictions of the expected number of failures and MRL at the component level cannot be estimated using the PL model. The reason is that the predicted values will be the same for all the components given that they have the same parameter values.

Table 3.1: Failures times for Air conditioners in 13 Airplanes

Airplane Number												
7907	7908	7909	7910	7911	7912	7913	7914	7915	7916	7917	8044	8045
194	413	90	74	55	23	97	50	359	50	130	487	102
209	427	100	131	375	284	148	94	368	304	623	505	311
250	485	160	179	431	371	159	196	380	309		605	325
279	522	346	208	535	378	163	268	650	592		612	382
312	622	407	710	755	498	304	290		627		710	436
493	687	456	722	994	512	322	329		639		715	468
	696	470	792		574	464	332				800	535
	865	494	813		621	532	347				891	594
		550	842		846	609	544				934	728
		570			917	689	732					880
		649				690	811					907
		733				706	899					921
		777				812	945					
		836					950					
		865					955					
		983					991					

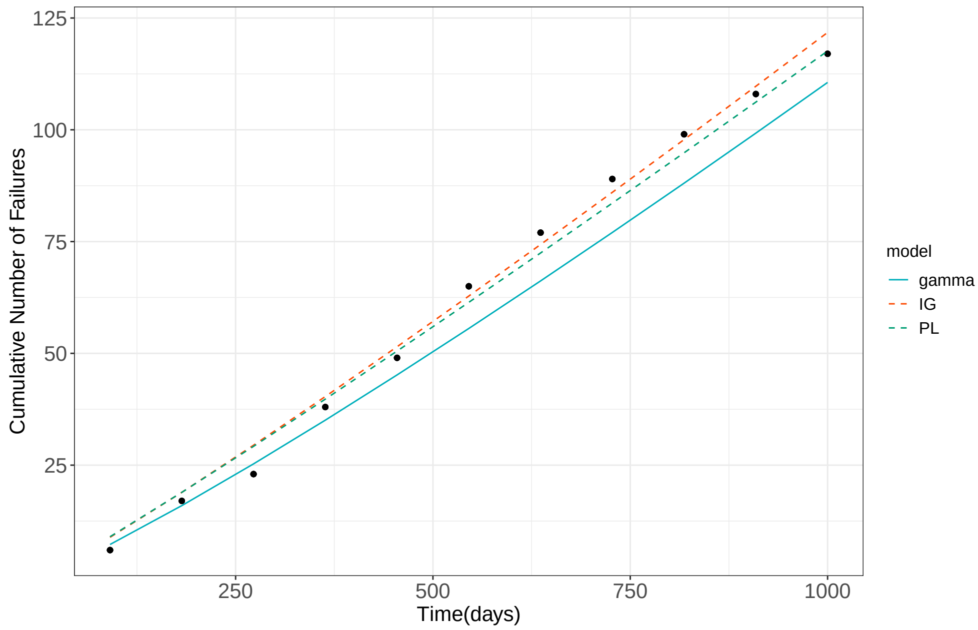


Figure 3.7: Plot of the expected number of failures predicted by the IG frailty model, gamma frailty model and PL model compared to the observed cumulative number of failures within the interval 0 to 1000 hours (system level prediction).

Table 3.2: Predictions of the number of failures from gamma, and IG frailty models and the PL model.

		Expected number of failures		
Time (days)	observed	gamma	IG	PL
0 - 90.90	6	7.2554	8.8649	8.9968
90 - 181.82	11	8.6956	10.0487	9.9176
181.83 - 272.73	6	9.3348	10.5468	10.2977
272.73 - 363.64	15	9.7746	10.8819	10.5525
363.64 - 454.55	11	10.1145	11.1369	10.7462
454.55 - 545.55	16	10.3935	11.3438	10.9031
545.55 - 636.36	12	10.6310	11.5183	11.0353
636.36 - 727.27	12	10.8385	11.6695	11.1498
727.27 - 818.18	10	11.0230	11.8030	11.2508
818.18 - 909.09	9	11.1894	11.9227	11.3414
909.09 - 1000	9	11.3412	12.0313	11.4235

Table 3.3: Parameter estimates of gamma, and IG frailty models and the PL model when fitted to Air conditioner failure time data.

	λ	Var λ	ρ	Var ρ	θ	Var θ
gamma	0.003353	0.0000035	1.142425	0.00224294	0.133469	0.0048281
IG	0.005861	0.0000046	1.118666	0.00454367	0.980574	0.3487936
PL	0.005502	0.0000036	1.0720	0.00189678		

3.5 Conclusion

In this chapter, we applied the IG frailty model for analysing failure data from heterogeneous repairable systems. The IG frailty model, which combines the PL model and IG distribution, assumes that the relative frailty distribution among survivors becomes more homogeneous over time. This is in contrast to the commonly used gamma frailty models which assume that the relative frailty distribution among survivors is independent of age. The main objective of this paper were to evaluate the application of IG frailty model for analysing failure data from heterogeneous repairable systems, compare its results with the gamma frailty model, and develop a method for event prediction based on the IG and gamma frailty models. To accomplish the objective, we developed the IG frailty model and a method for parameter estimation of the IG

Table 3.4: AIC and RMSE values of gamma, and IG frailty models and the PL model when fitted to Air conditioner failure time data.

	AIC	RMSE	RMSE (Cumulative failures)
gamma [293]	1335.281	2.893	7.490
IG	1331.92	2.856	3.207
PL [293]	1337.34	2.844	3.570

Table 3.5: Predictions of the cumulative number of failures from gamma, and IG frailty models and the PL model.

	Cumulative number of failures			
Time (days)	observed	gamma	IG	PL
90.90	6	7.2554	8.8649	8.9968
181.83	17	15.9510	18.9137	18.9145
272.73	23	25.2859	29.4605	29.2122
363.64	38	35.0605	40.3425	39.7648
454.55	49	45.1751	51.4794	50.5110
545.55	65	55.5686	62.8233	61.4142
636.36	77	66.1996	74.3416	72.4496
727.27	89	77.0381	86.0112	83.5994
818.18	99	88.0612	97.8143	94.8503
909.09	108	99.2507	109.7371	106.1917
1000	117	110.5919	121.7684	117.6153

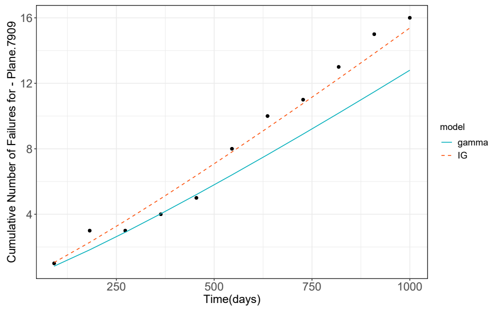
frailty model using maximum likelihood estimation and numerical methods. We found that the estimator is robust to different parameter values, and beyond a sample size of 10, the estimator performs well.

A comparison of the gamma and IG frailty models was conducted to examine whether both models are good alternatives of each other. Statistical fit of the gamma and IG frailty models as well as the prediction performance was thoroughly studied and compared. We found that regardless of the degree of heterogeneity or frequency of failures when early component behaviour is concerned, the probability of selecting a wrong model is low whether for model fit or for prediction purpose. A wrong model is only selected when the sample size is small.

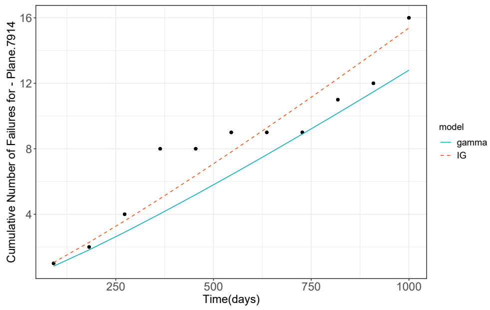
Furthermore, we compared the two frailty models and a PL model to a classic dataset where the gamma frailty model had been studied. Our results found that the

Table 3.6: Mean frailty $\overline{z_j^*}$, Mean residual life, and RMSE value for the expected number of failures of each Airplane

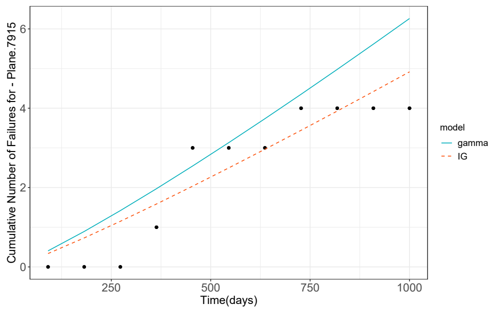
	$\overline{z_j^*}$		**Mean residual life (Days)**		**RMSE**		**RMSE (cumulative no. of failures)**	
Airplane ID	gamma	IG	gamma	IG	gamma	IG	gamma	IG
Plane.7907	0.8197	0.4849	117.3	136.5987	1.0137	1.0012*	1.5184	1.7143
Plane.7908	0.9412	0.6105	102.3372	108.7911	0.8338	0.8375	1.2177	1.2283
Plane.7909	1.4272	1.1561	67.8388	56.3213	0.8103	0.7582*	2.0142	0.7714*
Plane.7910	1.0019	0.6758	96.2018	98.3719	0.8536	0.8498*	1.0555	1.0336*
Plane.7911	0.8197	0.4849	117.3	136.5987	0.6726	0.6603*	0.905	0.5656*
Plane.7912	1.0627	0.7424	90.7624	89.6296	0.7873	0.7863*	1.0447	0.9446*
Plane.7913	1.2449	0.9469	77.5854	70.3957	1.3625	1.3511*	2.6173	1.8336*
Plane.7914	1.4272	1.1561	67.8388	56.3213	1.3958	1.3664*	2.0152	1.4194*
Plane.7915	0.6982	0.3696	137.3962	178.5138	0.6737	0.6463*	1.1075	0.6626*
Plane.7916	0.8197	0.4849	117.3	136.5987	0.807	0.7939*	0.6757	0.738
Plane.7917	0.5767	0.2703	165.8234	242.7943	0.4922	0.4183*	1.6388	0.7609*
Plane.8044	1.0019	0.6758	96.2018	98.3719	0.774	0.7828	1.7996	1.8478
Plane.8045	1.1842	0.8781	81.5381	75.8732	0.7651	0.7569*	1.1155	0.8169*



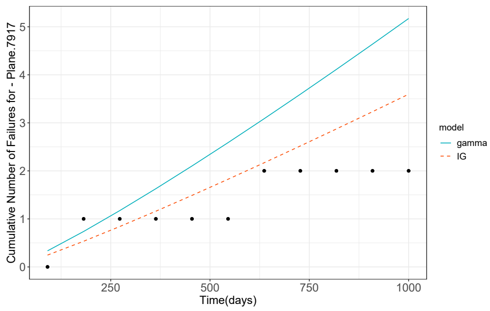
(a) Airplane 7909



(b) Airplane 7914



(c) Airplane 7915



(d) Airplane 7917

Figure 3.8: Plot of the expected number of failures predicted by the IG and gamma frailty models compared to the observed cumulative number of failures within the interval 0 to 1000 hours (component level prediction).

IG frailty model was better in terms of model fit and outperforms the gamma for system level prediction of expected number of failures when the number of days was greater than 500.

Using empirical Bayes framework, we developed a method for prediction of expected number of failures and mean residual life at the component level based on IG and gamma frailty models. We applied the developed method to the same classic dataset. Our results show that there is difference in the predictions of the observed number of failures and MRL for each of the Airplanes from the gamma and IG frailty models when time is more than 250 days.

Lastly, whilst still accounting for heterogeneity, the standard PL model is not suitable for predictions of the expected number of failures and MRL at the component level because the predicted values will be the same for all the components given that they have the same parameter values.

In the next chapter (chapter 4), we will develop a system reliability model for estimating the reliability of a load sharing multi-component system with spatial dependent components considering proximity effects.

Appendix - Chapter 3

A.1 Appendix A

The derivation of the Maximum likelihood estimation for gamma frailty model is presented below. If we let z_j be the frailty random variable following iid gamma distribution then the density of the two-parameter gamma distribution will be:

$$f(z_j) = \frac{z_j^{k-1} exp^{-\frac{z_j}{\delta}}}{\delta^k \Gamma(k)},$$

where $z_j \geq 0$, k is shape parameter and δ is scale parameter. Further more, E(z_j)=$k\delta$ and Var(z_j)=$k\delta^2$. To make gamma frailty model identifiable, restriction is used which requires expectation value to equal 1. Thus, if we set $k = 1$ and $\delta = \theta$, we will have a one-parameter gamma distribution given by:

$$f(z_j) = \frac{z_j^{\frac{1}{\theta}-1} exp^{-\frac{z_j}{\theta}}}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})},$$

where $\mathrm{E}(z_j) = 1$ and $\mathrm{Var}(z_j) = \theta$. The conditional likelihood function for a single component j with random effect z_j is given below as:

$$L_j(\lambda_0(t_{ij}) \mid z_j) = \left(\prod_{i=1}^{n_j} z_j \lambda_0(t_{ij})\right) exp^{(-z_j \Lambda(\tau))},$$

where $\lambda_0(t_{ij}) = \lambda \rho t_{ij}^{\rho-1}$ and $\Lambda_0(\tau) = \lambda \tau^{\rho}$. Since z_j is unobservable from the data, the contribution of component j to the full likelihood from the conditional likelihood can only be obtained by deriving the unconditional likelihood function with respect to z_j. This implies computing the expected value by integrating out z_j and thus deriving the gamma distribution parameter which directly affects the full likelihood. Thus, we have that:

$$L_j(\theta \mid \lambda_0(t_{ij})) = \frac{\prod_{i=1}^{n_j} \lambda_0(t_{ij})}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} \int_0^{\infty} z_j^{\frac{1}{\theta}-1+n_j} exp^{-\frac{z_j}{\theta}-z_j\Lambda(T)} dz_j.$$

For gamma frailty model, the expression of $L_{j(marg)}$ is given by:

$$L_{j(marg)} = \left[\frac{\prod_{i=1}^{n_j} \lambda_0(t_{ij})\Gamma(n_j + \frac{1}{\theta})}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})\left(\Lambda_0(t) + \frac{1}{\theta}\right)^{(n_j+\frac{1}{\theta})}}\right]. \tag{27}$$

It then follows that the full unconditional likelihood of all the components is given below:

$$L = \prod_{j=1}^{m} \left[\frac{\prod_{i=1}^{n_j} \lambda \rho t^{\rho-1}}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} \frac{\Gamma(n_j + \frac{1}{\theta})}{\left(\lambda(\tau)^{\rho} + \frac{1}{\theta}\right)^{(n_j+\frac{1}{\theta})}}\right].$$

Taking the derivative of the log likelihood function with respect to λ and ρ and θ leads to:

$$\frac{\delta l}{\delta \lambda} = \frac{n}{\lambda} - m\tau^{\rho} = 0,$$

where $n = \sum_{j=1}^{m} n_j$.

$$\frac{\delta l}{\delta \rho} = \frac{n}{\rho} + \sum_{j=1}^{m}\sum_{i=1}^{n_j} \log t_{ij} - m\lambda\tau^{\rho}\log\tau.$$

$$\frac{\delta l}{\delta \theta} = \sum_{j=1}^{m} \log\Gamma(n_j + \frac{1}{\theta}) - \left[m\log\Gamma(\frac{1}{\theta}) + m\log(\theta^{\frac{1}{\theta}}) + m(n_j + \frac{1}{\theta})\log[\lambda\tau^{\rho} + \frac{1}{\theta}]\right].$$

Thus estimators for the scale λ and shape ρ parameters are of the form:

$$\widehat{\lambda} = \frac{n}{m\tau^{\rho}},$$

and

$$\widehat{\rho} = \frac{n}{n\log\tau - \sum_{j=1}^{m}\sum_{i=1}^{n_j}\log t_{ij}}.$$

As can be seen, estimators for λ and ρ are explicitly derived whereas an analytic expression for the estimator of θ can not be easily derived. So a numerical optimization algorithm will be used to estimate the parameters.

B.2 Appendix B

The expressions for the posterior distribution, expected frailty value, expected number of failures and marginal reliability function for the IG frailty model is derived as follows.

The marginal reliability function of the j^{th} component is given as:

$$R_j(t) = \int_0^{\infty} R_j(t \mid z_j) f(z_j;\theta) dz_j, \tag{28}$$

where $f(z_j;\theta)$ is the IG distribution and $R_j(t \mid z_j) = exp^{-z_j\Lambda_0(t)}$. Thus,

$$R_j(t) = \int_0^{\infty} \frac{1}{\sqrt{2\pi\theta}} z_j^{-\frac{3}{2}} exp^{\left(-z_j\Lambda_0(t) - \frac{(z_j-1)^2}{2z_j\theta}\right)} dz_j, \tag{29}$$

and

$$R_j(t) = exp^{\left(\frac{1}{\theta}\left(1-\sqrt{1+2\theta\Lambda_0(t)}\right)\right)}, \tag{30}$$

where $R_j(t)$ is derived by replacing s with $\Lambda_0(t)$ in the Laplace transform for gamma distribution $L_j[s] = exp^{\left(\frac{1}{\theta}(1-\sqrt{1+2\theta s})\right)}$. (See Mundo et al [297] for derivation of Laplace transform of IG and Gamma frailty models).

The marginal expected number of failures of the j^{th} component between time t_1 and t_2 is given as:

$$\begin{aligned} E[N_j(t_1,t_2)] &= \Lambda_j(t_1,t_2) = -log(R_j(t_1,t_2)) \\ &= \frac{1}{\theta}(\sqrt{1+2\theta\Lambda_0(t_1,t_2)}-1). \end{aligned} \tag{31}$$

The marginal expected number of component failures in the system between time t_1 and t_2 is given as:

$$E[N_S(t_1,t_2)] = \sum_{j=1}^{m} E[N_j(t_1,t_2)] = \frac{m}{\theta}(\sqrt{1+2\theta\Lambda_0(t_1,t_2)}-1). \tag{32}$$

where m is the number of components in the system.

Next we derive the posterior distribution for z_j^* given a IG prior distribution. Replacing the expressions of $L_{j(marg)}$, $L_j(\lambda_0(t_{ij}) \mid z_j)$, and $f(z_j;\theta)$ in the Bayes formula Eq.(3.24), then the posterior distribution of z_j^* for the j^{th} component will be derived as:

$$h_j(z_j^* \mid \theta) = \frac{z_j^{n_j-\frac{3}{2}} exp^{\left(-z_j\frac{(2\theta\Lambda_0(t)+1)}{2\theta}-\frac{1}{2\theta z_j}\right)}}{2\,(2\theta\Lambda_0(t)+1)^{-(\frac{n_j-1/2}{2})}\,k_{(n_j-1/2)}\left[\frac{\sqrt{(1+2\theta\Lambda_0(\tau))}}{\theta}\right]}. \tag{33}$$

The value of the frailty term for component j is then derived as a point estimate by finding the mean of the updated frailty distribution $h_j(z_j^* \mid \theta)$. The mean of the updated frailty distribution $\overline{z_j^*} = E[z_j^*]$ is derived as:

$$E[z_j^*] = \int_0^\infty z_j^* h_j(z_j^* \mid \theta) dz_j^*. \tag{34}$$

Then

$$E[z_j^*] = \frac{\int_0^\infty z_j^* . z_j^{*\, n_j - \frac{3}{2}} exp^{\left(-z_j^* \frac{(2\theta\Lambda_0(t)+1)}{2\theta} - \frac{1}{2\theta z_j}\right)} dz_j^*}{2\left(2\theta\Lambda_0(t)+1\right)^{-\left(\frac{n_j-1/2}{2}\right)} k_{(n_j-1/2)}\left[\frac{\sqrt{(1+2\theta\Lambda_0(\tau))}}{\theta}\right]}, \tag{35}$$

and

$$E[z_j^*] = \frac{(2\theta\Lambda_0(t)+1)^{-\left(\frac{n_j+1/2}{2}\right)} k_{(n_j+1/2)}\left[\frac{\sqrt{(1+2\theta\Lambda_0(\tau))}}{\theta}\right]}{(2\theta\Lambda_0(t)+1)^{-\left(\frac{n_j-1/2}{2}\right)} k_{(n_j-1/2)}\left[\frac{\sqrt{(1+2\theta\Lambda_0(\tau))}}{\theta}\right]}, \tag{36}$$

where n_j is the number of events that component j has experienced. θ is the estimated variance parameter from the gamma frailty likelihood estimator, and $\Lambda_0(t)$ is the cumulative intensity of the component until time t. $k_{(n_j-1/2)}[.]$ is a modified bessel function of the second kind.

C.3 Appendix C

The expressions for the posterior distribution, expected frailty value, expected number of failures and marginal reliability function for the gamma frailty model is derived as follows.

The marginal reliability function of the j^{th} component is given as:

$$R_j(t) = \int_0^\infty R_j(t \mid z_j) f(z_j; \theta) dz_j, \tag{37}$$

where $f(z_j;\theta) = \frac{z_j^{\frac{1}{\theta}-1} exp^{-\frac{z_j}{\theta}}}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})}$ is a gamma distribution and $R_j(t \mid z_j) = exp^{-z_j\Lambda_0(t)}$.

Thus,

$$R_j(t) = \int_0^\infty exp^{-(z_j\Lambda_0(t))} \frac{z_j^{\frac{1}{\theta}-1} exp^{-\frac{z_j}{\theta}}}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} dz_j, \tag{38}$$

and

$$R_j(t) = (\theta\Lambda_0(t) + 1)^{-\frac{1}{\theta}}, \tag{39}$$

where $R_j(t)$ is derived by replacing s with $\Lambda_0(t)$ in the Laplace transform for gamma distribution $L_j[s] = (\theta s + 1)^{-\frac{1}{\theta}}$.

The expected number of failures of the j^{th} component between time t_1 and t_2 can be derived from the marginal reliability as:

$$\begin{aligned} E[N_j(t_1, t_2)] &= \Lambda_j(t_1, t_2) = -log(R_j(t_1, t_2)) \\ &= \frac{1}{\theta} log(\theta\Lambda_0(t_1, t_2) + 1)^{-\frac{1}{\theta}}. \end{aligned} \tag{40}$$

The expected number of component failures in the system between time t_1 and t_2 is given as:

$$E[N_S(t_1, t_2)] = \sum_{j=1}^{m} E[N_j(t_1, t_2)] = \frac{m}{\theta} log(\theta\Lambda_0(t_1, t_2) + 1)^{-\frac{1}{\theta}}. \tag{41}$$

where m is the number of components in the system.

Next we derive the posterior distribution for z_j^* given a gamma prior distribution. Replacing the expressions of $L_{j(marg)}$, $L_j(\lambda_0(t_{ij}) \mid z_j)$, and $f(z_j; \theta)$ for the gamma frailty model in the Bayes formula Eq.(3.24), then the posterior distribution of z_j^* for the j^{th} component will be derived as:

$$h_j(z_j^* \mid \theta) = \frac{z_j^{n_j+\frac{1}{\theta}-1} exp^{-z_j\left(\Lambda_0(t)+\frac{1}{\theta}\right)} (\Lambda_0(t) + \frac{1}{\theta})^{n_j+\frac{1}{\theta}}}{\Gamma(n_j + \frac{1}{\theta})}. \tag{42}$$

The expression for the mean of the updated frailty distribution $\overline{z_j^*} = E[z_j^*]$ is given as:

$$E[z_j^*] = \frac{\int_0^\infty z_j^*.{z_j^*}^{n_j+\frac{1}{\theta}-1} exp^{-z_j^*\left(\Lambda_0(t)+\frac{1}{\theta}\right)}(\Lambda_0(t)+\frac{1}{\theta})^{n_j+\frac{1}{\theta}} dz_j^*}{\Gamma(n_j+\frac{1}{\theta})}. \tag{43}$$

However, $h_j(z_j^* \mid \theta)$ is a gamma density function, thus the associated posterior mean is:

$$E[z_j^*] = \frac{(n_j+\frac{1}{\theta})}{(\Lambda_0(t)+\frac{1}{\theta})}, \tag{44}$$

where n_j is the number of events that component j has experienced. θ is the estimated variance parameter from the gamma frailty likelihood estimator, and $\Lambda_0(t)$ is the cumulative intensity of the component until time t.

Chapter 4

Reliability analysis of a load-sharing system with spatial dependence, and proximity effects

In chapter 3, we studied the problem of reliability prediction for a repairable system subject to unobserved heterogeneity. This chapter considers dependence in the form of load-sharing in multi-component systems. In section 2.4.2 of chapter 2, we established the following: First, no research has considered models for reliability prediction of load-sharing systems with spatial dependence and proximity effect. Second, the consequence of ignoring spatial effect in a load-sharing system has not been investigated.

Based on the established gaps in section 2.4.2 of chapter 2, our contribution in this chapter is twofold. First, we will develop a model for estimating the reliability of a load-sharing system with spatial dependent components. Second, using the developed model, we will examine the importance and significance of the spatial effect. We will extend the classic capacity flow model to characterize the relationship between the failure rate of a component and the load imposed on the component. The proposed model will capture the load-life relationship of a component considering spatial dependence. In the setup of the chapter, proximity effect will be considered in load-sharing, that is, we will consider the case that the redistribution of a failed component's load depends on proximity to the failed component. A Markov model will be used to characterize the deterioration process of the entire system. Furthermore, a modified Euler's method will be used to derive the system state probabilities and corresponding system reliability. We will then formulate an algorithm based on Monte Carlo simulation to validate the derived reliability estimations. Finally, numerical analysis will be conducted to assess the developed model.

The remainder of this chapter is outlined as follows: Section 4.1 introduces the System description and load-sharing rule for the study. Section 4.2 presents the methods for modelling spatial dependence and proximity effect when distance information is either available or not available. Section 4.3 shows the derivation of the load function expression. Section 4.4 presents the formulation of the system state transition and failure rate function for homogeneous and heterogeneous components. Section 4.5 presents parameter estimation. Sections 4.6, 4.7 and 4.8 presents illustrations of the developed model through numerical examples. Section 4.9 presents the conclusion.

4.1 System description and load-sharing rule

Consider a load-sharing system that consists of n components connected in parallel. The following assumptions are made to better position our study:

- Each component can either be in a working or failed state while the system is multi-state and can function at different performance levels depending on the states of its components.
- The lifetimes of the components are load dependant and follow exponential distributions.
- The system fails if the sum of the loads on each working component at time t is less than the total system load L.

- When a component fails, its working spatial neighbours in any direction (as long as they are in close proximity) can take up the failed component's load.

The components are spatial dependent and the system structure is known beforehand. The number of spatial neighbours a component has is determined by the number of links with other components. The system operates such that at time $t = 0$ the total load L of the system is shared by all the components. The components share the constant system load L in the proportion $\gamma_1, \gamma_2, \ldots, \gamma_n$, where γ_i is the proportion taken by component i at time $t = 0$ and n is the number of components in the system. If the components equally share the load, then we can simply let $\gamma_i = \gamma_j$ for any $1 \leq i, j \leq n$. Let i denote a working component, and j denote the index of a failed neighbour of component i. The load z_i taken by component i at initial time $t = 0$ is given by [248]:

$$z_i = \frac{\gamma_i L}{\sum_{i=1}^{n} \gamma_i}. \tag{4.1}$$

Denote $z_i(j)$ as load on component i given that its neighbour component j has failed. We have

$$z_{i(j)} = z_i + \vartheta_{ij} z_j, \tag{4.2}$$

where ϑ_{ij} denotes the ratio of a failed component's load that affects the failure rate of a working adjacent component. It can be viewed as the proportion of a failed component's load that its working spatial neighbour will take on. The proportion of extra load is a function of whether or not a failed component is proximate to a working component. ϑ_{ij} takes values between 0 and 1 as a result. If $\vartheta_{ij} = 0$ a working component is not impacted by the failure of component j. We will refer to ϑ_{ij} as the *proximity effect* and refer to the expression $z_{i(j)}$ as the *interacting load function.*

Similarly, the load on component i after the failure of its second neighbour k, $z_{i(j,k)}$, is given by:

$$z_{i(j,k)} = z_{i(j)} + \vartheta_{ik} z_k.$$

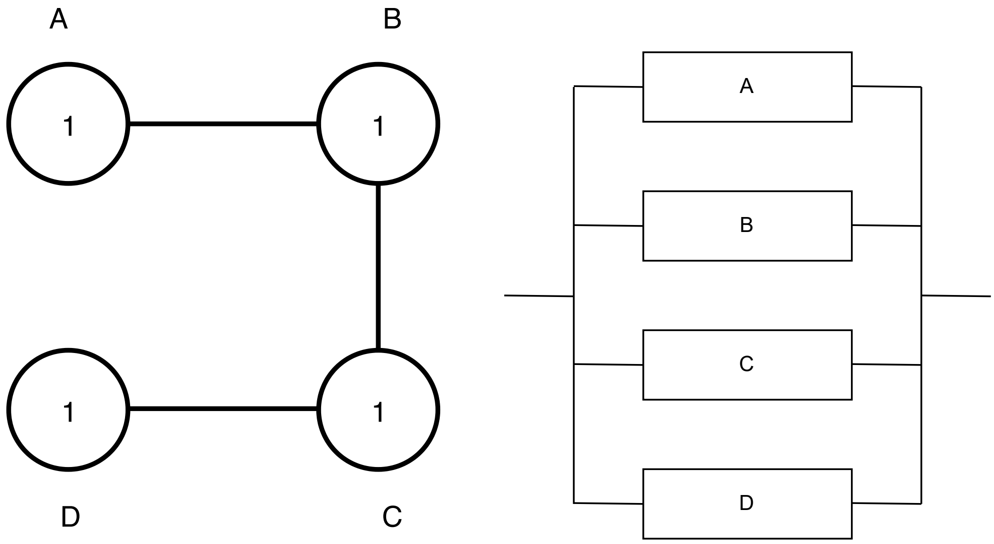
The system fails when the sum of loads on the working components is less than the load placed on the system, i.e., $\sum_{i=1}^{n} z_i < L$.

In this chapter, we focus on parallel system because parallel structures are more common in practice for load-sharing systems. However, the parallel system considered here is different from traditional parallel systems. Traditional parallel systems fail when all components fail, while the system under investigation here fails when the system cannot bear a specified load. As such, our system is more analogous to a performance-based system that fails when it cannot sustain its performance (load).

4.2 Modelling spatial dependence and proximity effect

In order to account for components' interaction with each other, we will define spatial dependence in terms of a given system structure. We assume that if we know the spatial arrangement of components and connections between them, we may be able to infer their dependency. We introduce a conceptual dependence model to account for spatial dependence between components. The dependence model is based on the assumption that pairs of components with a direct link (in the form of solid lines) between them are close enough to interact while pairs of components without a link (i.e, have no line) are independent even if they are positioned next to each other.

To illustrate the dependence concept, we consider the four-component system depicted in Figure 4.1. Note that in Figure 4.1b, the system is in a parallel structure, and in Figure 4.1a the links between the components indicate proximity of the components. The systems are composed of four components indexed as A, B, C and D. Components A, B, and C have solid lines between them while A and D have no line indicating that even though components A and D are spatial neighbours in terms of their position next to each other, they do not influence each other. By definition, component B is spatially dependent on A and C while component A is only spatially dependent on component B. Component C is spatially dependent on component B and D while component D is only spatially dependent on component C. While the performance of components B and C could be influenced by its two spatial neighbours, components A and D's performance is only influenced by one spatial neighbour. In contrast, component pair A and D in



(a) Spatial connections - three links (b) Block diagram of parallel connections

Figure 4.1: Visual depiction of four component system

the two structures are not spatially dependent so if one of them is in failed state, the other one's reliability is not affected.

4.2.1 Modelling the proximity effect

In this section, we will assume that the spatial arrangement (spatial pattern) of components in the system is known and that there is no information about the distance between components. In section 4.2.2 we will develop proximity models assuming that we know the distance between components. We assume that one can derive the proximity effect (which describes the proportion of load that a working proximate component takes from a failed component) between components using their spatial arrangement and the link between them (that is, the spatial dependence). An example of a load-sharing system that the method could be applied to is an intelligent air conditioning system. In the literature, the lattice, star and circular structures of the intelligent air conditioning has been used to infer spatial dependence and load distribution between components (see [52, 53, 54]). Let θ_{ij} represent spatial dependence of two components such that if they have a direct link $\theta_{ij} = 1$ otherwise $\theta_{ij} = 0$. We will introduce a dependence matrix to capture each θ_{ij}.

Let us assume that θ_{ij} is updated according to the state of the components then at time $t = 0$ when all components are working, the dependence matrix $[\boldsymbol{\theta_{ij}}]$ will be:

$$[\boldsymbol{\theta_{ij}}] = \begin{bmatrix} \theta_{11} & \theta_{12} & \dots \\ \vdots & \ddots & \\ \theta_{n1} & & \theta_{nn} \end{bmatrix},$$

where the j^{th} column of matrix $[\boldsymbol{\theta_{ij}}]$ denotes the index of a failed component while the i^{th} row is the index of its working neighbours.

In order to ensure that the sum of proportion of the j^{th} failed component's load taken up by all its working neighbours at time t sum up to one, the elements of each column in matrix $[\boldsymbol{\theta_{ij}}]$ will be column normalized. If we assume that the j^{th} column of $[\boldsymbol{\vartheta_{ij}}]$ represents the proportion of its load that a failed component j will transfer to its working proximate neighbours i at time t, then the normalization equation will be:

$$\vartheta_{ij} = \frac{\theta_{ij}}{\sum_{i=1}^{n} \theta_{ij}}. \tag{4.3}$$

After column normalizing, we would derive a new matrix $[\boldsymbol{\vartheta_{ij}}]$ with normalized values ϑ_{ij} which represents the proximity effect, i.e.,

$$[\boldsymbol{\vartheta_{ij}}] = \begin{bmatrix} \vartheta_{11} & \vartheta_{12} & \dots \\ \vdots & \ddots & \\ \vartheta_{n1} & & \vartheta_{nn} \end{bmatrix},$$

where $\vartheta_{ii} = 0$ and ϑ_{ij} is non-negative and the values in each column have unit sum, i.e., $\sum_{i=1}^{n} \vartheta_{ij} = 1$, for any $j = 1, \dots, n$.

4.2.2 Proximity models with distance information

For load-sharing systems with spatial dependent components, [32] defined the load-sharing rule as a rule in which the load on a failed component is transferred to proximate components, and the proportion of the load that the working components inherit depends on their distance to the failed component. They mentioned that examples of this kind of systems include cables supporting bridges and other structures, composite materials with bounding matrix joins, and transmission. Following the load distribution rule, a working component will be affected by its failed neighbour j with an increased load $z_{i(j)}$ as:

$$z_{i(j)} = z_0(1 + p_{ij}), \tag{4.4}$$

where p_{ij} describes how much a failed component's load affects its working proximate neighbours. If $p_{ij} = 0$, it means two components are not close enough to have load-sharing interaction whereas $p_{ij} \leq 1$ represents the degree to which two proximate component affect each other's performance.

In this section we consider the case that there is information about the distance d_{ij} between components and one can derive the proximity values p_{ij} which would describe how much a failed component's load affects its working proximate neighbours. According to [32], the proportion of the load inherited by the working component depends on their distance to the failed component thus we will assume that p_{ij} is a function of the distance d_{ij} between components. We will introduce two distance decay spatial weights matrix methods applied in geostatistics for deriving the proximity values. We shall refer to the distance decay spatial weights methods as *constant proximity model* and an *exponential proximity model*. The methods are based on the idea that areas closer to an area of interest have more of an influence than those further away and thus are weighted as such in [298].

Let us assume that one can represent all the distances between components as elements of a distance matrix $[D_{ij}]$ given as:

$$[D_{ij}] = \begin{bmatrix} d_{11} & d_{12} & \dots \\ \vdots & \ddots & \\ d_{n1} & & d_{nn} \end{bmatrix},$$

where $d_{11} = d_{22} = \ldots = d_{nn} = 0$. Let us assume that distance is an important criterion of the degree of influence between components and that there is a threshold distance beyond which there is no influence between components. Let α denote the threshold distance. First, we introduce the *constant proximity model* which assumes that all component pairs within a distance threshold α have the same proximity effect p_{ij} regardless of their distance apart while component pairs with distance more than the threshold do not interact [298]. Each p_{ij} is derived by:

$$p_{ij} = \begin{cases} c, & 0 \leq d_{ij} \leq \alpha \\ 0, & d_{ij} > \alpha, \end{cases} \tag{4.5}$$

where c is a predefined constant value and is the same for all pair of components within the threshold distance α. All other components outside the threshold distance have $p_{ij} = 0$.

In contrast, the exponential proximity model assumes a diminishing proximity effect which reduces as the distance between components increase up to the threshold α. The exponential proximity model could find application to a cable-strut system in suspension bridge where the booms far from the failed one are only subjected to the indirect force and the transfer effect on them is less compared with those on the proximate booms [50]. The exponential proximity model unlike the constant proximity model allows for variability of the p_{ij}. The exponential proximity model is given by:

$$\widehat{p}_{ij} = e^{-\frac{d_{ij}}{\alpha}}, \tag{4.6}$$

where α is the threshold and d_{ij} is the distance between component i and j. If we take the limit of the exponential proximity model as d_{ij} is close to zero i.e.,

$$\lim_{d_{ij} \to 0} e^{-\frac{d_{ij}}{\alpha}} = 1. \tag{4.7}$$

As the distance between two components is reduced to zero, the influence component j can have on a neighbour i increases to one. In contrast, if we take the limit of the exponential proximity model as d_{ij} tends to ∞, i.e.,

$$\lim_{d_{ij} \to \infty} e^{-\frac{d_{ij}}{\alpha}} = 0. \tag{4.8}$$

In the operation of industrial systems in real-life, highly distant components would barely interact directly. As a result, we set every element $\widehat{p}_{ij} = 0$ for all $(d_{ij} > \alpha)$. In this work, we will consider the exponential model. We also assume that the threshold α can be derived as the average of pairs of distances in the system (without repetition) and $\alpha = \frac{\sum d_{ij}}{l}$. l is the number of inter-component links and $d_{ij} = d_{ji}$. α ensures that all components j whose distances d_{ij} from a component i less than or equal to mean distance $(d_{ij} \leq \alpha)$ are influenced by i while components j with distances greater than the mean distance $(d_{ij} > \alpha)$ are not affected. Note that α is a scaling parameter used to scale the effect of distance. In practice, average distance is used to scale the distance (e.g., see [298]). Mathematically other scaling factors can also be explored. In this research, we adopt the average distance to model α. α is calculated by taking an average of the upper or lower triangular matrix instead of the entire matrix as the distance values will be doubled.

After deriving each proximity value $\widehat{p}_{ij}$ using an exponential proximity model, we represent the proximity effects between components of a system as elements of a proximity matrix:

$$\left[\widehat{P}_{ij}\right] = \begin{bmatrix} 1 & \widehat{p}_{12} & \dots \\ \vdots & \ddots & \\ \widehat{p}_{n1} & & 1 \end{bmatrix}.$$

The diagonal elements $\widehat{p}_{ii}$ of the $\left[\widehat{P}_{ij}\right]$ matrix contains unit values indicating that a component is in close proximity with itself. If the unit values of each $\widehat{p}_{ii}$ elements are left, they would affect how the load of a failed component is shared by its working neighbours. In order to remove the unit values, we will subtract unit matrix from $\left[\widehat{P}_{ij}\right]$ so that we have:

$$\left[P^*_{ij}\right] = \begin{bmatrix} 1 & \widehat{p}_{12} & \dots \\ \vdots & \ddots & \\ \widehat{p}_{n1} & & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & \dots \\ \vdots & \ddots & \\ 0 & & 1 \end{bmatrix} = \begin{bmatrix} 0 & p^*_{12} & \dots \\ \vdots & \ddots & \\ p^*_{n1} & & 0 \end{bmatrix}.$$

Similar to θ_{ij} in section 4.2.1, p^*_{ij} functions according to the state of the components and the matrix $[\boldsymbol{P_{ij}}^*]$ provides proximity information at time $t = 0$ when all components are working.

Similar to $[\boldsymbol{\vartheta_{ij}}]$, the proximity elements of the $\left[P^*_{ij}\right]$ matrix will be column normalized, assuming that the j^{th} column of $\left[P^*_{ij}\right]$ contains all proximity effects of neighbours influenced by the failed state of component j. The normalization equation is given by:

$$p_{ij} = \frac{p^*_{ij}}{\sum_{i=1}^{n} p^*_{ij}}. \tag{4.9}$$

After column normalizing, we derive a new matrix $[P_{ij}]$ with normalized proximity values of p_{ij}. The proximity values in each column are normalized to have unit sum, i.e.,

$$\sum_{i=1}^{n} p_{ij} = 1, j = 1, \ldots, n \quad , \tag{4.10}$$

where $p_{ii} = 0$ and p_{ij} is non-negative.

To illustrate the methodology, let us consider the four-component system illustrated in section 4.2.1. If we are given the distance between each pair of component in the distance matrix $[D_{ij}]$ as

$$[D_{ij}] = \begin{bmatrix} 0 & 30 & 110 & 20 \\ 30 & 0 & 10 & 100 \\ 110 & 10 & 0 & 50 \\ 20 & 100 & 50 & 0 \end{bmatrix}.$$

The average distance $\alpha = 53.33333$ is derived by taking the mean of elements in the upper triangle of the distance matrix. We then applied the exponential proximity model to derive the proximity values in matrix $[P_{ij}]$ below:

$$[P_{ij}] = \begin{bmatrix} 1 & 0.5697828 & 0 & 0.6872893 \\ 0.5697828 & 1 & 0.8290291 & 0 \\ 0 & 0.8290291 & 1 & 0.3916056 \\ 0.6872893 & 0 & 0.3916056 & 1 \end{bmatrix},$$

$$[P_{ij}^*] = \begin{bmatrix} 1 & 0.5697828 & 0 & 0.6872893 \\ 0.5697828 & 1 & 0.8290291 & 0 \\ 0 & 0.8290291 & 1 & 0.3916056 \\ 0.6872893 & 0 & 0.3916056 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$[P_{ij}^*] = \begin{bmatrix} 0 & 0.5697828 & 0 & 0.6872893 \\ 0.5697828 & 0 & 0.8290291 & 0 \\ 0 & 0.8290291 & 0 & 0.3916056 \\ 0.6872893 & 0 & 0.3916056 & 0 \end{bmatrix}.$$

The column normalized proximity matrix $[\widehat{p_{ij}}]$ at time $t = 0$ is given by:

$$[\widehat{p_{ij}}] = \begin{bmatrix} 0 & 0.4073334 & 0.0000000 & 0.6370308 \\ 0.4532618 & 0 & 0.6791787 & 0.0000000 \\ 0.0000000 & 0.5926666 & 0 & 0.3629692 \\ 0.5467382 & 0.0000000 & 0.3208213 & 0 \end{bmatrix}.$$

If at time t_1 component B fails, and its two neighbours A and C are working then component A and C will take up 40.7 and 59.3 percent of component B's load respectively. However, if at time t_2 component C fails while component A still works, then $\widehat{p_{AB}}$ is recalculated as 1 because $\theta_{CB} = 0$ at time t_2.

4.3 A simple example: derivation of load function for a multi-component system

We use a multi-component system to illustrate the derivation of the load functions shown in Eq.(4.4). Let us consider a system that consists of n components connected in parallel. Without generality, consider the case that the components share the constant system load L in the proportion $\gamma_1, \gamma_2, \ldots, \gamma_n$, where γ_i is the proportion taken by component i at time $t = 0$. From Eq.(4.1) the system load is derived as:

$$L = \frac{\sum_{i=1}^{n} \gamma_i}{\gamma_i} z_i, \tag{4.11}$$

where z_i is the load taken by component i at initial time $t = 0$. Let $\gamma_i = \gamma_j$ for any $1 \leq i, j \leq n$ then for $t \geq 0$ the system load can be rewritten as:

$$L = \sum_{i=1}^{n-j} z_{i(j)} = z_{i(j)} \times (n - j), \tag{4.12}$$

where i denotes a working component, and j denotes the number of failed components in the system. $z_{i(j)}$ is the load of the i^{th} working component after j number of components have failed. Eq.(4.12) is the system load when we consider that the components equally share the system load. When $j = 0$ Eq.(4.12) and Eq.(4.11) are equal.

If we consider that an increased load on each working component i for $i = 1, 2, \ldots, n-j$ is a function of both its own independent load z_i and the load of a failed component z_j then $z_{i(j)}$ can be written as:

$$z_{i(j)} = z_i + l_j, \tag{4.13}$$

where l_j is the proportion of the load of the j^{th} component taken up by component i. l_j is given by:

$$l_j = \frac{\gamma_i}{\sum_{i \neq j} \gamma_i} z_j. \tag{4.14}$$

When the first component failure occurs, $z_{i(j)}$ equals:

$$z_{i(1)} = z_i + l_1.$$

When the second component failure occurs, $z_{i(j)}$ will become:

$$z_{i(2)} = z_{i(1)} + l_2 = z_i + l_1 + l_2.$$

After the j^{th} component failure has occurred, $z_{i(j)}$ can be written as:

$$\begin{aligned} z_{i(j)} &= z_{i(j-1)} + l_j = z_i + \textstyle\sum_j l_j \\ &= z_i + \tfrac{\gamma_i}{\sum_{i \neq j} \gamma_i} \textstyle\sum_j z_j. \end{aligned} \tag{4.15}$$

Let $e_{ij} = \frac{\gamma_i}{\sum_{i \neq j} \gamma_i}$ be the proportion of the load of component j taken by component i. From Eq.(4.15), e_{ij} does not have a spatial element. However, consider the load-sharing rule such that an increased load on a component is a function of both its own independent load z_i and the load of its influencing proximate component z_j. To introduce spatial dependence between components in $z_{i(j)}$ we set $e_{ij} = p_{ij}$ if distance information is known otherwise $e_{ij} = \vartheta_{ij}$.

If we consider that $\gamma_i = \gamma_j$ for any $1 \leq i, j \leq n$ in Eq.(4.1) then $z_i = z_j = z_0$ and we derive our expression for the load function where equal load is allocated to all components at time $t = 0$. This is written as:

$$z_{i(j)} = z_0 + \sum_{j \neq i} e_{ij} z_0.$$

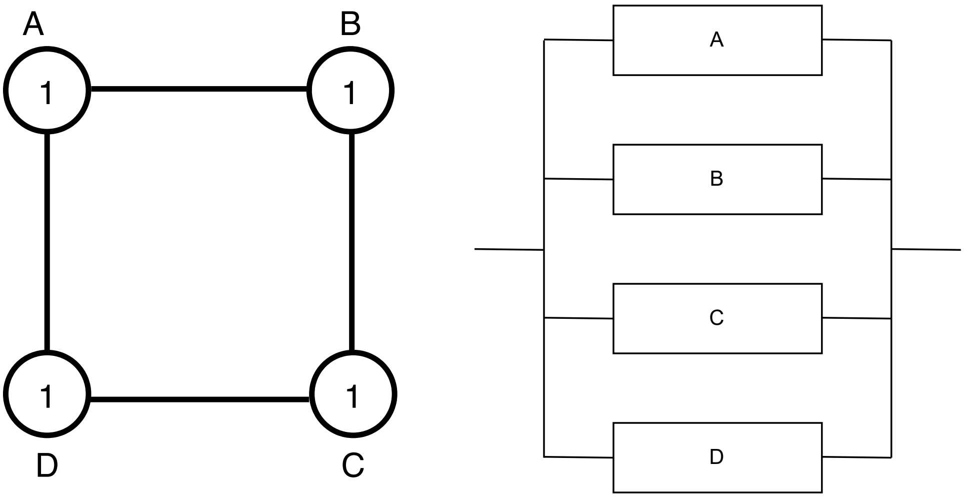
4.4 Formulation of system state transition

4.4.1 Failure rate function for homogeneous components

When a component fails, the load of the failed component is added to the proximate components and causes an increase in their failure rates. In this chapter, the relation between component i and the load of its failed proximate neighbour j is defined as follows:

$$\lambda_{i(j^r)} = \begin{cases} \lambda_0, & \text{if } j = 0, \\ \lambda_0 (z_{i(j)})^\beta = \lambda_0 (z_0 + \vartheta_{ij} z_0)^\beta, & j \neq 0, \end{cases} \tag{4.16}$$

where r is the number of working neighbours of component j that shares its load. z_0 represents the load imposed on components i and j when $z_i = z_j$. λ_0 is the baseline failure rate (also called hazard rate) function. Whenever a working neighbour of component j fails after component j has failed, the load of component j is redistributed to the remaining working neighbours and r changes. We assume that the initial load im-



(a) Spatial connections - four links (b) Block diagram of parallel connections

Figure 4.2: The four component system structure

posed on each working component i cannot be less than 1. Otherwise, the load shared between components will have a reducing effect on a working component's failure rate. β is the load factor which describes the influence of the increased load on the failure rate of a working component. When $\beta = 0$, an increased load has no impact on a component's failure rate. The higher β is, the higher the impact of load on the working neighbours; thus, the working neighbours have higher failure rates. At time $t = 0$, we assume that all the components have the same initial failure rate given as $\lambda_i = \lambda_0$. Obviously, we have $\lambda_i \leq \lambda_{i(j^r)}$.

To illustrate the load-sharing concept, let us consider a four-component system whose components are indexed by A, B, C, and D as seen in Figure 4.2. Using the spatial dependence concept introduced in section 4.2 component pairs (A, B), (A, D), (B, C), and (C, D) are spatially dependent while (A, C) and (B, D) are independent as depicted in Figure 4.2a. If component B first failed at time t_1 then its two working neighbours component A and C will have increased loads of $z_{A(B)}$ and $z_{C(B)}$ given by:

$$z_{A(B)} = z_0 + \vartheta_{AB} z_0,$$

and

$$z_{C(B)} = z_0 + \vartheta_{CB} z_0.$$

Their failure rate will become $\lambda_{A(B^2)} = (z_0 + \vartheta_{AB} z_0)^{\beta} \lambda_0$ and $\lambda_{C(B^2)} = (z_0 + \vartheta_{CB} z_0)^{\beta} \lambda_0$ respectively.

On the other hand, component D will remain as z_0 with failure rate λ_D since it is not spatially dependent on component B.

If one of components A or C fails at time t_2 given that component B has failed, the dependence matrix is updated and the only working neighbour of B will take all the load of B. The load and failure rate of the working component will be:

$$z_{i(B)} = z_0 + \vartheta_{i(B)} z_0,$$

and

$$\lambda_{i(B^1)} = (z_0 + \vartheta_{i(B)} z_0)^{\beta} \lambda_0,$$

where i indicates component A or C, and $\vartheta_{i(B)} = 1$.

However, if components A and C are still working and both have two neighbours in failed state, say component D has failed at time t_2 after component B failed, then components A and C will have increased load and failure rate functions of:

$$z_{i(B,D)} = z_{i(B)} + \vartheta_{iD} z_D = z_0 + \vartheta_{iB} z_0 + \vartheta_{iD} z_0,$$

and $\lambda_{i(B^2,D^2)} = (z_0 + z_0\vartheta_{iB} + z_0\vartheta_{iD})^{\beta} \lambda_0$, where i denotes node A and C.

On the other hand, if we assume that there is information about the distance d_{ij} between components. Then given the distance values d_{ij} between each pair of components one can derive proximity values p_{ij} and the failure rate in Eq.(4.16) will become:

$$\lambda_{i(j^r)} = (z_0 + p_{ij} z_0)^{\beta} \lambda_0.$$

4.4.2 System state transition with homogeneous components

In order to describe the behaviour of the load-sharing system with proximity and spatial dependence, we will apply the aforementioned proximity concept and failure rate model in a Markov model. The proximity model will be used to infer proximity between components, while the failure rate model model will be used to account for effect of component load on the transition rates in the Markov model. Markov models are frequently used in reliability and maintainability work where events, such as the failure or repair of a component, can occur at any point in time. The Markov model evaluates the probability of jumping from one known state into the next logical state. The process continues until the system being considered has reached the final failed state or until a particular mission time is achieved. The model assumes that the conditional distribution of a future state is independent of the past states of the process i.e., the behaviour of the system in each state is memoryless [145]. Thus, the sojourn time of each state is exponentially distributed and the transition probability to each state is independent of the process history. An advantage of Markov models is that they are simple to generate even though they require a complicated mathematical approach.

To apply the Markov model, we will construct a state diagram of load-sharing system representing all possible states of the system. The transition from one state to another will be specified by an arrow whose direction indicates the direction of transition. The failure rate expressions in section 4.1 will denote the rate parameter of the transition from one state to another.

Consider the four-component system in Figure 4.2. The components in the system are linked in a way that each component is connected to its two immediate neighbours but distant from the third component. The system fails whenever the sum of loads on the working components is less than the load placed on the system. Thus the system failure occurs after three components in the system has failed.

Consider that the system's state can be in any one of a discrete set of states $S_0, S_1, \ldots, S_4$ at time t. Let each system's state S_i for $i = 0, 1, \ldots, 4$ be described by the vector of its components' states. The definitions of state S_i are given as follows

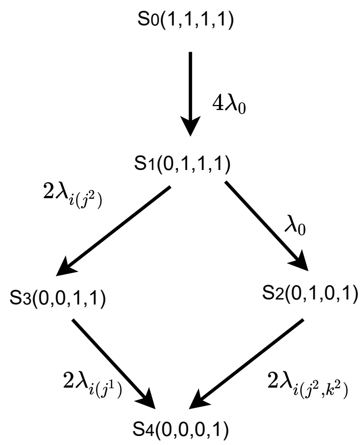


Figure 4.3: State transitions for the four-component system with homogeneous components

(see Figure 4.3). Let a vector of zeros and ones represent the states of components in a system such that a working component is denoted by 1 and a failed component is denoted by 0. If we let the vector also denote the arrangement of components in the system and assume that the order of components in the system is a function of the first component to fail. Therefore (1,1,1,1) will denote that all components in the system are working. If one component fails, say component A first fails, then (0,1,1,1) will denote that component A is in failed state while components B, C, and D are working. Otherwise if any other component failed first, say component C, then (0,1,1,1) will mean that the working components are D, A and B in that order. If two components fail, say components A and B, then (0,0,1,1) denotes that the two proximate components are in failed state. If two non spatially dependent components are in failed state, say components A and C, then (0,1,0,1) will denote failed state of the two components. (0,0,0,1) denotes that three spatially dependent components A, B and C are in failed state while component D is still working. (0,0,0,0) denotes that all components have failed. For other forms of system states, (0,1,0,1) and (1,0,1,0) indicates that two components are in failed state and there is a working component between the two failed components. (0,1,1,0) denotes that two proximate working components each have one failed neighbour.

The four-component system's evolution is determined by the transitions among states. State transitions of the system goes as follows. The transition from state (0,1,1,1) to state (0,0,1,1) means that one of the two working components that is spatially dependent on the failed component has now failed. The transition from state (0,1,1,1) to state (0,1,0,1) means that one working component which is not spatially dependent on the failed component has now failed. If a component has one failed spatial dependent neighbour, its transition rate will become $\lambda_{i(j^r)}$. So the transition from state (0,1,1,1) to state (0,0,1,1) will have a transition rate of $\lambda_{i(j^2)}$. For example, since component B is spatially dependent on component A then the transition from state (0,1,1,1) to (0,0,1,1) will have a transition rate of $\lambda_{i(j^2)}$ while the transition from state (0,1,1,1) to state (0,1,0,1) would have a transition rate of λ_0 because component C is not spatially dependent on component A. If we consider a transition from state (0,1,0,1) to state (0,0,0,1), it shows that the working component between two failed neighbours has now failed. Since the component has two failed neighbours, its transition rate will be $\lambda_{i(j^2,k^2)}$. Transition from state (0,1,1,0) to state (0,0,1,0) denotes that one of two working components with a failed neighbour has failed with transition rate $\lambda_{i(j^1)}$.

The states where the sum of loads on the working components are less than the load placed on the system (i.e. $\sum_{i=1}^{n} z_i < L$) are system failure states. For instance, if we consider transition from state (0,1,1,0) to state (0,0,1,0). The transition denotes that one of the last two working components in the system has failed. Thus, the last working component will take up all the load of its two proximate neighbour and the load of all other failed components will be lost. Due to the load distribution, the total load on the working component in state (0,0,1,0) will be less than the required and the system fails.

4.4.3 Failure rate function for heterogeneous components

This section considers the load-sharing system with heterogeneous components, which is very common in reality. An example of a system with non-identically distributed components is a tri-engine airplane which can operate when at least both of its wing engines or its central engine are working. The wing engines and the central engine are not necessarily identical [13]. A representative sample of such study is [248] that suggested a new model for load-sharing k-out-of-n: G system with different component. The model was developed as an extension of the capacity flow model. The model describes the increase of the component's failure rates under different loads. [250] extended the model by introducing common cause failure. An example can be observed in a load-sharing system with identical types of components, but having a mix of old (used) and new components. This section extends the work of [248] and [250] that both studied systems with equal load-sharing and heterogeneous components by considering spatial dependence and proximity effects in load-sharing systems. Thus we extend the load-sharing model developed in the previous sections by considering systems with heterogeneous components.

Assume that there is no information about the distance between components and one cannot estimate how close the components are to each other. Whilst other assumptions are maintained, we assume that the life times of the components are load dependant and follow the non-identical exponential distributions. In this case, at time $t = 0$ all components have the different initial failure rates given by:

$$\lambda_i = \lambda_i^*, \tag{4.17}$$

where $\lambda_i \neq \lambda_j$ for i and $j = 1, 2, \ldots, n$.

If we assume for example that all the components in the system take up varied load proportions z_i derived from Eq.(4.1), then when one component fails, its neighbours will have an increased failure rate of:

$$\lambda_{i(j^r)} = (z_{i(j)})^\beta \lambda_i^* = (z_i + \vartheta_{ij} z_j)^\beta \lambda_i^*, \tag{4.18}$$

where $\lambda_i < \lambda_i(j^r)$.

When the second neighbour of a working component fails, the working component will have an increased load of $\lambda_{i(j^r,k^r)}$ given by:

$$\lambda_{i(j^r,k^r)} = (z_{i(j)} + \vartheta_{ik} z_k)^\beta \lambda_i^*. \tag{4.19}$$

4.4.4 System state transition with heterogeneous components

In this section, we will formulate a method of accounting for the states of the system with heterogeneous components using Markov model. We will then integrate the formulated failure rates into the Markov model. To illustrate the method, we will consider the four-component system in Figure 4.2 with heterogeneous components indexed as A, B, C, and D. The definition of each state of the system will be as follows (see Figure 4.4). If any one component in the system is in failed state, say component B, then components A and C will have increased failure rates of $\lambda_{A(B^2)}$ and $\lambda_{C(B^2)}$ respectively while component D remains the same as λ_i since it is not spatially dependent on component B. If component A has two spatial dependent components in failed state, say component B and D, then component A and C will have increased failure rates of $\lambda_{A(B^2,D^2)}$ and $\lambda_{C(B^2,D^2)}$ respectively.

The four-component system's evolution will be determined by the transitions among states. When all components are working, we indicate this initial state by 0 meaning that zero components have failed. When one component fails we denote this state by 1. The procedure is continued until the system fails.

The transition rate from 0 to 1 when one component has failed is $\lambda_A+\lambda_B+\lambda_C+\lambda_D$ implying that anyone of the four components could fail. The transition rate of one possible path from 1 to 2 is $2\lambda_{i(j^2)}+\lambda_i$. For example, using the four-component system, if component B first failed, then anyone of component A, C and D could be the next to fail with transition rate $\lambda_{A(B^2)}+\lambda_D+\lambda_{C(B^2)}$. Likewise, if we assume that component C is the second component to have failed, then anyone of component A, or D could be the next to fail with transition rate $\lambda_{A(B^1)}+\lambda_{D(C^1)}$. Here component A is not affected by the failed state of component C, likewise component D is also not affected by the failed state of component B. In contrast, component A and D are affected by the failed state of components B and C in a way that A and D take up 100 percent of their load. When either component A or D fails, the system fails because the sum of loads on the working component will be less than the load placed on the system. The process is applied to derive all possible state transition equations for the system.

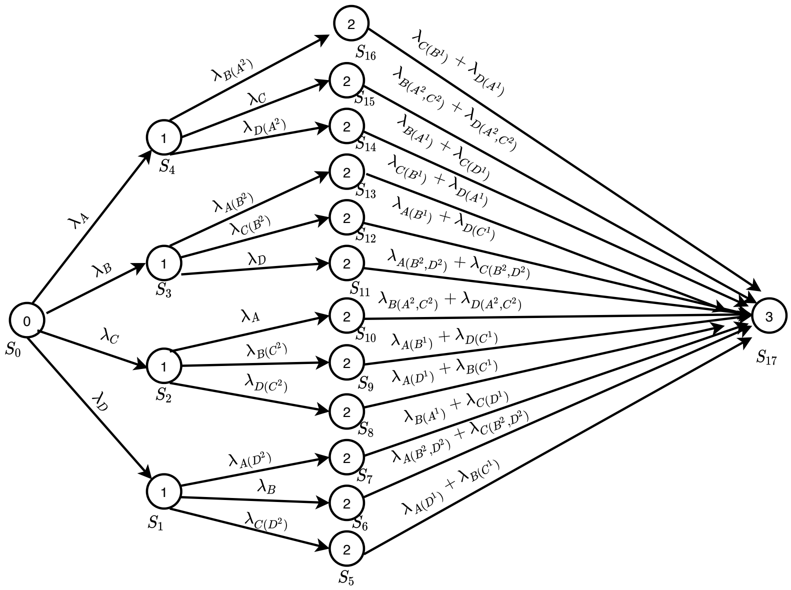


Figure 4.4: State transitions for the four-component system with heterogeneous components.

4.5 Parameter estimation

In this section, we extend the method introduced by [40] for equal load models to estimate parameters λ_0 and β. Each $\lambda_{i(j^r)}$ values in the system model can be calculated from the estimated values of λ_0 and β. The system load L, component spatial dependence matrix $[\boldsymbol{\theta_{ij}}]$, and initial load-sharing proportion γ_i are assumed to be estimated from operators of the system. Let us consider p load-sharing parallel systems each having n components that are put on test and their failure times are recorded. Let $T_{si}(s = 1, 2, ..., p; i = 1, 2, ..., n)$ be the failure time of the i^{th} component in the s^{th} parallel system. Let $\lambda_s^{(J)}$ be the system failure rate given that J number of components have failed. J is the sum of all components j that has failed in the system. Note that J counts the number of failed components while j is an index of a failed component. When all components are working, $J = 0$ and

$$\lambda_s^{(0)} = \sum_{i=1}^{n} \lambda_i. \tag{4.20}$$

After the failure of the J^{th} component, the failure rates of $n-J$ working components will be a mixture of $\lambda_{i(j^r)}$ and λ_i. The system failure rate after the J^{th} failure becomes $\lambda_s^{(J)}$ which is the sum of $\lambda_{i(j^r)}$ and λ_i for $n-J$ remaining components. Let $\sigma_J = (z_i + \vartheta_{ij} z_j)^\beta$ such that after the J^{th} failure, $\lambda_{i(j^r)} = \sigma_J \lambda_i$.

To illustrate the estimation method, let us consider the four-component system in Figure 4.2 with homogeneous components. If all the four components are working, then any of the components can fail with failure rate $\lambda_i = \lambda_0$, and $\lambda_s^{(0)} = 4\lambda_0$.

After the first component failure has occurred, the failure rates of the three remaining components becomes $\lambda_{i(j^2)}$ for two components and λ_0 for the third component. $\lambda_{i(j^2)} = \sigma_1 \lambda_0$. Thus, the system failure rate becomes $\lambda_s^{(1)} = 2\sigma_1\lambda_0 + \lambda_0$. Note that $\sigma_0 = 1$.

After the second component failure has occurred, the failure rates of the two remaining components becomes $2\sigma_2\lambda_0$. $\sigma_2 = (z_0 + z_0\vartheta_{i1} + z_0\vartheta_{i2})^\beta$. The system failure rate becomes $\lambda_s^{(2)} = \lambda_{i(j^2)} + \lambda_{i_2(j^2)} + \lambda_0$.

After the second component failure has occurred, the failure rates of the two remaining components becomes $2\sigma_2\lambda_0$. $\sigma_2 = (z_0 + z_0\vartheta_{i1} + z_0\vartheta_{i2})^\beta$. The system failure rate becomes $\lambda_s^{(2)} = 2\sigma_2\lambda_0$. The system fails when one of the two components fails because the sum of loads on the last working component will be less than the load placed on the system.

The likelihood function for the s^{th} load-share parallel system model is:

$$
\begin{aligned}
L_s(\lambda_0; t_{s1}, t_{s2}, t_{s3}) = &\left(4\lambda_0 exp^{(-4\lambda_0 t_{s1})}\right)\left((2\sigma_1+1)\lambda_0 exp^{(-(2\sigma_1+1)\lambda_0 t_{s2})}\right) \times \dots \\
&\left(2\sigma_2\lambda_0 exp^{(-2\sigma_2\lambda_0 t_{s3})}\right),
\end{aligned}
\tag{4.21}
$$

$$
L_s(\lambda_0; t_{s1}, t_{s2}, t_{s3}) = 8\left((2\sigma_1+1)\sigma_2\right)\lambda_0^3 exp^{-(\lambda_0[4t_{s1}+(2\sigma_1+1)t_{s2}+2\sigma_2 t_{s3}])}. \tag{4.22}
$$

The likelihood function corresponding to p systems is:

$$L(T \mid \lambda_0, \sigma_1, \sigma_2) = 8^p \left((2\sigma_1 + 1)^p \sigma_2^p\right) \lambda_0^{3p} exp^{-\left(\lambda_0 \sum_{s=1}^{p} [4t_{s1} + (2\sigma_1+1)t_{s2} + 2\sigma_2 t_{s3}]\right)}. \tag{4.23}$$

The log likelihood function is:

$$\begin{aligned} l = log(L) = plog(8) + 3plog(\lambda_0) + plog(2\sigma_1 + 1) + plog(\sigma_2) \ldots \\ - \left(\lambda_0 \sum_{s=1}^{p} [4t_{s1} + (2\sigma_1 + 1)t_{s2} + 2\sigma_2 t_{s3}]\right). \end{aligned} \tag{4.24}$$

Taking the derivative of the log likelihood function with respect to λ_0, σ_1, and σ_2 leads to:

$$\frac{\delta l}{\delta \lambda_0} = \frac{3p}{\lambda_0} - \sum_{s=1}^{p} [4t_{s1} + (2\sigma_1 + 1)t_{s2} + 2\sigma_2 t_{s3}] = 0, \tag{4.25}$$

$$\frac{\delta l}{\delta \sigma_1} = \frac{2p}{(2\sigma_1 + 1)} - \lambda_0 \sum_{s=1}^{p} 2t_{s2} = 0, \tag{4.26}$$

$$\frac{\delta l}{\delta \sigma_2} = \frac{p}{\sigma_2} - \lambda_0 \sum_{s=1}^{p} 2t_{s3} = 0. \tag{4.27}$$

From Eq.(4.25), (4.26), and (4.27), one cannot obtain a closed form solution for λ_0, σ_1, and σ_2. However, from Eq.(4.25), we have:

$$\hat{\lambda_0} = \frac{3p}{\sum_{s=1}^{p} [4t_{s1} + (2\sigma_1 + 1)t_{s2} + 2\sigma_2 t_{s3}]}. \tag{4.28}$$

Thus, on using Eq. 4.26, 4.27, and 4.28, one gets:

$$\hat{\sigma_1} = \frac{2p}{(2\sigma_1 + 1)} - \frac{3p \sum_{s=1}^{p} 2t_{s2}}{\sum_{s=1}^{p} [4t_{s1} + (2\sigma_1 + 1)t_{s2} + 2\sigma_2 t_{s3}]}, \tag{4.29}$$

$$\hat{\sigma_2} = \frac{p}{\sigma_2} - \frac{3p \sum_{s=1}^{p} 2t_{s3}}{\sum_{s=1}^{p} [4t_{s1} + (2\sigma_1 + 1)t_{s2} + 2\sigma_2 t_{s3}]}. \tag{4.30}$$

Eq.(4.26), and (4.27) can be solved for $\hat{\sigma_1}$ and $\hat{\sigma_2}$ by using some suitable iterative procedure

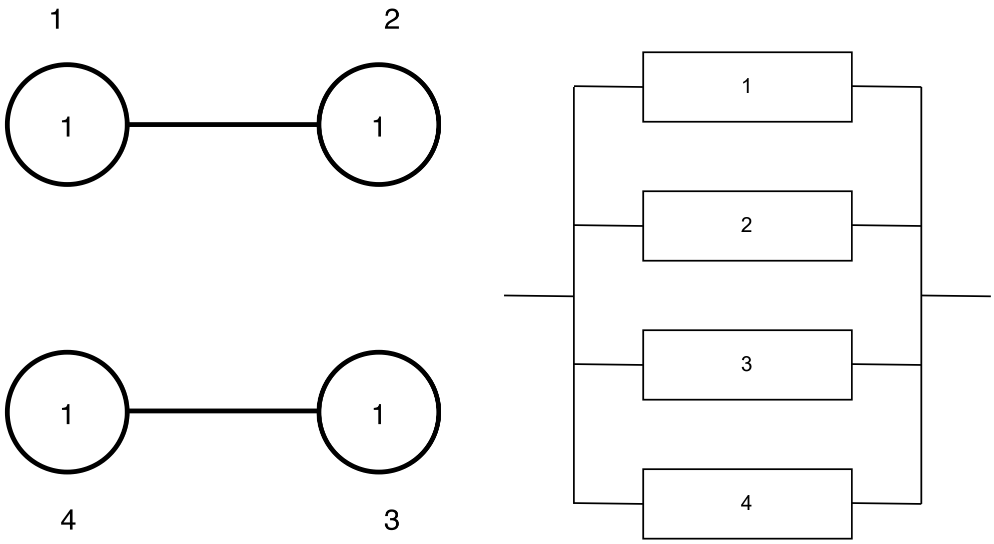
After each σ_J has been estimated, β will be derived as:

$$\beta = \frac{log(\sigma_J)}{log(z_i + \vartheta_{ij} z_j)}, \tag{4.31}$$

where z_i, ϑ_{ij}, and z_j are known. Based on the method introduced by Kim and Kvam [40], the first, second, ..., $J+1^{th}$ failure times are needed to estimate λ_0, σ_1, $\ldots \sigma_J$ respectively. However, for systems with large number of components, the challenge lies in the exponential explosion of the large number of states and the number of σ_J parameters to be estimated. Given that β can be calculated from any σ_J using Eq.(4.31), and other σ_J values can be calculated using the estimated λ_0 and β, we will estimate only σ_1 in this thesis while other σ_J's will be ignored. Furthermore, we will estimate λ_0 and σ_1, using the first and second component failure times of any system studied.

4.6 Numerical example 1: a simple four-component system

In this section we use a four-component system with a simple structure to illustrate the developed model. The purpose is to highlight the importance of spatial effect and demonstrate the proposed spatial model through numerical example. The structure of the system is depicted in Figure 4.5. Note that the components are connected in parallel as seen in Figure 4.5b, while the links between the components indicate spatial dependence. The system has four components indexed as component 1, 2, 3, and 4. From Figure 4.5a, we note that the components are linked together in a way that the system is composed of two subsystems (that is, subsystem having components 1 and 2 and another subsystem having components 3 and 4). The system fails whenever the sum of loads on the working components is less than the load placed on the system. When one component fails, its immediate working neighbour takes up all the load so that the system still works. If a third failure occurs, two components in one of the subsystems must have failed, hence the system fails. We assume that there is no information about



(a) Spatial connections - two links (b) Block diagram of parallel connections

Figure 4.5: The four-component systems

the distance between components in the system so that the component arrangement will be used to infer component proximity. In the analysis, we will consider two systems; one with homogeneous components and another with heterogeneous components. We assume that the parameters to be used for the numerical analysis are known. The parameters for this numerical analysis are taken from [250].

We examine the effect of the load factor on the reliability estimations of a load-sharing system with spatial dependence and proximity effect. We also examine the system's reliability at different performance levels. Consider a system composed of homogeneous components with equal load allocations at time $t = 0$. Since we aim to highlight the importance of spatial effect and demonstrate the proposed spatial model through numerical example, we do not assume any unit for time. The state transitions are presented as Figure 4.6. The system failed or "down" states are S_3 and S_4 while the system "up" states are S_0, S_1, and S_2. Let $Q_i(t) = \Pr(\text{System is in state } S_i \text{ at time } t)$. In order for the system to be in state S_0 at time $t+\Delta t$, the system must be in state S_0 at time t, and no transition occurs from state S_0 in time $(t, t+\Delta t)$. Thus, we have:

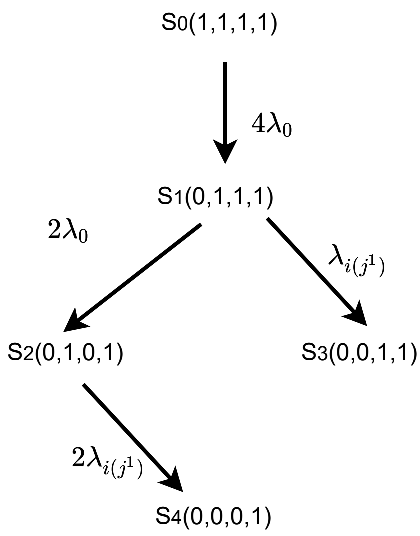


Figure 4.6: State transitions for the four-component system with homogeneous components.

$$\begin{aligned} &Q_0(t+\Delta t) = Q_0(t)(1 - 4\lambda_0 \Delta t) \\ &\lim_{\Delta t \to 0} \tfrac{Q_0(t+\Delta t) - Q_0(t)}{\Delta t} = -4\lambda_0 Q_0(t) \\ &\tfrac{d}{dt} Q_0(t) = Q_0'(t) = -4\lambda_0 Q_0(t). \end{aligned} \tag{4.32}$$

Likewise for the system to be in state S_1 at time $t + \Delta t$, either the system is in state S_1 at time t and no transition occurs during $(t, t+\Delta t)$ or the system is in state S_0 at time t and a transition $S_0 \to S_1$ occurs in $(t, t+\Delta t)$. Thus, we have:

$$\begin{aligned} &Q_1(t+\Delta t) = Q_1(t)(1 - (2\lambda_0 + \lambda_{i(j^1)})\Delta t) + Q_0(t) 4\lambda_0 \Delta t \\ &\lim_{\Delta t \to 0} \tfrac{Q_1(t+\Delta t) - Q_1(t)}{\Delta t} = -(2\lambda_0 + \lambda_{i(j^1)}) Q_1(t) + 4\lambda_0 Q_0(t) \\ &Q_1'(t) = -(2\lambda_0 + \lambda_{i(j^1)}) Q_1(t) + 4\lambda_0 Q_0(t). \end{aligned} \tag{4.33}$$

Using the same logic, we can derive the corresponding differential equation for $Q_2'(t)$ as follows:

$$Q_2'(t) = -2\lambda_{i(j^1)} Q_2(t) + 2\lambda_0 Q_1(t), \tag{4.34}$$

$Q_3(t)$ and $Q_4(t)$ are not needed for calculating the system reliability because they are the probabilities of the system being in failed states S_3 and S_4.

In contrast, if we assume that the components in the system are heterogeneous (non-identically distributed), then the state transitions can be derived as shown in Figure 4.7. The system failed or "down" states are S_5, S_8, S_{13}, S_{16}, and S_{17} while the system "up" states are S_0, S_1, S_2, S_3, S_4, S_6, S_7, S_9, S_{10}, S_{11}, S_{12}, S_{14}, S_{15}. The set of differential equations derived for the up states are given as:

$$\begin{aligned}
Q'_0(t) &= -(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)Q_0(t) \\
Q'_1(t) &= -(\lambda_1 + \lambda_2 + \lambda_{3(4^1)})Q_1(t) + \lambda_4 Q_0(t) \\
Q'_2(t) &= -(\lambda_1 + \lambda_2 + \lambda_{4(3^1)})Q_2(t) + \lambda_3 Q_0(t) \\
Q'_3(t) &= -(\lambda_{1(2^1)} + \lambda_3 + \lambda_4)Q_3(t) + \lambda_2 Q_0(t) \\
Q'_4(t) &= -(\lambda_{2(1^1)} + \lambda_3 + \lambda_4)Q_4(t) + \lambda_1 Q_0(t) \\
Q'_6(t) &= -(\lambda_{1(2^1)} + \lambda_{3(4^1)})Q_6(t) + \lambda_2 Q_1(t) \\
Q'_7(t) &= -(\lambda_{2(1^1)} + \lambda_{3(4^1)})Q_7(t) + \lambda_1 Q_1(t) \\
Q'_9(t) &= -(\lambda_{1(2^1)} + \lambda_{4(3^1)})Q_9(t) + \lambda_2 Q_2(t) \\
Q'_{10}(t) &= -(\lambda_{2(1^1)} + \lambda_{4(3^1)})Q_{10}(t) + \lambda_1 Q_2(t) \\
Q'_{11}(t) &= -(\lambda_{1(2^1)} + \lambda_{3(4^1)})Q_{11}(t) + \lambda_4 Q_3(t) \\
Q'_{12}(t) &= -(\lambda_{1(2^1)} + \lambda_{4(3^1)})Q_{12}(t) + \lambda_3 Q_3(t) \\
Q'_{14}(t) &= -(\lambda_{2(1^1)} + \lambda_{3(4^1)})Q_{14}(t) + \lambda_4 Q_4(t) \\
Q'_{15}(t) &= -(\lambda_{2(1^1)} + \lambda_{4(3^1)})Q_{15}(t) + \lambda_3 Q_4(t).
\end{aligned} \tag{4.35}$$

We will use the derived equations to estimate the system reliability for both the homogeneous and heterogeneous systems.

The failure rate of the component depends on the load function and proximity effect of the components. To derive the proximity effect of a failed component on the working components, the column normalized values of the working components are derived. Therefore, the spatial dependence matrix of components in the system is given as matrix $[\boldsymbol{\theta_{ij}}]$:

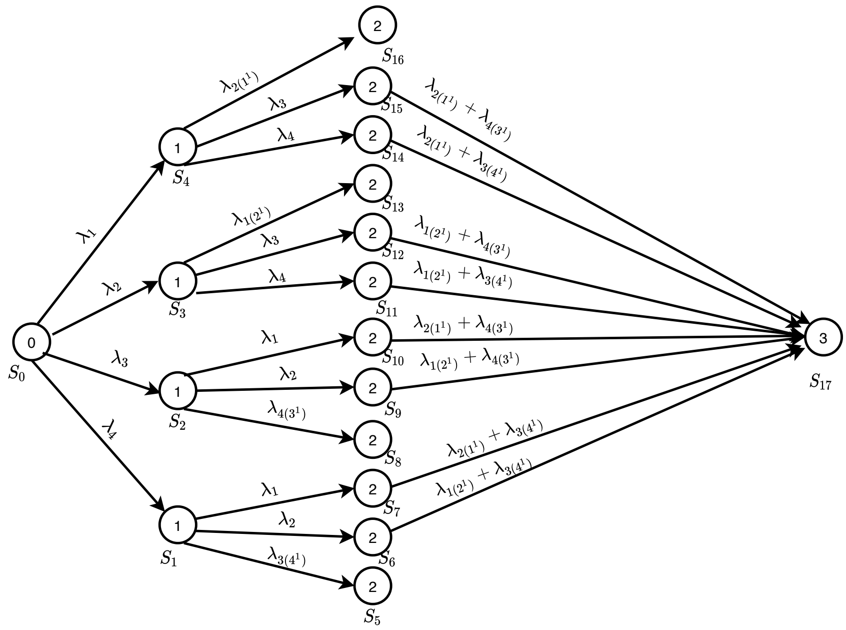


Figure 4.7: State transitions for the four-component system with Heterogeneous components.

$$[\boldsymbol{\theta_{ij}}] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The corresponding proximity matrix $[\boldsymbol{\vartheta_{ij}}]$ for components at time $t = 0$ is given by:

$$[\boldsymbol{\vartheta_{ij}}] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Consider the failure rates for the components if they are homogeneous. When all components are working, their failure rate is $\lambda_i = \lambda_0$ for $i = 1, 2, 3, 4$. Using the proximity effects of ϑ_{ij} derived in the column normalized matrix, when one component in the system has failed, the failure rate of the neighbour in the same subsystem becomes:

$$\lambda_{i(j^1)} = (z_0 + \vartheta_{ij} \times z_0)^{\beta} \, \lambda_0,$$

where $\vartheta_{ij} = 1$ for $i = 1, 2, \ldots, 4$, and where $z_0 = \frac{\gamma_i L}{\sum_{i=1}^{4} \gamma_i} = L/n$ where $\gamma_i = \gamma_j$ for any $1 \leq i, j \leq n$. If the working components are non-neighbours of the failure component their failure rate would remain as λ_0.

For the system with heterogeneous components, when all four components are in a working state, their failure rate is λ_i^* where $i = 1, 2, 3, 4$ and $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_4$. When one component has failed, the failure rate of its spatial neighbour is given by one of the following failure rates:

$$\begin{aligned}
\lambda_{1(2^1)} &= (z_1 + \vartheta_{12} z_2)^{\beta} \, \lambda_1^* \\
\lambda_{2(1^1)} &= (z_2 + \vartheta_{21} z_1)^{\beta} \, \lambda_2^* \\
\lambda_{3(4^1)} &= (z_3 + \vartheta_{34} z_4)^{\beta} \, \lambda_3^* \\
\lambda_{4(3^1)} &= (z_4 + \vartheta_{43} z_3)^{\beta} \, \lambda_4^*,
\end{aligned}$$

where $z_i = \frac{\gamma_i L}{\sum_{i=1}^{4} \gamma_i}$ and $\gamma_i \neq \gamma_j$ for any $1 \leq i, j \leq n$.

From this, we can estimate the reliability of the system for varied values of β. Let $R(t)$ be the probability that the system is functioning at time t, the reliability of the system is given by:

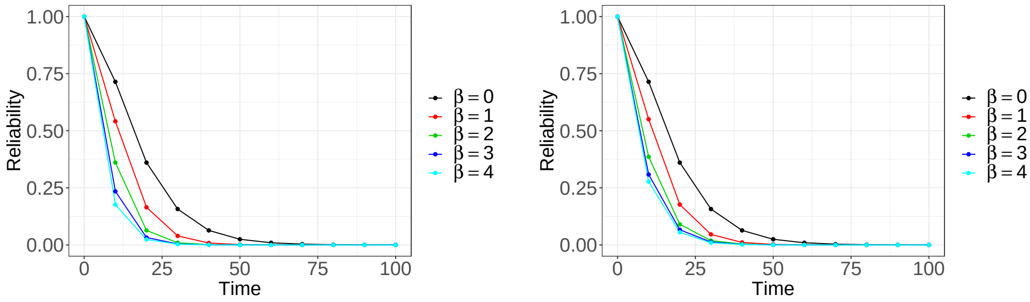
$$R(t) = \sum_{i=0}^{s-h} Q_i(t), \tag{4.36}$$

where s is the number of system states, h is the number of system down states and $Q_i(t)$ are the state probabilities for the system up states. The state probabilities $Q_i(t)$ can be computed by solving the Kolmogorov equations (i.e., the set of differential equations) expressed in Eq.(4.32), (4.33), (4.34), and (4.35). Let the initial conditions be set as $Q_0(0) = 1$ and $Q_i(0) = 0$ for $i = 1, 2, \ldots, s$. Calculating the system's reliability could be cumbersome depending on the method applied and the number of components considered. The main challenge lies in the exponential explosion of number of states. The large number of states makes it difficult to calculate system reliability. Appropriate methods should be used to ease the computational burden for a large-scale system. Analytical methods such as Laplace transform can be applied to derive the system's reliability function from the Kolmogorov equation. Laplace transform converts differential equations to algebraic forms by means of a transformation defined as $L_{(s)} = \int_0^\infty exp^{(-st)} f(t)dt$ [299]. The algebraic equations may then be easily solved, the inverse of the transformation applied to obtain the final solution [299]. However, the derivation of the analytic expression of the reliability function could be time consuming and cumbersome when large number of states are involved (see appendix A of chapter 4 for a comparison of reliability estimations using Laplace and Euler method). On another hand, numerical methods such as the modified Euler's method can be used to easily derive the system state probabilities and corresponding reliability estimate. The Euler method is a first-order numerical procedure for solving ordinary differential equations (ODEs) with a given initial value. Euler's method uses the simple formula, $y(x+h) = y(x) + hf(x, y)$ to construct the tangent at the point x and obtain the value of $y(x+h)$, whose slope is $\frac{dy}{dx}$ [300]. In Euler's method, one can approximate the curve of the solution by the tangent in each interval (that is, by a sequence of short line segments), at steps of h [300]. [301] and [302] used a numerical method to solve a system of differential-equations and demonstrated that the approximate results obtained by numerical method match considerably well with the results obtained by laplace transform technique. We will implement the method using deSolve package in R. Table 4.1 presents the parameters used to study the effects of β where β values will be made to vary between 0 to 4 by steps of 1.

Table 4.1: Parameters used to study the effect of β

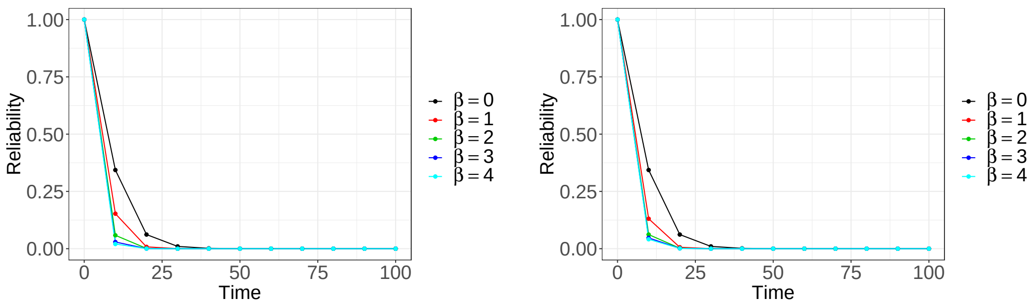
	β	λ_0	λ_1^*	λ_2^*	λ_3^*	λ_4^*	L	γ_1	γ_2	γ_3	γ_4
Homogeneous	$(0, 4)$	0.05	-	-	-	-	4	0.5	1	1.5	2
Heterogeneous	$(0, 4)$	-	0.05	0.10	0.15	0.12	4	0.5	1	1.5	2

Figures 4.8 and 4.9 show the variation of system reliability with respect to β. Figures 4.8a and 4.8b present two cases for systems with homogeneous components. Figure 4.8a represents reliability estimations when the system load is assumed to be equally distributed to the components at time $t = 0$. Figure 4.8b represents reliability estimations when the system load is not equally distributed. For the case of equal initial



(a) Components share equal initial load (b) Components share different initial load

Figure 4.8: Variation of system reliability with load factor β considering homogeneous components



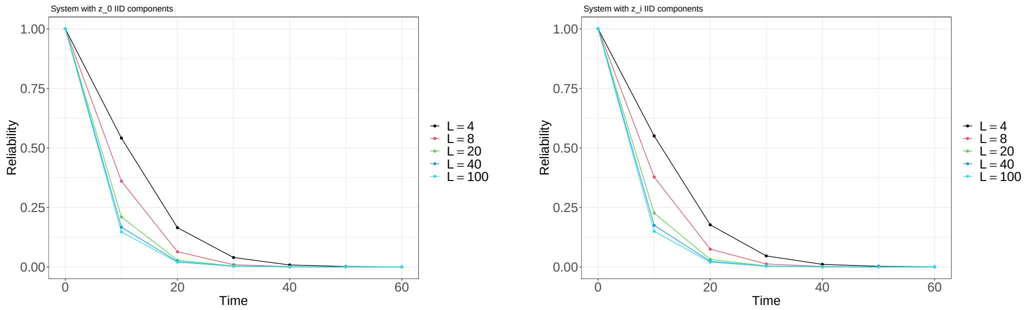
(a) Components share equal initial load (b) Components share different initial load

Figure 4.9: Variation of system reliability with load factor β considering heterogeneous components

load distribution, when $\beta = 0$, which indicates that the system is not load-sharing, the system reliability hits zero at time $t = 60$. However, when $\beta = 1$ indicating the presence of load-sharing, the system's reliability reduces. As β is increased, indicating that the impact of load-sharing between components is increased, the system's fails faster. For the case of unequal initial load distribution, a similar pattern is observed.

In Figure 4.9a and 4.9b we present systems with heterogeneous components. In both cases, we observe that $\beta = 0$, indicating that the system is not load-sharing, the system reliability hits zero at time $t = 27$. However, as the impact of load-sharing between components is increased, the system's reliability reduces sharply.

Next, we investigate the system reliability for different performance levels. We use the same parameters in Table 4.1 except that the system performance level (load L) will be made to vary. $L = 4, 8, 20, 40$, and 100 are the values used for the study. β will



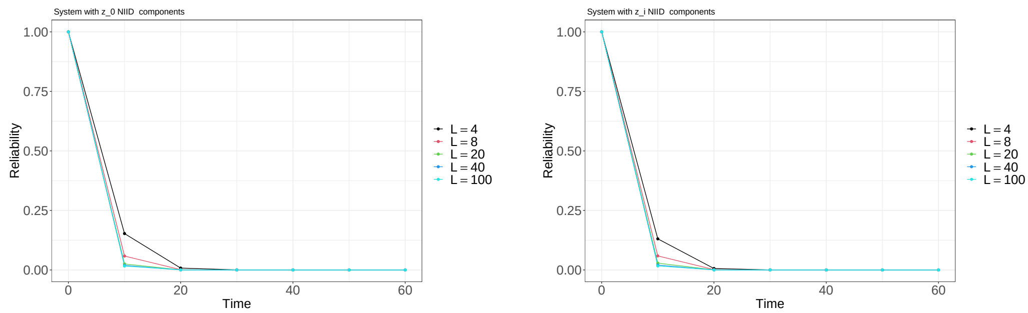
(a) Components share equal initial load (b) Components share different initial load

Figure 4.10: Variation of system reliability with performance level L considering homogeneous components

be set to 1 to indicate the presence of load-sharing in the system. Figures 4.10 and 4.11 show the variation of system reliability with respect to load L. Figures 4.10a and 4.10b present two cases for systems with homogeneous components. Figure 4.10a represents reliability estimations when the system load is assumed to be equally distributed to the components at time $t = 0$. Figure 4.10b represents reliability estimations when the system load is not equally distributed. For the case of equal initial load distribution, when $L = 4$, which indicates that the load placed on the system is low, the system reliability hits zero at time $t = 50$. However, when $L = 8$ indicating that an increased load is placed on the system, the system's reliability reduces. As L is increased, indicating that a heavy load is placed on the system's components, the system fails faster. For the case of unequal initial load distribution, a similar pattern is also observed.

Figure 4.11a and 4.11b presents systems with heterogeneous components. In both cases, when $L = 4$ and 8, we observe that when a light load is placed on the system, the system reliability hits zero at time $t = 20$. However, when the level of load on the system is increased beyond 20, the system's reliability reduces sharply to zero at time $t = 10$. Overall, when compared with the reliability of the system with homogeneous components, the reliability of a system with heterogeneous components decreases sharply over time.

We compare our developed models with existing models to investigate the importance and significance of the spatial effect. We compare our spatial model with capacity flow models developed for a k-out-of-n: G systems, in the setting of equal load-sharing



(a) Components share equal initial load (b) Components share different initial load

Figure 4.11: Variation of system reliability with performance level L considering heterogeneous components

with homogeneous components [248] and heterogeneous components [248, 250]. The four-component system fails when two components in the same subsystem fails, i.e. components 1 and 2, or components 3 and 4. This behaviour can be likened to those of k-out-of-n: G system in which the system works when k components work and the system fails when $n - k + 1$ components have failed. The case where the system described in Figure 4.5 fails due to the failure of two components in the same subsystem can be likened to 3-out-of-4: G system. Likewise, the case where the system fails when three components fail can be likened to 3-out-of-4: G system. As a result, we will compare reliability estimation from our spatial model with models of equal load-sharing for 2-out-of-4 and 3-out-of-4: G systems. We will examine the performance of the three models for reliability estimation of the four-component system with homogeneous and heterogeneous components. We use Monte Carlo simulation (MCS) to estimate the system reliability and compare with our models. The simulation model is used to compare the underlying data generating process with the modelling approaches. Using MCS, it is possible to modify the failure behaviour of the working components after one of the proximate components have failed. We denote the Monte Carlo simulation as Spatial simulation. The algorithm for generating the system lifetime T_{sys} is described as follows.

Algorithm 1 Algorithm for generating lifetime data of a system with load-sharing and spatial dependence

Require: Load factor β; component failure rate λ; system load L; number of components n; proximity matrix $[\boldsymbol{\vartheta_{ij}}]$.

Ensure: System lifetime T_{sys} at the end of the observation period.

start at $k = 1$.
repeat
  for all working components $i \in \{k, k+1, \ldots, n\}$ **do**
    calculate $t_i = \frac{-\ln(U_{0,1})}{\lambda_i}$ where $U_{0,1}$ is a random value drawn from a uniform distribution.
  end for
  let the kth component failure time $T_k = min(t_k, t_{k+1}, \ldots, t_n)$;
  calculate $z_{i(j)} = (z_i + \vartheta_{ij} z_j)$ for proximate neighbours of failed component j.
  replace λ_i by $\lambda_{i(j^r)} = (z_i + \vartheta_{ij} z_j)^\beta \lambda_i$ for proximate neighbours of failed component j.
  recalculate the proximity matrix $[\boldsymbol{\vartheta_{ij}}]$ by excluding the failed component.
  set $k = k + 1$.
until $\sum_{i=1}^{n-k} z_i < L$;
return $T_{sys} = \sum_k T_k$.

If one generates a large number of system lifetimes T_{sys}, the cumulative distribution function of the system can be evaluated. Using the above algorithm, 1000 system lifetimes T_{sys} are generated and used to evaluate the system's reliability $R_{sys}(t) = \frac{m(t)}{1{,}000}$ where $m(t)$ denotes the number of times the system survived beyond time t. Table 4.2 presents the parameters for the two forms of the four-component systems used for the comparison. Figure 4.12 and 4.13 present system reliability when homogeneous and heterogeneous components are considered respectively.

An example of a system where k-out-of-n system model maybe suitable for reliability analysis of systems with local load-sharing is when the system has n components and each component is a neighbour of all the remain n-1 components. Such system structure will likely have a regular shape.

In addition, Table 4.3 presents root mean squared error (RMSE) of reliability estimations from the three compared models and the Spatial simulation predictions. It can be observed that the spatial model's estimation error is way less that those by the models that assume equal load-sharing. In other words, the simulation results validate the results from the spatial model for heterogeneous and homogeneous components. It can be observed that when heterogeneous components are considered, the model for 3-out-of-4 systems had a better performance than the model for 2-out-of-4 systems. From Figure 4.13, it can be observed that compared with the model for 2-out-of-4 systems, the reliability estimations by the model for 3-out-of-4 systems closely matched those from the spatial model from $t = 0$ to $t = 10$. We also observe how close the results are based on the RMSE values indicating that the model for 3-out-of-4 systems could be suitable for reliability estimation of the four-component system when heterogeneous components are considered. When homogeneous components are considered as seen in Figure 4.12, it can be observed that from time $t = 5$ to $t = 20$ reliability estimations from the models for 2-out-of-4 and the model for 3-out-of-4 systems diverge from the spatial simulation. The result indicates that when homogeneous components are considered, neither the models for 2-out-of-4 nor the model for 3-out-of-4 system is suitable. Overall, for the four-component system studied in this section, when heterogeneous components are considered, the spatial effect may be ignored and a simpler model such as the model for 3-out-of-4 systems which does not account for spatial effects could be applied. However, when homogeneous components are considered for the four-component system, predictions by the models which assumes equal load-sharing for 2-out-of-4 and 3-out-of-4 systems either overestimate or underestimate the reliability of the four-component system.

Table 4.2: Parameters used for model comparison

	β	λ_0	λ_1^*	λ_2^*	λ_3^*	λ_4^*	L	γ_1	γ_2	γ_3	γ_4
Homogeneous	1	0.05	-	-	-	-	4	-	-	-	-
Heterogeneous	1	-	0.05	0.10	0.15	0.12	4	0.5	1	1.5	2

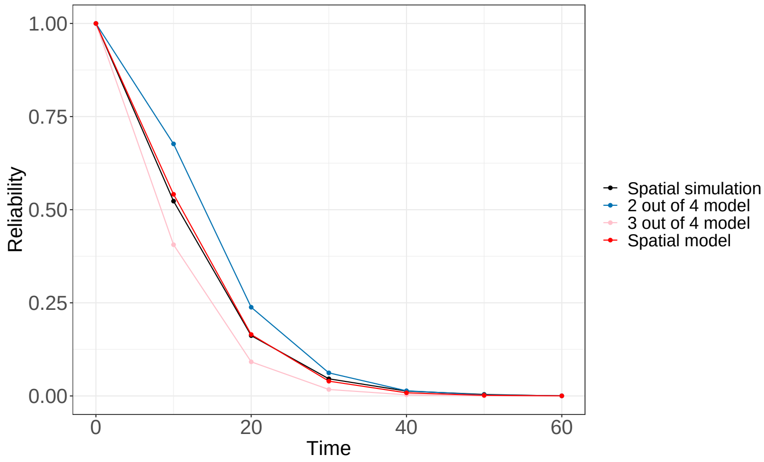


Figure 4.12: Comparison of reliability estimation using spatial simulation, the spatial model and existing models considering homogeneous components

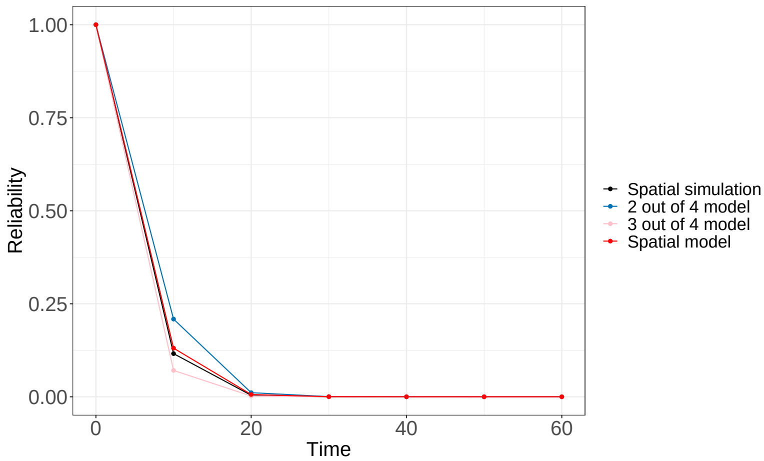
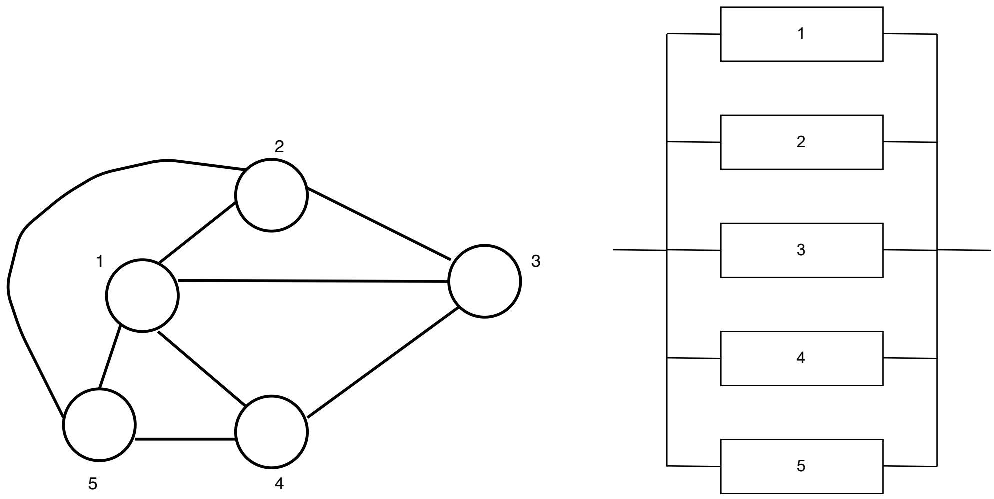


Figure 4.13: Comparison of reliability estimation using spatial simulation, the spatial model and existing models considering heterogeneous components

Table 4.3: Root mean squared error of the reliability estimations by the models

RMSE	**Homogeneous**		
	2-out-of-4 model	Spatial model	3-out-of-4 model
	0.0650	0.0076	0.0528
	Heterogeneous		
	2-out-of-4 model	Spatial model	3-out-of-4 model
	0.0351	0.0056	0.0170



(a) Spatial connections

(b) Block diagram of parallel connections

Figure 4.14: The five-component system

4.7 Numerical example 2: a five-component system with complex spatial structure

We extend the example in Section 5 to consider a more complex spatial structure with 5 components. For brevity, we only consider homogeneous components for the system depicted in Figure 4.14, however, the modelling could be easily extended for heterogeneous components. The components are linked together in a way that the system is composed of one dominant component (component 1) and four secondary components (component 2, 3, 4, and 5). As indicated by Figure 4.14a, the system fails when the dominant component and at least three secondary components fail. In this case the load on the only working component is less than the load placed on the system. We assume that there is no information about the distance between components in the system so that the component arrangement will be used to infer component proximity. We compare our spatial model to an equal load-sharing model for a 1-out-of-5 system.

Table 4.4 presents the parameters used in the model while the system state transition diagram is presented in Figure 4.15.

Table 4.5 presents RMSE from comparing the spatial model with the 1-out-of-5

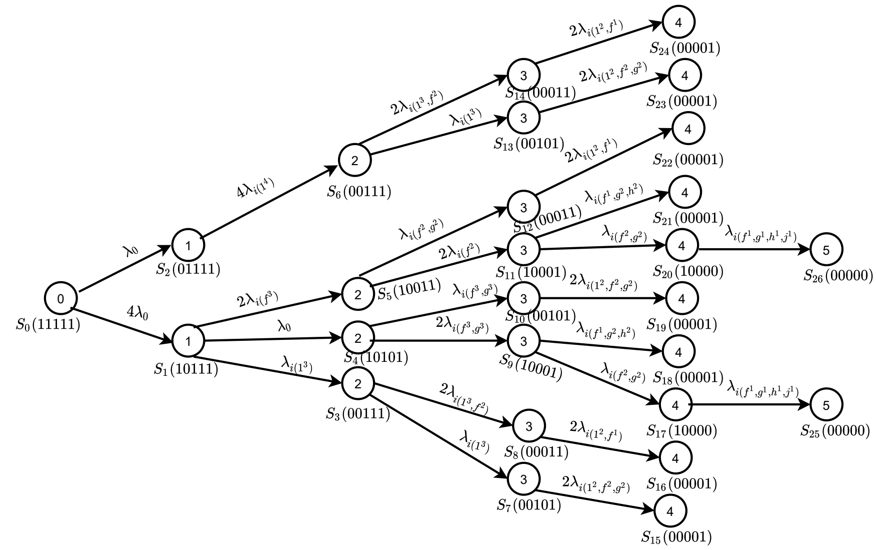


Figure 4.15: Five-component system state transition diagram.

Table 4.4: Parameters used for model comparison on the five-component system

	β	λ_0	L
Homogeneous	1	0.05	5

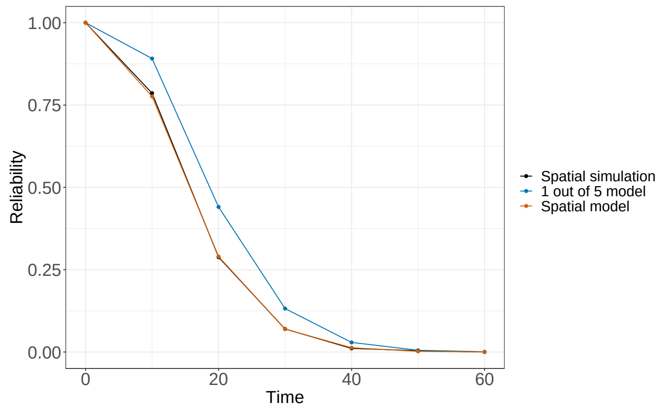


Figure 4.16: Comparison of reliability estimation using the spatial model and capacity flow model.

model. We observe that our spatial model's estimation error is less than those by the 1-out-of-5 model that assumes equal load-sharing. Figure 4.16 present the system reliability predictions by our spatial model and the 1-out-of-5 model where the predictions from the 1-out-of-5 model at time $t = 1$ and $t = 30$ is far from the Spatial simulation prediction compared to the spatial model whose prediction closely matches the Spatial simulation. Hence, it suggests that our spatial model is more accurate at modelling these more complex scenarios when homogeneous components are considered. Our results demonstrate that our model is more accurate than standard load-sharing models in evaluating system reliability when spatial effects exist.

Table 4.5: Root mean squared error of the reliability estimations by the models

RMSE	**Homogeneous**	
	1-out-of-5 model	Spatial model
	0.0741657	0.003929705

4.8 Numerical example 3: Parameter estimator

In this section, we assess the performance of the parameter estimation method developed in section 4.5 using simulated data. We use the four-component system studied in section 4.6. The components in the system are connected in parallel, while the links between the components indicate spatial dependence. The system has four components indexed as component 1, 2, 3, and 4. The components are linked together in a way that the system is composed of two subsystems (that is, subsystem having components 1 and 2 and another subsystem having components 3 and 4). The system fails whenever the sum of loads on the working components is less than the load placed on the system. When one component fails, its immediate working neighbour takes up all the load so that the system still works. If a third failure occurs, two components in one of the subsystems must have failed, hence the system fails. We assume that there is no information about the distance between components in the system so that the component arrangement is used to infer component proximity. In the analysis, we will consider that the system has homogeneous components. We assume that λ_0, and β are parameters to be estimated while the system load L, component spatial dependence matrix $[\boldsymbol{\theta_{ij}}]$, and initial load-sharing proportion γ_i are known. The performance of the estimators will be examined with respect to bias and mean squared error (MSE) in the estimates of λ_0, and β given some known input parameter values. The expression for the MSE is $MSE = Var(\hat{p}) + Bais^2(\hat{p})$. See Eq.(3.20) of section 3.3.2 for the expression of the Bias.

Spatial simulation described in section 4.6 is used to generate data in order to study the performance and asymptotic properties of the Maximum Likelihood Estimation approach. The Spatial simulations were based on 1,000 replications. In each simulation, we drew random samples of various sizes (s = 25, 50, 100, 1000) of the same system described. Sample size in this section refers to the number of parallel systems drawn. Table 4.2 presents the input values of λ_0, and β used for the simulation.

If all the four components in the system are working, then any of the components can fail with failure rate $\lambda_i = \lambda_0$, and $\lambda_s^{(0)} = 4\lambda_0$.

After the first component failure has occurred, the failure rates of the three remaining components becomes λ_0 for two components and $\lambda_{i(j^1)}$ for the third component. $\lambda_{i(j^1)} = \sigma_1\lambda_0$. Thus, the system failure rate becomes $\lambda_s^{(1)} = (\sigma_1 + 2)\lambda_0$. $\sigma_1 = (z_0 + z_0\vartheta_{ij})^{\beta}$.

As stated in section 4.5, λ_0 and σ_1 will be estimated using the first and second failure times of the studied system then β will be calculated using Eq.(4.31).

The likelihood function for the s^{th} load-share parallel system model is:

$$L_s(\lambda_0; t_{s1}, t_{s2}, t_{s3}) = \left(4\lambda_0 exp^{(-4\lambda_0 t_{s1})}\right)\left((\sigma_1+2)\lambda_0 exp^{(-(\sigma_1+2)\lambda_0 t_{s2})}\right), \tag{4.37}$$

$$L_s(\lambda_0; t_{s1}, t_{s2}, t_{s3}) = 4(\sigma_1+2)\lambda_0^2 exp^{-(\lambda_0[4t_{s1}+(\sigma_1+2)t_{s2}])}. \tag{4.38}$$

The likelihood function corresponding to p systems is:

$$L(T \mid \lambda_0, \sigma_1, \sigma_2) = 4^p(\sigma_1+2)^p\lambda_0^{2p} exp^{-\left(\lambda_0 \sum_{s=1}^{p}[4t_{s1}+(\sigma_1+2)t_{s2}]\right)}. \tag{4.39}$$

The log likelihood function is:

$$\begin{aligned} l = log(L) = plog(4) + 2plog(\lambda_0) + plog(\sigma_1+2) + \dots \\ - \left(\lambda_0 \sum_{s=1}^{p}[4t_{s1} + (\sigma_1+2)t_{s2}]\right). \end{aligned} \tag{4.40}$$

From Table 4.6 we observe that the bias and MSE in the estimated λ values were less than 1 % for all sample sizes and improved as sample size increased. However, the bias in the estimated σ_1, and β values were -0.15 and -0.20 and reduced as the sample size increased. The MSE in the estimated σ_1, and β values were quite large for sample size 25 and reduced as sample size increased. Overall, the estimator performs well beyond sample size 25.

Table 4.6: Bias and MSE of the estimators.

	λ_0	σ_1	β
Size	0.10	4	2
s=25			
Bais	0.01019	-0.15426	-0.20498
MSE	0.00073	2.83202	0.54028
s=50			
Bais	0.00259	0.09251	-0.02962
MSE	0.00025	1.44490	0.18995
s=100			
Bais	0.00146	0.05104	-0.01254
MSE	0.00011	0.70987	0.09087
s=1000			
Bais	0.00054	-0.0105	-0.00752
MSE	0.00001	0.08380	0.01078

4.9 Discussion of results and Conclusion

In this chapter, we studied the problem of estimating the reliability of a load-sharing system with spatial dependence and proximity effects. We have assumed that a system exists that operates in a way that the load on a failed component is taken up only by its working spatial neighbours in close proximity. The system fails whenever the sum of loads on the working components is less than the load placed on the system i.e. $\sum_{i=1}^{n} z_i < L$. We developed a model to evaluate system reliability, extending the capacity flow model to take into account load-sharing interactions and proximity effects. The model was developed for both homogeneous or heterogeneous components and illustrated through two numerical examples. Sensitivity analysis of the load factor was conducted to examine its effect on reliability estimations. We found that an increased load on the system reduced the system reliability faster.

We examined the impact of using a wrong model for reliability estimation of a four-component load-sharing system with spatial dependent components. The analysis was conducted by comparing our spatial model with existing equal load-sharing models for 2-out-of-4 and 3-out-of-4 systems. Monte Carlo simulation was used to generate cumulative distribution function of the four-component system and to validate the reliability estimation from the compared models. We found that when heterogeneous components are considered, the spatial effect may be ignored and a simpler model such as the model for 3-out-of-4 systems which does not account for spatial effects could be applied. However, when homogeneous components are considered, an application of equal load-sharing model for reliability analysis of a load-sharing system with spatial dependence could lead to either overestimated or underestimated reliability prediction. A similar analysis was made on a five-component system with a more complex structure. Comparison was made between the spatial model and an equal load-sharing model for 1-out-of-5. It was found that when homogeneous components are considered, the equal load model had overestimated reliability assessment.

From the analysis of both the four-component and five-component systems, we can infer that considering homogeneous components regardless of the level of component connections and system size, if it exists, spatial effect should not be ignored in reliability prediction of a system with load-sharing.

Lastly, when the number of components in the system is more than 7 or 5 for systems with homogeneous or heterogeneous components respectively, the computation of the system reliability becomes very lengthy and tedious because we have to consider all possible permutations of the failure sequences of the components. R software was used to develop the computational program to analyse the reliability of the systems.

Appendix - Chapter 4

A.1 Appendix A

Solution of the differential equations describing the state probabilities of the four-component system described in section 4.6 using laplace transforms is presented below. A comparison of the results of Laplace transform method and Euler method used in sections 4.6.

Let us consider the four-component system with homogeneous components. Given the initial conditions of $Q_0(0) = 1$ and $Q_i(0) = 0$ for $i = 1, 2, \ldots, s$, the differential equations describing the state probabilities in Eq.(4.32), (4.33), and (4.34) can be rewritten using laplace transform as:

$$SL_0(s) - 1 = -4\lambda_0 L_0(s), \tag{41}$$

$$SL_1(s) = -(2\lambda_0 + \lambda_{i(j^1)})L_1(s) + 4\lambda_0 L_0(s), \tag{42}$$

$$SL_2(s) = -2\lambda_{i(j^1)}L_2(s) + 2\lambda_0 L_1(s). \tag{43}$$

Next, we rearranging the three equations. For Eq.(41), we get:

$$\begin{aligned} &SL_0(s) + 4\lambda_0 L_0(s) = 1, \\ &L_0(s) = \tfrac{1}{(S+4\lambda_0)}. \end{aligned} \tag{44}$$

For Eq.(42), we get:

$$\begin{aligned} &SL_1(s)(S + 2\lambda_0 + \lambda_{i(j^1)}) = 4\lambda_0 L_0(s) \\ &L_1(s) = \tfrac{4\lambda_0 L_0(s)}{(S+2\lambda_0+\lambda_{i(j^1)})} \\ &L_1(s) = \tfrac{4\lambda_0}{(S+2\lambda_0+\lambda_{i(j^1)})(S+4\lambda_0)}. \end{aligned} \tag{45}$$

Using partial decomposition with repeated factors on Eq.(45) given that $2\lambda_0 + \lambda_{i(j^1)} = 4\lambda_0$, we get:

$$\begin{aligned} &L_1(s) = \tfrac{4\lambda_0}{(S+2\lambda_0+\lambda_{i(j^1)})(S+4\lambda_0)} \\ &\tfrac{4\lambda_0}{(S+4\lambda_0)(S+4\lambda_0)} = \tfrac{C}{(S+4\lambda_0)} + \tfrac{D}{(S+4\lambda_0)^2}. \end{aligned} \tag{46}$$

After solving the equation, $C = 0$ and $D = 4\lambda_0$, we have:

$$L_1(s) = \frac{4\lambda_0}{(S+4\lambda_0)^2}. \tag{47}$$

Using the same steps for Eq.(43), we get:

$$\begin{aligned}
&SL_2(s) + 2\lambda_{i(j^1)}L_2(s) = 2\lambda_0 L_1(s)\\
&L_2(s) = \frac{2\lambda_0 L_1(s)}{(S+2\lambda_{i(j^1)})}\\
&L_2(s) = \frac{8(\lambda_0)^2}{(S+4\lambda_0)^2}.
\end{aligned} \tag{48}$$

Using partial decomposition with repeated factors, we get:

$$L_2(s) = \frac{8(\lambda_0)^2}{(S+4\lambda_0)^3}. \tag{49}$$

Applying Laplace inverse transforms on Eq.(44), (47), and (49) then replace $4\lambda_0$ with $2\lambda_0 + \lambda_{i(j^1)}$ and $2\lambda_{i(j^1)}$ in Eq.(47) and (49) respectively to get:

$$Q_0(t) = exp^{(-4\lambda_0 t)}, \tag{50}$$

$$Q_1(t) = 4\lambda_0 texp^{(-2\lambda_0+\lambda_{i(j^1)}t)}, \tag{51}$$

$$Q_2(t) = 4(\lambda_0)^2 t^2 exp^{(-2\lambda_{i(j^1)}t)}. \tag{52}$$

The reliability of the system is given by:

$$R(t) = \sum_{i=0}^{2} Q_i(t). \tag{53}$$

Next,if we assume that the components in the system are heterogeneous (non-identically distributed), then given the initial conditions of $Q_0(0) = 1$ and $Q_i(0) = 0$ for $i = 1, 2, \ldots, s$, the differential equations describing the state probabilities in Eq.(4.35) can be rewritten using laplace transform as:

$$\begin{aligned}
&SL_0(s) - 1 = -(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)L_0(s)\\
&SL_1(s) = -(\lambda_1 + \lambda_2 + \lambda_{3(4^1)})L_1(s) + \lambda_4 L_0(s)\\
&SL_2(s) = -(\lambda_1 + \lambda_2 + \lambda_{4(3^1)})L_2(s) + \lambda_3 L_0(s)\\
&SL_3(s) = -(\lambda_{1(2^1)} + \lambda_3 + \lambda_4)L_3(s) + \lambda_2 L_0(s)\\
&SL_4(s) = -(\lambda_{2(1^1)} + \lambda_3 + \lambda_4)L_4(s) + \lambda_1 L_0(s)\\
&SL_6(s) = -(\lambda_{1(2^1)} + \lambda_{3(4^1)})L_6(s) + \lambda_2 L_1(s)\\
&SL_7(s) = -(\lambda_{2(1^1)} + \lambda_{3(4^1)})L_7(s) + \lambda_1 L_1(s)\\
&SL_9(s) = -(\lambda_{1(2^1)} + \lambda_{4(3^1)})L_9(s) + \lambda_2 L_2(s)\\
&SL_{10}(s) = -(\lambda_{2(1^1)} + \lambda_{4(3^1)})L_{10}(s) + \lambda_1 L_2(s)\\
&SL_{11}(s) = -(\lambda_{1(2^1)} + \lambda_{3(4^1)})L_{11}(s) + \lambda_4 L_3(s)\\
&SL_{12}(s) = -(\lambda_{1(2^1)} + \lambda_{4(3^1)})L_{12}(s) + \lambda_3 L_3(s)\\
&SL_{14}(s) = -(\lambda_{2(1^1)} + \lambda_{3(4^1)})L_{14}(s) + \lambda_4 L_4(s)\\
&SL_{15}(s) = -(\lambda_{2(1^1)} + \lambda_{4(3^1)})L_{15}(s) + \lambda_3 L_4(s).
\end{aligned} \tag{54}$$

After using partial decomposition without repeated factors and applying Laplace inverse transforms on Eq.(54), we have:

$$Q_0(t) = exp^{-(\lambda_1+\lambda_2+\lambda_3+\lambda_4)t}, \tag{55}$$

$$Q_1(t) = \lambda_4 texp^{-(\lambda_1+\lambda_2+\lambda_{3(4^1)})t}, \tag{56}$$

$$Q_2(t) = \lambda_3 texp^{-(\lambda_1+\lambda_2+\lambda_{4(3^1)})t}, \tag{57}$$

$$Q_3(t) = \lambda_2 texp^{-(\lambda_{2(1^1)}+\lambda_3+\lambda_4)t}, \tag{58}$$

$$Q_4(t) = \lambda_1 texp^{-(\lambda_{1(2^1)}+\lambda_3+\lambda_4)t}, \tag{59}$$

$$Q_6(t) = \left(\tfrac{\lambda_2\lambda_4}{2}\right) t^2 exp^{-(\lambda_{1(2^1)}+\lambda_{3(4^1)})t}, \tag{60}$$

$$Q_7(t) = \left(\tfrac{\lambda_1\lambda_4}{2}\right) t^2 exp^{-(\lambda_{2(1^1)}+\lambda_{3(4^1)})t}, \tag{61}$$

$$Q_9(t) = \left(\tfrac{\lambda_2\lambda_3}{2}\right) t^2 exp^{-(\lambda_{1(2^1)}+\lambda_{4(3^1)})t}, \tag{62}$$

$$Q_{10}(t) = \left(\frac{\lambda_1\lambda_3}{2}\right) t^2 exp^{-(\lambda_{2(1^1)}+\lambda_{4(3^1)})t}, \tag{63}$$

$$Q_{11}(t) = \left(\frac{\lambda_2\lambda_4}{2}\right) t^2 exp^{-(\lambda_{1(2^1)}+\lambda_{3(4^1)})t}, \tag{64}$$

$$Q_{12}(t) = \left(\frac{\lambda_2\lambda_3}{2}\right) t^2 exp^{-(\lambda_{1(2^1)}+\lambda_{4(3^1)})t}, \tag{65}$$

$$Q_{14}(t) = \left(\frac{\lambda_1\lambda_4}{2}\right) t^2 exp^{-(\lambda_{2(1^1)}+\lambda_{3(4^1)})t}, \tag{66}$$

$$Q_{15}(t) = \left(\frac{\lambda_1\lambda_3}{2}\right) t^2 exp^{-(\lambda_{2(1^1)}+\lambda_{4(3^1)})t}, \tag{67}$$

The reliability of the system is given by:

$$R(t) = \sum_{i=0}^{15} Q_i(t). \tag{68}$$

From Table 7 we observe that the system reliability estimations using laplace and numerical method (Euler method) are the same for the systems with homogeneous and heterogeneous components.

Table 7: Comparison of reliability estimation using Laplace transform method and Euler method considering homogeneous and heterogeneous components.

	Homogeneous components		**Heterogeneous components**	
Time	Laplace method	Euler method	Laplace method	Euler method
0	1.000000	1.000000	1.000000	1.000000
10	0.541341	0.541341	0.082052	0.082052
20	0.164840	0.164840	0.018221	0.018220
30	0.039660	0.039659	0.000033	0.000032
40	0.008386	0.008386	0.000000	0.000000
50	0.001634	0.001633	0.000000	0.000000
60	0.000301	0.000300	0.000000	0.000000

Chapter 5

Reliability modelling and preventive maintenance of load-sharing systems with spatial dependence, proximity effects and external shocks

In chapter 4, we studied the problem of reliability estimation of a load-sharing system with spatial dependence and proximity effects. This chapter considers reliability estimation and maintenance of systems subject to load-sharing, spatial dependence, proximity effects and external shocks. In section 2.4.3 of chapter 2, we established the following: First, no research has considered models for reliability prediction for load-sharing system with spatial dependence, proximity effect and external shocks. We also established that models for maintenance optimization of load-sharing systems with spatial dependence, proximity effect and external shocks have not been developed. Lastly, we established that the impact of ignoring spatial dependence if it exists in a load-sharing system subject to spatial dependence and external shocks has not been investigated.

Based on the established gaps in section 2.4.3 of chapter 2, our contribution in this chapter is threefold. First, we will develop a model for estimating the reliability of a load-sharing system with spatial dependent components and random shock processes. Second, we will develop a model for preventive maintenance of load-sharing system with spatial dependence, proximity effect and external shocks. Lastly, using the developed models, we will conduct an assessment of the impact of ignoring spatial dependence if it exists in a load-sharing system subject to external shocks.

The remainder of this chapter is outlined as follows: Section 5.1 introduces the system description for the study. Section 5.2 presents component reliability quantities and system reliability quantities. Section 5.3 presents expressions for the age-replacement policy. Section 5.4, 5.5 and 5.6 presents illustrations of the developed methods through numerical examples. Section 5.7 presents the conclusion.

5.1 System description

The system consists of m components connected in parallel. The system starts with a certain amount of load equally distributed to its components. The following assumptions are made to better position our study:

- Each component can either be in a working or failed state while the system is multi-state and can function at different performance levels depending on the states of its components.
- The system fails if the sum of the loads on each working component at time t is less than the total system load L.
- When a component j fails, its working proximate component i in any direction can take up the failed component's load.
- The effect of a failed component's load acts multiplicatively on the failure rate of a proximate component.
- Component i is operating in a random environment modelled by the HPP of shocks $N(t), t \geq 0$, with rate v, where $N(t)$ is the number of shocks in $[0, t)$.

- Shocks on its own do not cause immediate component failure but they cause some damage.
- The effect of external shock acts additively on the failure rate of a working component.

5.2 Component and system reliability analysis

5.2.1 Component reliability analysis

Let us consider a component i in the load-sharing system described in section 5.1. Let $N_s(t)$ be the number of shocks from the beginning of the planning horizon until time t. Given that component i is affected by extra loads and external shocks, whose occurrence are random, we introduce the piece wise failure rate of component i at time t considering a sample path in Figure 5.1. Provided that one spatial neighbour j has failed and there were two external shocks from the beginning of the planning horizon until time t, the failure rate is:

$$\lambda_{i(j^r)}(t) = \begin{cases} \lambda_0, & j = 0, \quad 0 < t < t_{s1} \\ \lambda_0 + \phi, & j = 0, \quad t_{s1} \leq t < t_{j1} \\ \lambda_0 z_{i(j)} + \phi, & j > 0, \quad t_{j1} \leq t < t_{s2} \\ \lambda_0 z_{i(j)} + 2\phi, & j > 0, \quad t_{s2} \leq t \end{cases}, \tag{5.1}$$

where t_s denotes the arrival time of shocks and t_j denotes the failure time of component j. $z_{i(j)} = (z_i + \vartheta_{ij} z_j)^{\beta}$. z_i is the initial load imposed on component i, and λ_0 is the baseline failure rate (also called hazard rate) function. β is the load factor which describes the influence of the increased load on the failure rate of component i. When $\beta = 0$, an increased load has no impact on a component's failure rate. j denote the index of a failed neighbour of component i. r is the number of working neighbours of component j that shares its load. ϑ_{ij} denotes the percentage of a failed component's load that affects the failure rate of a working proximate component. If $\vartheta_{ij} = 0$ a working component is not impacted by the failure of component j. ϕ is a constant deterministic jump that occurs on each shock arrival. ϕ denotes the shock

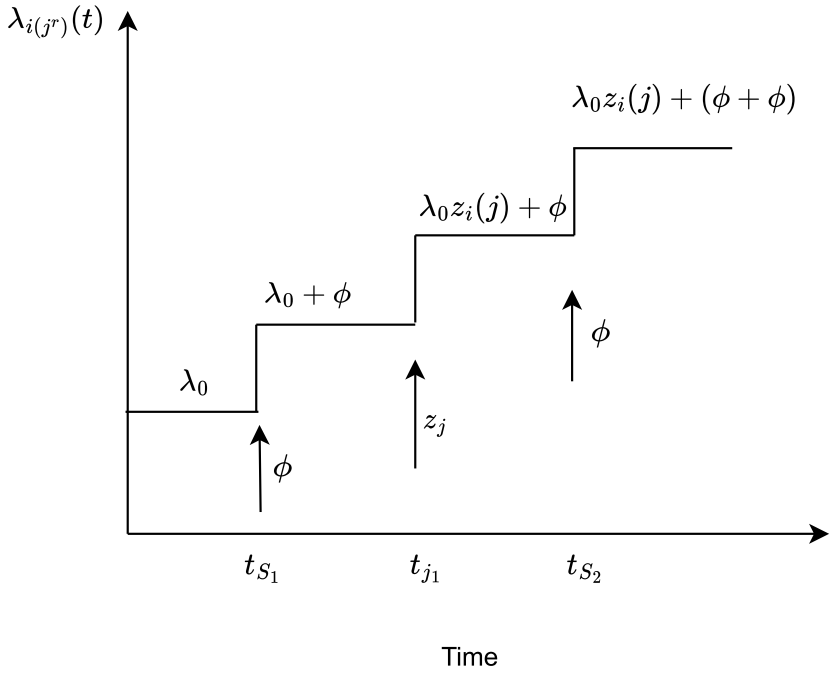


Figure 5.1: Failure rate function under two shocks with increments ϕ and extra load z_j.

size and it represents the effect that each shock has on the failure rate. If $\phi = 0$ then the shocks do not influence the failure rate. Thus, the shock process and the lifetime of component i are independent. If there is no external shock until time t, then the failure rate of component i becomes Eq.(4.16) which was studied in chapter 4. Note that external shocks on the system affects all the working components at the same time.

5.2.2 System reliability analysis

Consider a system consisting of m components in parallel. Assume that the duration of one mission, which is equal to the system's lifecycle is T. The system may start with new or used components. The age of new components when the system starts operating is zero whereas the age of a used component is not zero. Components of the system might not have the same initial age at the beginning of the mission, even though they are both subject to the same amount of load and external shocks. Therefore, their failure rates may vary to reflect their different initial ages. The components may or may not fail during a mission. If the system works at the end of a mission, its components are preventively maintained otherwise they are correctively maintained. As such, the following cases may arise.

- *Case 1*: a number of components fail during the mission but the system still works because the sum of loads on the working components still equals the system load. In this case, the system is preventively maintained by replacing the failed components at a predetermined time.
- *Case 2*: a number of components fail during the mission and the system fails because the sum of loads on the working components is less than the system load. In this case, the system is correctively maintained by replacing all components in the system before the predetermined time.
- *Case 3*: None of the components fail during the mission. The components are preventively maintained.

In each case, the probability of system failure, depend on the sequence of the component failures. For instance, let us consider case 2 for a two-component system whose components are indexed as A and B. Component A could fail first and its load is transferred to component B after the failure before component B fails. In the same way, component B could fail first before component A fails. Both scenarios present a possible path to the system failure. However, when external shock arrivals are considered in each case, the cases become even more complex as the arrival of shocks are also random. The stochastic behaviour of external shocks and failures of each component j makes it difficult to have a closed form failure rate function as seen in section 5.2.1. Thus, two approaches for modelling the system reliability will be introduced in this chapter. First modelling approach considers a complex case and is based on simulation. For the complex case, Monte Carlo simulation is used to derive the system reliability quantities. Monte Carlo methods are useful for estimating system reliability that cannot be derived analytically [16, 131]. The idea of Monte Carlo method is the generation of random events which are repeated many times and the number of a specific event is counted

[132]. The second approach considers a simple case where the effect of each external shock is added on all working component after the failure of a component in the system. For the simple case an analytic model is used to derive the system reliability quantities.

Let us consider the complex case described earlier. Consider that there were n number of external shocks from the beginning of planning horizon until time t which occur according to HPP with constant failure rate v. Thus, we introduce algorithm 2 for generating the system reliability and failure distribution as follows.

Algorithm 2 Algorithm for generating lifetime data of a system with load-sharing, spatial dependence and external shocks

Require: shock size ϕ; load factor β; shock inter-arrival time X_s; component failure rate $\lambda_{i(j^r)}$; baseline failure rate λ_0; observation time τ; number of shocks n; arrival rate of shocks v; system load L; number of components m; proximity matrix $[\boldsymbol{\vartheta_{ij}}]$.

Ensure: System lifetime T_{sys} at the end of the observation period.

1: compute $T_{sys} = \sum_k T_k$ for $k = m + n$.
2: set up the value of parameters: v, τ, ϕ, L, λ_0, and β.
3: generate a random number n from Pois$(V(\tau))$ where $V(\tau) = v\tau$.
4: generate n shock inter-arrival times $X_s = \frac{-\ln(U[0,1])}{v}$ where $s = 1, 2 \ldots, n$.
5: set $p = 0$, $h = 0$ and $z_{i(j)} = 1$.
6: start at $k = 1$.
7: **repeat**
8:     **for all** $i \in \{k, k+1, \ldots, m\}$ **do**
9:         simulate m random variate u_i from $U[0, 1]$.
10:         calculate $t_i = \frac{-\ln(u_i)}{\lambda_{i(j^r)}}$ where $\lambda_{i(j^r)} = \lambda_0 z_{i(j)} + \phi p$.
11:     **end for**
12:     select the component with least failure time $t_j = min(t_k, t_{k+1}, \ldots, t_m)$.
13:     set the next shock inter-arrival time $t_s = X_s$ in order from $s = 1$ to $s = n$.
14:     select $T_k = min(t_i, t_s)$.
15:     if $T_k = t_s$, then shock arrived before component j failed. go to step 17.
16:     if $T_k = t_j$, then component failed before shock arrival. go to step 18.

17: set $p = p + 1$ then update p in $\lambda_{i(j^r)} = \lambda_0 z_{i(j)} + \phi p$. go to step 21.
18: calculate $z_{i(j)} = (z_i + \vartheta_{ij} z_j)$ for proximate neighbours of failed component j.
19: set $h = h + 1$ then evaluate whether $\sum_{i=1}^{m-h} z_i = L$.
20: recalculate $z_{i(j)}$ in $\lambda_{i(j^r)} = \lambda_0 z_{i(j)} + \phi p$ for proximate neighbours of failed component j.
21: set $k = k + 1$.
22: **until** $\sum_{i=1}^{m-h} z_i < L$;
23: **return** $T_{sys} = \sum_k T_k$.

The probability that the system survives till time T is:

$$R_{sys}(T) = \frac{r(T)}{q}. \tag{5.2}$$

where r(T) denotes the number of times the system survived beyond time T. q is the number of system lifetimes generated.

The probability that the system fails before time T is:

$$F_{sys}(T) = 1 - \frac{r(T)}{q}. \tag{5.3}$$

Given the failed components and the number of external shocks, we can derive the failure rate for component i as:

$$\lambda_{i(j^r)}(t) = \begin{cases} \lambda_0 + \phi n, & j = 0 \\ \lambda_0 z_{i(j)} + \phi n, & j > 0 \end{cases}. \tag{5.4}$$

For the simple case, we consider a specific scenario where shocks arrive before the failure of any component in the system. We assume that one shock arrives before the failure of a component in the systems and the shock does not cause component failure. We also assume that the shock arrival time is not known. However, the effect of each external shock is added on all working component after the failure of a component in the system.

Let us consider a component in a system with the sample path in Figure 5.2. Provided that two spatial neighbours j and k have failed and there were two external shocks from the beginning of the planning horizon until time t, the failure rate is:

$$\lambda_{i(j^r)}(t) = \begin{cases} \lambda_0, & j = 0, 0 < t < t_{j1} \\ \lambda_0 z_{i(j)} + \phi, & j > 0, t_{j1} < t < t_{k2} \\ \lambda_0 z_{i(j,k)} + 2\phi, & j > 0, k > 0, t_{j1} < t < t_{k2} \end{cases}. \tag{5.5}$$

Note that the failure rate of component i can be generalised as:

$$\lambda_{i^s(j^r)}(t) = \begin{cases} \lambda_0 + \phi s, & j = 0, s = 0, 1, ... \\ \lambda_0 z_{i(j)} + \phi s, & j > 0, s = 0, 1, ... \end{cases}, \tag{5.6}$$

where s is the number of shocks that have arrived. Given that the effect of shocks influences each components' failure rate after a component has failed s is then linked to the number of failed components in the system. Note that s counts the number of shocks as well as the numbers failed components while j is an index of a failed spatial neighbour of component i. When all components are working, and no shock arrival $s = 0$. For the simple case, the probability density function (pdf) of component i at time t is given by:

$$f_i(t) = \lambda_{i^s(j^r)}(t) e^{-\int_0^t \lambda_{i^s(j^r)}(u)du}. \tag{5.7}$$

The cumulative distribution function (cdf) of component i is given by:

$$F_i(t) = 1 - e^{-\int_0^t \lambda_{i^s(j^r)}(u)du}. \tag{5.8}$$

The reliability function is:

$$R_i(t) = e^{-\int_0^t \lambda_{i^s(j^r)}(u)du}. \tag{5.9}$$

Next, for the simple case, we consider the system reliability quantities. The system failure distribution and density will be derived by utilizing the paths to the system failure event. To aid in identifying all the paths that could lead to the system failure event, we introduce a modified failure sequence diagram (MFSD) as an extension of the failure sequence diagram (FSD) introduced by Wang et al. [303] for linear consecutive

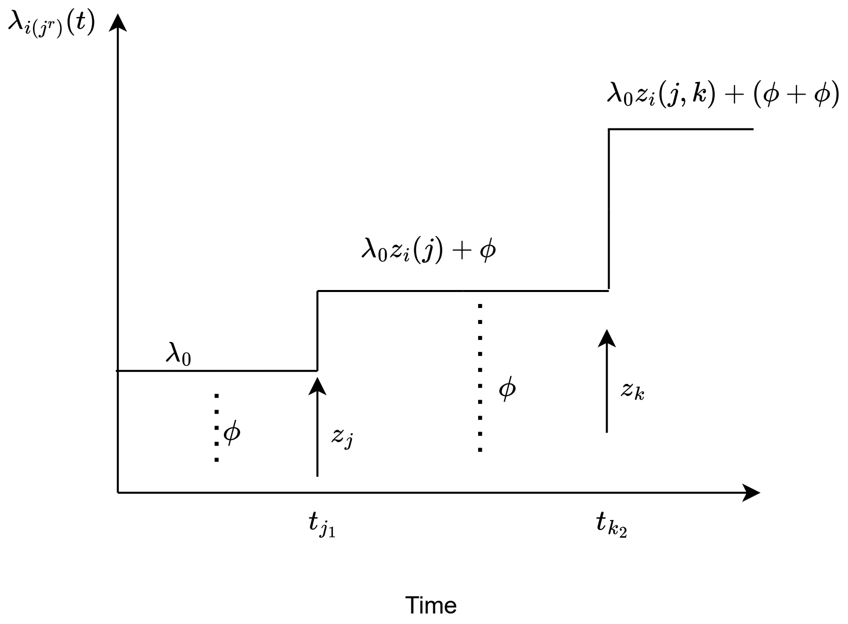


Figure 5.2: Failure rate function under two shocks and two component failures with increments ϕ and extra loads z_j and z_k.

k-out-of-n system. The failure sequence diagram is a tree diagram that provides a visual representation of the all the possible path to the system failure (e.g., see Figure 5.3). Unlike FSD which considers only adjacent neighbour interactions, MFSD accounts for proximity effect between components in the system. Whilst the visual representation of system's failure path derived by the MFSD is like the system state diagram of a Markov model in the way that both models evaluate the probability of jumping from one known state into the next logical state until the system being considered has reached the final failed state or until a particular mission time is achieved. The difference is that using the MFSD allows the modelling of the sojourn time of each state by arbitrary lifetime distributions whose failure rates are functions of the baseline failure rate, load redistributed and external shocks. The use of system state diagram, outside a Markov model,which captures arbitrary distributions for load-sharing system reliability analysis has been considered in the load-sharing literature. For example, Liu [304] studied the reliability of a load-sharing k-out-of-n:G system with arbitrary load-dependent component lifetime distributions. In their study, it was assumed that after load redistribution, the component age changes according to an accelerated failure-time model. Li et al [305], conducted reliability analysis of a two component hydraulics system with arbitrary lifetime distribution. The algorithm for constructing the MFSD is as follows:

1. Start the MFSD construction with the representation of the entire system indicated by an initial node "S". If there are no failed components in the entire system during a mission, we indicate it by adding node 0. When a component fails, we add a node j which indicates that component j has failed where $j = 1, 2, \ldots, m$. The failure rate of component i before failure is indicated on the arrow pointing to node i. We indicate the updated failure rates of the working components below node i.

2. According to the characteristics of the load-sharing system, that the system fails if the sum of loads on the working components is less than the load on the system, we need to judge if the system has failed or not in any path. If the condition is satisfied, then we indicate that the system fails by adding node Cm to the path. If the system survives till the end of the mission after component j has failed, then we indicate by adding node Pm.

3. Once the jth level for the MFSD has been constructed, the $(j + 1)$th level for the MFSD can be constructed similarly to step 2. The construction of the MFSD stops when all the paths end either at node 0, node Pm or node Cm.

To illustrate the developed model, we will consider a two-component system in Figure 5.4 with heterogeneous components indexed as 1 and 2. Let us assume that both components have ages zero. The system fails when both components fail before T. Using the paths derived in Figure 5.4 we will derive the probabilities for the system to fail or work at time t for each path. Let f_{jkl} and F_{jkl} denote the pdf and cdf of the system's life given that components j, k and l failed. Let P_{jkl} indicate probability that the system survives till time t given that components j, k and l failed. The probability that the system survives till the end of mission denoted by time t and no component

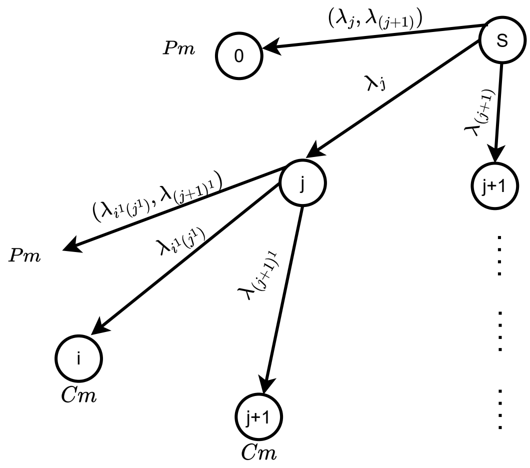


Figure 5.3: MFSD path description

failed is given by:

$$P_0(t) = R_1(t|0, \lambda_1)R_2(t|0, \lambda_2). \tag{5.10}$$

The probability that the system survives till the end of mission denoted by time t and only component 1 failed is given by:

$$P_1(t) = \int_0^t f_1(t_1|0, \lambda_1)R_2(t_1|0, \lambda_2))R_2(t - t_1|t_1, \lambda_{2^1(1^1)})dt_1. \tag{5.11}$$

The probability that the system survives till time t and only component 2 failed is given by:

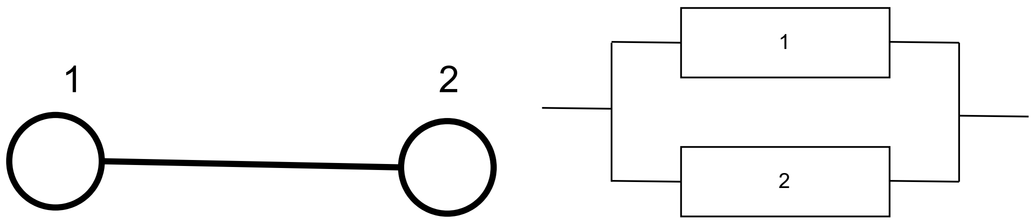
$$P_2(t) = \int_0^t f_2(t_1|0, \lambda_2)R_1(t_1|0, \lambda_1)R_1(t - t_1|t_1, \lambda_{1^1(2^1)})dt_1. \tag{5.12}$$

The density of the system failure at a time T before end of mission time t given that component 1 failed first is:

$$f_{12}(T) = \int_0^t f_1(t_1|0, \lambda_1)R_2(t_1|0, \lambda_2)f_2(T - t_1|t_1, \lambda_{2^1(1^1)})dt_1. \tag{5.13}$$

The density of the system failure before time T given that component 2 failed first is:

$$f_{21}(T) = \int_0^t f_2(t_1|0, \lambda_2)R_1(t_1|0, \lambda_1)f_1(T - t_1|t_1, \lambda_{1^1(2^1)})dt_1. \tag{5.14}$$



(a) Spatial connections - two links (b) Block diagram of parallel connections

Figure 5.4: The two-component system

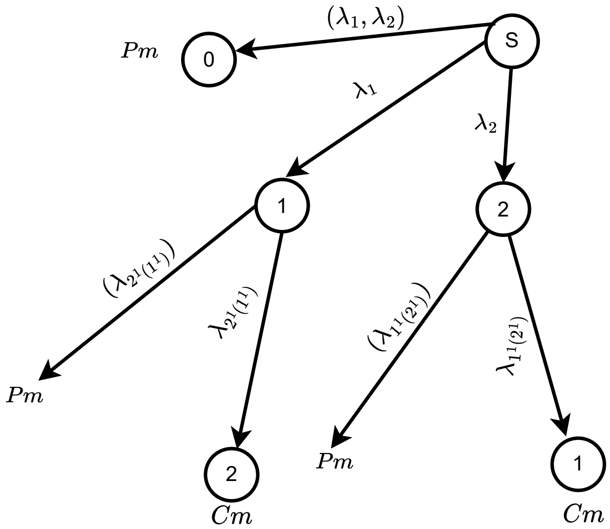


Figure 5.5: MFSD for the two-component system.

Using derived expressions, we can derive approximations of the system reliability quantities as follows.

The probability that the system survives till time T is:

$$R_{sys}(T) = \sum_{jk..} P_{jk}(T). \tag{5.15}$$

The density of the system failure time is:

$$f_{sys}(T) = \sum_{jk..} f_{jk}(T). \tag{5.16}$$

The probability that the system fails before time T is:

$$F_{sys}(T) = \sum_{jk..} \left(\int_0^t f_{jk}(t)dt \right) \quad ,$$

where $F_{sys}(T) + R_{sys}(T) = 1$.

Alternatively, $f_{sys}(T) = \sum_{jk..} \left(\frac{dF_{jk}(T)}{dT} \right)$.

The system's failure rate is:

$$h_{sys}(T) = \frac{f_{sys}(T)}{R_{sys}(T)}. \tag{5.17}$$

5.3 Age-based replacement policy

In this section, we consider the problem of finding the optimal replacement time which minimizes the total long-run average cost per unit time. According to the classical age-replacement policy, the system is replaced at time t_{PM} or at failure, whichever occurs first. It is assumed that a component's failure can be known at any time. If the sum of loads on the working components equals that placed on the system at time t_{PM}, the system is preventively maintained (PM) by replacing all failed components with new components. If the system failure occurs before t_{PM}, corrective maintenance (CM) is done by replacing the system. Assume that repair times are negligible.

Let c_{CM} denote the replacement cost at failure (the cost of CM) and c_{PM} denote the cost of PM, where $c_{CM} > c_{PM}$. If c_R is the cost of component replacement, then the expected cost of replacing the system is:

$$C_T(t_{PM}) = \begin{cases} c_{CM} + c_R E(N_j(T) \mid T < t_{PM}), & \text{if } T < t_{PM} \\ c_{PM} + c_R E(N_j(t_{PM}) \mid T \geq t_{PM}), & \text{if } T \geq t_{PM}, \end{cases} \tag{5.18}$$

where $N_j(t_{PM})$ denote the total number of failed components by time t_{PM}. $E(N_j(t_{PM}) \mid T \geq t_{PM}) = \sum_{jk..} pP_{jk}(t_{PM})$ is the expected number of failed components given that the system survives till t_{PM}. In contrast, $E(N_j(T) \mid T < t_{PM}) = \sum_{jk..} pF_{jk}(t)$ is the expected number of failed components given that the system failed before t_{PM}.

Assume that if the system is replaced at time t_{PM}, then all failed components are replaced by new components. The long-run expected cost per unit of time is equal to the expected cost per unit of time for one renewal cycle because of the renewal reward theorem (see, e.g. [306]). The expected cost for one cycle is defined as:

$$\begin{aligned} C_T(t_{PM}) &= [c_{CM} + c_R E(N_j(T) \mid T < t_{PM})]\, F_{sys}(t_{PM}) \\ &+ [c_{PM} + c_R E(N_j(t_{PM}) \mid T \geq t_{PM})]\, R_{sys}(t_{PM}), \end{aligned} \tag{5.19}$$

where $R_{sys}(t)$ is the probability that the system will be preventively maintained. $F_{sys}(t)$ is the probability that the system will be correctively maintained.

Thus, the expected cost rate can be represented as is:

$$\begin{aligned} &E(CR(t_{PM})) = \tfrac{C_T(t_{PM})}{E(min(T,t_{PM}))} \\ &= \tfrac{[c_{CM}+c_R E(N_j(T)|T<t_{PM})]F_{sys}(t_{PM})+[c_{PM}+c_R E(N_j(t_{PM})|T\geq t_{PM})]R_{sys}(t_{PM})}{\int_0^{t_{PM}} R_{sys}(u)du} \end{aligned}, \tag{5.20}$$

where $\int_0^{t_{PM}} R_{sys}(u)du = t_{PM}R_{sys}(t_{PM}) + \int_0^{t_{PM}} uf_{sys}(u)du$. We are interested in minimizing $E(CR(t_{PM}))$ with respect to the decision variable t_{PM}. Note that we ignore the cost of rejuvenating all non-failed components at time t_{PM}. A component is as good as new after rejuvenation. The cost of rejuvenating a component can easily be included in the model.

For the complex case, we introduce algorithm 3 for generating the expected cost rate as follows (see Appendix A for the flowchart).

Algorithm 3 Algorithm for generating the expected cost rate of a system with load-sharing, spatial dependence and external shocks

Require: preventive maintenance time t_{PM}; Optimal preventive maintenance time t^*_{PM}; number of shocks n; shock inter-arrival time X_s; component failure time t_j; cycle time corrective maintenance $CLCM$; cycle time preventive maintenance $CLPM$; simulation number SN; number of simulation runs SRN; number of failed components NF; number of working systems NWS; number of system failures NSF; number of failed components at corrective maintenance $NFCM$; number of failed components at preventive maintenance $NFPM$.

Ensure: Expected cost rate $E(CR(t_{PM}))$ at the end of the simulation.

1: set up the value of parameters: $t_{PM(max)}$, v, ϕ, L, λ_0, and β.

2: set $t_{PM} = 1$.

3: set $SN = 1$, $NF = 0$, $NWS = 0$, $NSF = 0$, $NFPM = 0$, $NFCM = 0$, $CLPM = 0$ and $CLCM = 0$.

4: generate a random number n from $\text{Pois}(V(t))$ where $V(t) = vt_{PM}$.

5: generate n shock inter-arrival times $X_s = \frac{-\ln(U[0,1])}{v}$ where $s = 1, 2 \ldots, n$.

6: set $p = 0$, $h = 0$ and $z_{i(j)} = 1$.

7: start at $k = 1$.

8: **repeat**

9: **for all** $i \in \{k, k+1, \ldots, m\}$ **do**

10: simulate m random variate u_i from $U[0,1]$.

11: calculate $t_i = \frac{-\ln(u_i)}{\lambda_{i(j^r)}}$ where $\lambda_{i(j^r)} = \lambda_0 z_{i(j)} + \phi p$.

12: **end for**

13: select the component with least failure time $t_j = min(t_k, t_{k+1}, \ldots, t_m)$.

14: set the next shock inter-arrival time $t_s = X_s$ in order from $s = 1$ to $s = n$.

15: select $T_k = min(t_i, t_s)$.

16: if $T_k = t_s$, then $NF = NF + 0$, and $p = p + 1$.

17: if $T_k = t_j$, then $NF = NF + 1$, $h = h + 1$ and recalculate $z_{i(j)}$ in $\lambda_{i(j^r)} = \lambda_0 z_{i(j)} + \phi p$ for proximate neighbours of failed component j.

18: $T_{sys} = \sum_k T_k$.

19: if $T_{sys} \geq t_{PM}$, then $NWS = NWS + 1$, $NFPM = NFPM + 1$, and $CLPM = CLPM + t_{PM}$ and go to step 23.

20: if $T_{sys} < t_{PM}$ and $\sum_{i=1}^{m-h} z_i = L$, set $k = k + 1$ then go to step 9.

21: if $T_{sys} < t_{PM}$ and $\sum_{i=1}^{m-h} z_i < L$, then $NSF = NSF + 1$, $NFCM = NFCM + 1$, and $CLCM = CLCM + t_{PM}$ and go to step 22.

22: if $SN < SRN$ then $SN = SN + 1$ and go to step 4. Else if $SN = SRN$ then go to step 23.

23: calculate $R_{sys}(t_{PM}) = \frac{\sum NWS}{SRN}$.

24: calculate $F_{sys}(t_{PM}) = \frac{\sum NSF}{SRN}$.

25: calculate $E(N_j(T) \mid T < t_{PM}) = \frac{\sum NFCM}{SRN}$.

26: calculate $E(N_j(T) \mid T \geq t_{PM}) = \frac{\sum NFPM}{SRN}$.

27: calculate $E(min(T, t_{PM})) = \frac{\sum(CLCM+CLPM)}{SRN}$.

28: calculate $E(CR(t_{PM})) = \frac{C_T(t_{PM})}{E(min(T,t_{PM}))}$, then go to step 29.

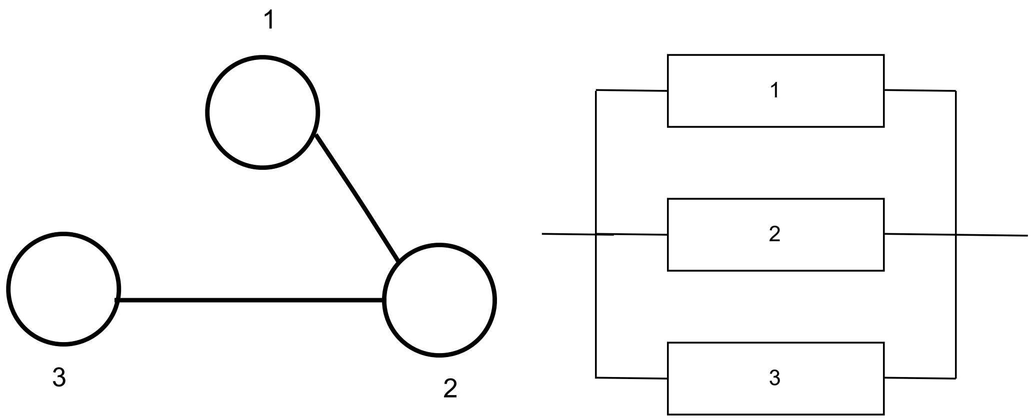
29: set $t_{PM} = t_{PM} + 1$.

30: **until** $t_{PM} < t_{PM(max)}$;

31: **return** $E(CR(t^*_{PM})) = min(E(CR(t_{PM})))$ and t^*_{PM} where $t_{PM} = 1, 2 \ldots, t_{PM(max)}$.

5.4 Numerical example: three-component systems

In this section we use a three-component system with a simple structure to illustrate the developed spatial model for the complex case. The purpose is to present the behaviour of the developed model, and to investigate the importance and significance of the spatial effect in maintenance decision making if it exists. We refer to the structure as an Isosceles structure. The Isosceles structure is depicted in Figure 5.6. Note that the components in the system are connected in parallel as seen in Figure 5.6b, while the links between the components in Figure 5.6a indicate spatial dependence. Components in the Isosceles system are linked together in a way that the system is composed of one dominant component (component 2) and two secondary components (component 1, and 3). As indicated by Figure 5.6a, the system fails when the dominant component and one secondary component fail. In this case the load on the only working component is less than the load placed on the system. We assume that there is no information about the distance between components in the system so that the component arrangement will be used to infer component proximity. We assume that the combined effects of load-sharing and shocks could have a significant impact on the whole system and that the impact on the system could vary depending on the form of load-sharing between components in the system, that is, whether load-sharing is uniform (equal) or non-uniform (with spatial dependence). The effects of load-sharing considering spatial dependence and external shocks will be investigated from the perspective of how long the load-sharing system should be in usage? and at what cost? to highlight the importance and significance of spatial effect.



(a) Spatial connections - two links (b) Block diagram of parallel connections

Figure 5.6: Isosceles system structure

Load-sharing in the Isosceles system is such that after the failure of component 2, components 1 and 3 take up an equal amount of extra load. However, if either of component 1 or component 3 fails, component 2 takes up all the extra load. We compare optimal replacement time and cost from our developed model with a model that assumes equal load-sharing and shocks. We extend the capacity flow model [248] which models equal load-sharing to account for shocks similar to Eq.(5.4). The extended equal load model is given by:

$$\lambda_{i(J)}(t) = \begin{cases} \lambda_0 + \phi n, & J = 0 \\ \lambda_0 \frac{L}{(m-J)} + \phi n, & J > 0 \end{cases}, \tag{5.21}$$

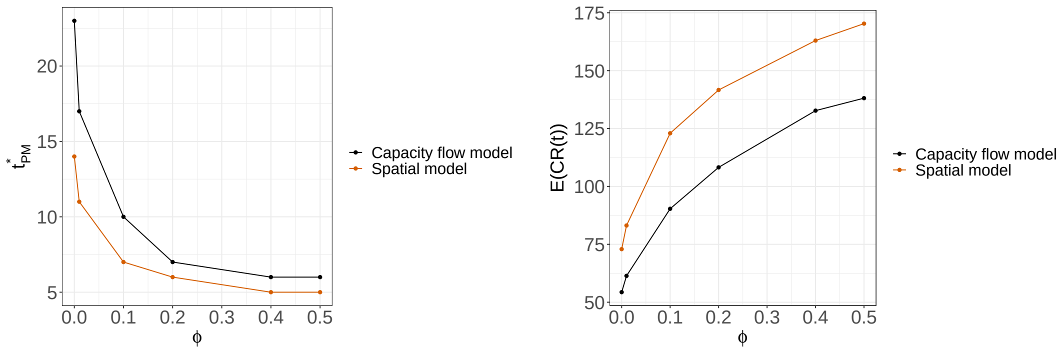
where m is the number of components in the system. J is the number of components that has failed. Note that J counts the number of failed components while j is an index of a failed spatial neighbour of component i. L is the total load placed on the system.

We assume that the parameters to be used for the numerical analysis are known. The following parameters are used in the numerical analysis:

$$L = 3, \beta = 1, \lambda_i = 0.05, \phi = 0, \phi = 0.01, \phi = 0.1, \phi = 0.2, \phi = 0.4, \phi = 0.5, v = 0.25$$

L is the load placed on the system. $\beta = 1$ indicates the presence of load-sharing which has an impact on the failure rate. $\phi = 0$ assumes that the system is not affected by external shocks. When the system is not subject to external shocks, the system is similar to the load-sharing system studied in chapter 4. ϕ from 0.01 to 0.5 indicates low to high impact of shocks on the system. We assume that the cost of component replacement $c_R = 50$. Cost of corrective maintenance $c_{CM} = 1000$ and cost of preventive maintenance $c_{PM} = 400$.

We use algorithm 3 in section 5.3 to estimate the system reliability quantities for the two models which were applied in the age replacement model. The results are generated from 1000 simulation runs. To ease the computation of the system reliability quantities, we implement the method using R. The optimal preventive maintenance time and expected cost rate from the two models are presented in Figure 5.7a and 5.7b respectively. If spatial effect is not known to exist in the Isosceles system even though it exists and maintenance decisions are based on the capacity flow model due to existence of load-sharing, then the following outcomes will be recommended. If the system is operating in an environment where it is either not subject to shocks or the impact of shocks on the system is very low, the optimal time for preventive maintenance is $t_{PM} = 23$ and $t_{PM} = 17$ while the cost rates are \$54.34 and \$61.37 respectively. If the impact of shocks on the system is high, that is from $\phi = 0.1$ to $\phi = 0.5$, the system is recommended for preventive maintenance ranging from $t_{PM} = 10$ to $t_{PM} = 6$ at a high cost rate ranging from \$90.35 to \$138.12. If maintenance decisions for the system are based on the developed spatial model, which captures the spatial effect, then the following outcomes will be observed. If the system is operating in an environment where it is either not subject to shocks $\phi = 0$ or the impact of shocks on the system is very low $\phi = 0.01$, the optimal time for preventive maintenance is $t_{PM} = 14$ and $t_{PM} = 11$ while the cost rates are \$72.92 and \$83.13 respectively. If the impact of shocks on the system is high, that is from $\phi = 0.1$ to $\phi = 0.5$, the system is recommended for preventive maintenance ranging from $t_{PM} = 7$ to $t_{PM} = 5$ at a high cost ranging from \$122.96 to \$170.30. It can be noticed from Figure 5.7a that the optimal t_{PM} derived based on the two models decreases as the shock size ϕ are increased. Implying that it is better to schedule preventive maintenance faster for a system subject to increased shocks due to the cumulative damage on the system. Furthermore, it can be noticed



(a) Optimal preventive maintenance time (b) Expected cost rate at t^*_{PM}

Figure 5.7: Comparison of the optimal preventive maintenance time and expected cost rate using the spatial model and capacity flow model

from Figure 5.7b that the expected cost rate from both models increases as impact of shocks increases.

In general, when load-sharing exists but the impact of shocks on the system is low or does not exist, the optimal preventive maintenance time based on the spatial model is almost half the time and the expected cost rates are less than those recommended by the capacity flow model. If decisions are made based on the capacity flow model, one implication could be that the system might fail before the scheduled preventive maintenance time. Another implication could be that the budget for maintenance would be underplanned by almost half of what is needed. In contrast, when load-sharing exists and the impact of shocks on the system is high, the difference in the expected cost rate from both models is significant whereas the difference in the optimal preventive maintenance time is not significant. The closeness of the optimal preventive maintenance time from both models may be because the impact of spatial effect on the system cannot be distinguished by the two models given the high effect of shocks.

Given that spatial effect on the Isosceles system cannot be distinguished between the two models when the effects of shocks are high, the effect can be ignored and simple equal load-sharing models such the capacity flow model would be suitable for modelling load-sharing and shocks. However, when the effect of shocks on the system is low and spatial effect exists then the model that accounts for spatial effect is suitable.

5.5 Numerical example: four-component systems

In this section, we use two four-component systems structures to further investigate the importance and significance of the spatial effect if it exists. We consider the complex case. The structure of the two four-component systems are depicted in Figure 5.8. The components in each of the four-component systems are connected in parallel as depicted in Figure 5.8c, while the links between the components indicate spatial dependence. For structure one depicted in Figure 5.8a, the components are linked together in a way that the system is composed of two subsystems (that is, subsystem having components 1 and 2 and another subsystem having components 3 and 4). The system fails when two connected components have fail. For structure two depicted in Figure 5.8b, the components are linked together in a way that every component is linked with two other components. The system fails when three components fail. We assume that load distribution in the two structures are not uniform given the presence of spatial effect. The level of non-uniform load distribution changes from high to low as more components are connected. Non-uniform load distribution in the two systems naturally becomes uniform load distribution when each component is connected to the other 3 components. Structure one has a high level of non-uniform load distribution while structure two has a low level of non-uniform distribution given the number of connected components. The purpose of this section is first to identify how the degree of spatial dependence which is controlled by the number of connected components can influence maintenance decision making. Second, how well the extended capacity flow model performs given the varied spatial dependence. Similar to example one in section 5.4, we assume that there is no information about the distance between components in the two systems so that the component arrangement will be used to infer component proximity. We will consider that the two structures have homogeneous components.

We consider the same parameters as those used in section 5.4. The following parameters are used in the numerical analysis:

$$L = 4, \beta = 1, \lambda_i = 0.05, \phi = 0, \phi = 0.01, \phi = 0.1, \phi = 0.2, \phi = 0.4, \phi = 0.5, v = 0.25$$

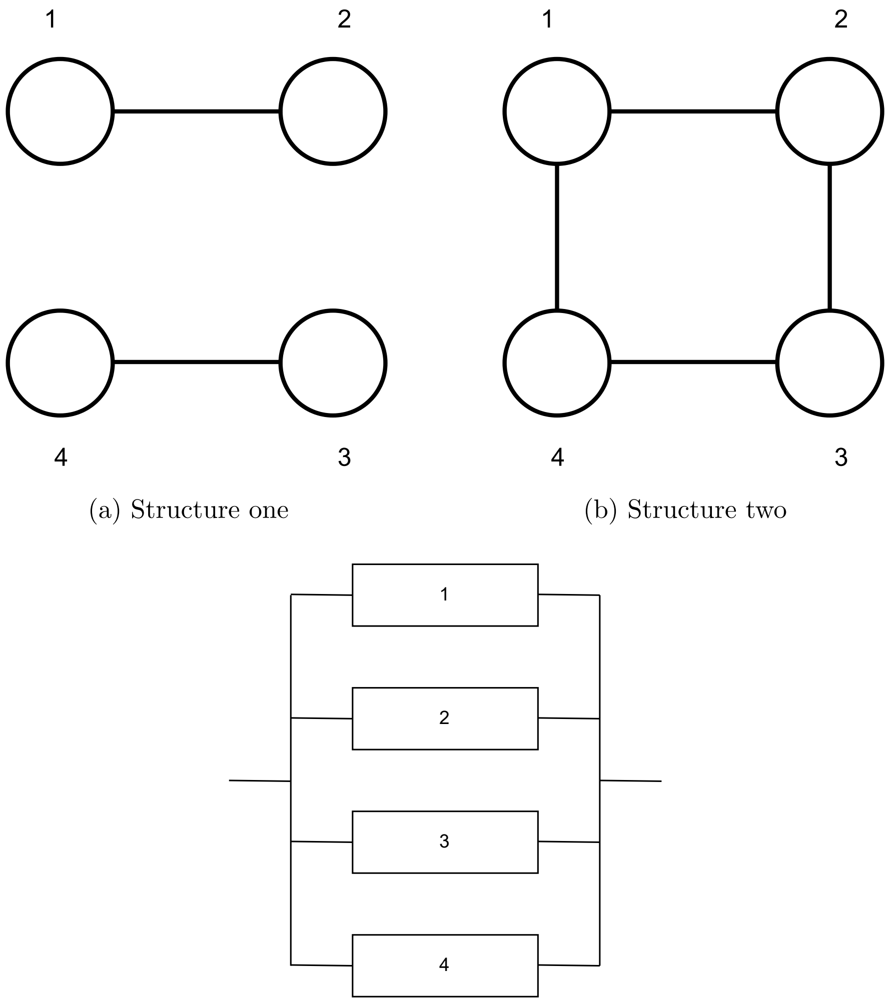
L is the load placed on the system. $\beta = 1$ indicates the presence of load-sharing which has an impact on the failure rate. $\phi = 0$ assumes that the system is not affected by external shocks. When the system is not subject to external shocks, the system is similar to the load-sharing system studied in chapter 4. ϕ from 0.01 to 0.5 indicates low to high impact of shocks on the system. We assume that the cost of component replacement $c_R = 50$. Cost of corrective maintenance $c_{CM} = 1000$ and cost of preventive maintenance $c_{PM} = 400$.

The optimal preventive maintenance time and expected cost rate for the two structures are presented in Figures 5.9a, 5.9b, 5.10a and 5.10b. When the systems are subject to zero shocks or low shock sizes, capacity flow model estimates the same optimal preventive replacement time and expected cost rate for the two structures. In addition, estimates of the optimal preventive maintenance time are 2 times that of the spatial model, for example, capacity flow estimates $t_{PM} = 32$ and $t_{PM} = 27$ for cases of the system being either not subjected to shocks or low shocks.

From all the figures, it can be noticed first that the estimate of optimal preventive maintenance time and expected cost rate of the capacity flow model is the same for the two structures. The similarity in the time and cost values indicates that capacity flow model assumes identical structures by ignoring the connections between the components. Furthermore, from Figures 5.9a and 5.10a, it can be noticed that the optimal t^*_{PM} decreases as the shock size ϕ increases in the two structures whereas the expected cost rates are increasing as the shock size increases.

A comparison of the spatial model's estimated optimal preventive maintenance time shows that a system with few component connections should be scheduled for maintenance earlier than a system with more component connections due to high level of non-uniform distribution of load caused by fewer links. In addition, comparing the expected cost rate of the two structures from the spatial model, it is concluded that a system with high level of non-uniform distribution of loads is more expensive to maintain. The result further indicates that the degree of component connections (spatial dependence) should be considered in maintenance modelling.

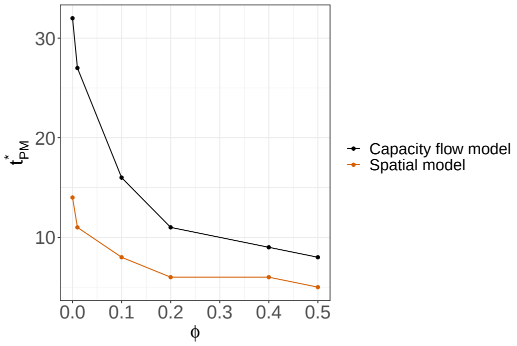
Moreover, compared with structure one, we can observe a closer gap between the



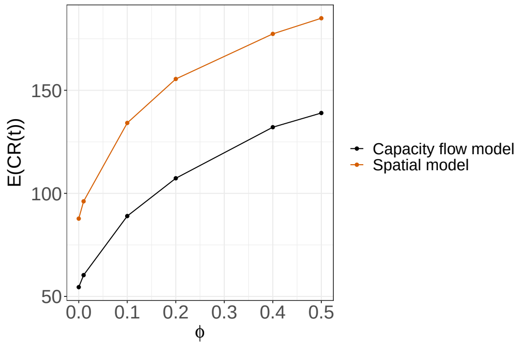
(a) Structure one (b) Structure two

(c) Block diagram of parallel connections

Figure 5.8: Four-component system structures

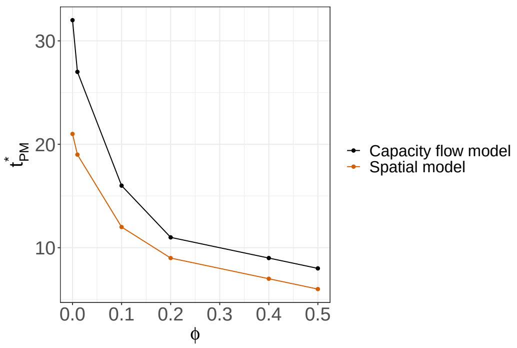


(a) Optimal preventive maintenance time

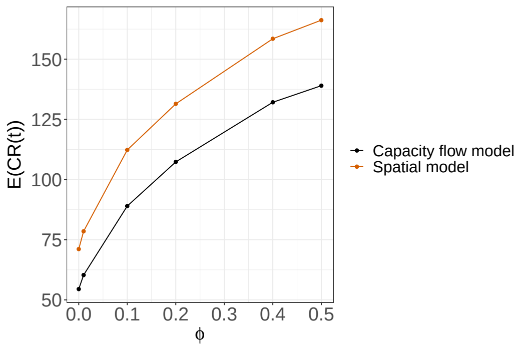


(b) Expected cost rate at t^*_{PM}

Figure 5.9: Comparison of the optimal preventive maintenance time and expected cost rate using the spatial model and capacity flow model considering structure one



(a) Optimal preventive maintenance time



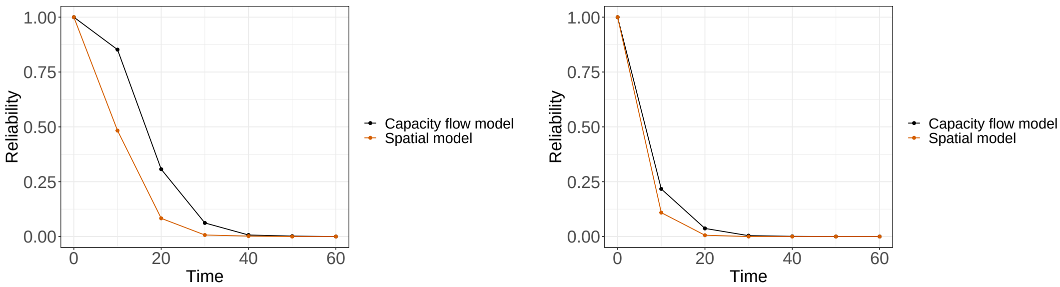
(b) Expected cost rate at t^*_{PM}

Figure 5.10: Comparison of the optimal preventive maintenance time and expected cost rate using the spatial model and capacity flow model considering structure two

optimal preventive maintenance and expected cost rate from the two models in structure two. The close gap could be because of the low level of non-uniform load distribution due to the presence of more connections between components in structure two compared with structure one that exhibits high level of non-uniform load distribution.

Compared to the numerical analysis of Isosceles system in section 5.4 where the capacity flow model is suitable for cases that the system is subject to high shock sizes, for structure one and structure two we observe a huge difference in the optimal preventive maintenance time and expected cost rate from the two models, indicating the poor performance of capacity flow models as the number of components increases.

Using algorithm 2, we compare reliability estimations of the spatial model with



(a) Shock size $\phi = 0.01$

(b) Shock size $\phi = 0.5$

Figure 5.11: Comparison of reliability estimation using spatial model and extension of the capacity flow considering the complex case

the extended capacity flow model for structure one. We examine the difference in the reliability estimation of the two models given that structure one is subject to load-sharing, spatial dependence and shocks. Figure 5.11 presents system reliability for cases when ϕ is 0.01 and 0.5. Figure 5.11a represents reliability estimations when $\phi = 0.01$ indicating that the shock size is low. When $\phi = 0.01$ the system reliability from the spatial model and capacity flow model hits zero at time $t = 50$ and $t = 60$ respectively. For the spatial model, the system reliability at times $t = 10$, 20, 30 and 40 are 0.483, 0.083, 0.007, and 0.002 respectively. In contrast, for the capacity flow model, the system reliability at times $t = 10$, 20, 30 and 40 are 0.852, 0.307, 0.062, and 0.007 respectively. Figure 5.11b represents reliability estimations when $\phi = 0.5$ indicating that the shock size is high. When $\phi = 0.5$ the system reliability from the spatial model and capacity flow model hits zero at time $t = 30$ and $t = 50$ respectively. For the spatial model, the system reliability at times $t = 10$, and $t = 20$ are 0.109, and 0.006 respectively. In contrast, for the capacity flow model, the system reliability at times $t = 10$, 20, 30 and 40 are 0.217, 0.037, 0.004, and 0.001 respectively.

It can be observed that when $\phi = 0.01$ there is difference between the reliability estimations of the two models and reliability estimations from the capacity flow model is twice those from the spatial model. The result indicates that when shock size is low, both models cannot be used as alternatives to the other. In this case, spatial dependence should not be ignored if it exists. When $\phi = 0.5$ the difference between the reliability estimations of the two models is reduced indicating that when shock size is increased, reliability estimations from both models become closer. In this case, spatial dependence could be ignored and a model that does not account for spatial dependence might be appropriate. The finding supports the earlier findings with respect to preventive maintenance optimization of structure one.

5.6 Numerical example: three-component systems with NHPP shocks

In this section we use the three-component system studied in section 5.4 to investigate the importance and significance of the spatial effect in the context that the system is operating in an environment modelled by NHPP of shocks. The system is referred to as an Isosceles structure. The complex case is considered here. The Isosceles structure is depicted in Figure 5.6. Note that the components in the system are connected in parallel, while the links between the components indicate spatial dependence. Components in the Isosceles system are linked together in a way that the system is composed of one dominant component (component 2) and two secondary components (component 1, and 3). As indicated by Figure 5.6, the system fails when the dominant component and one secondary component fail. In this case the load on the only working component is less than the load placed on the system. We assume that there is no information about the distance between components in the system so that the component arrangement will be used to infer component proximity.

A power law NHPP model is employed to model the rate of shocks as the rate of shocks is considered changing over time. Let the expected number of shocks be expressed using power law model:

$$V(t) = at^b. \tag{5.22}$$

A NHPP model with a power law intensity function can be written as follows:

$$P(N_s(t) = n) = \frac{(V(t))^n}{n!} exp^{(-V(t))}, n = 0, 1, 2, \ldots \tag{5.23}$$

where a is the scale parameter that stretch out or squeezes the distribution and b is the shape parameter that affects the general shape of the distribution. If b is not equal to 1, the behaviour of shocks is a non-stationary process as the shock rate changes with respect to time. That is, if $b > 1$, the shock rate increases with t, and if $b < 1$, otherwise.

We consider the same parameters as those used in section 5.4. The following parameters are used in the numerical analysis:

$$L = 3, \beta = 1, \lambda_i = 0.05, \phi = [0, 0.01, 0.1, 0.2, 0.4, 0.5], a = 1, b = [0.5, 1.5]$$

L is the load placed on the system. $b = 0.5$ indicates decreasing shock rate of shocks while $b = 1.5$ indicates increasing shock rate of shocks. $\beta = 1$ indicates the presence of load-sharing which has an impact on the failure rate. $\phi = 0$ assumes that the system is not affected by external shocks. When the system is not subject to external shocks, the system is similar to the load-sharing system studied in chapter 4. ϕ from 0.01 to 0.5 indicates low to high impact of shocks on the system. We assume that the cost of component replacement $c_R = 50$. Cost of corrective maintenance $c_{CM} = 1000$ and cost of preventive maintenance $c_{PM} = 400$. We present algorithm 4 to generate the expected cost rate when shocks are modelled by NHPP.

Algorithm 4 Algorithm for generating the expected cost rate of a load-sharing system with spatial dependence and shocks modelled by NHPP

Require: preventive maintenance time t_{PM}; Optimal preventive maintenance time t^*_{PM}; number of shocks n; shock arrival time X'_s; shock gap time X_s; component failure time t_j; cycle time corrective maintenance $CLCM$; cycle time preventive maintenance $CLPM$; simulation number SN; number of simulation runs SRN; number of failed components NF; number of working systems NWS; number of system failures NSF; number of failed components at corrective maintenance $NFCM$; number of failed components at preventive maintenance $NFPM$.

Ensure: Expected cost rate $E(CR(t_{PM}))$ at the end of the simulation.

1: set up the value of parameters: $t_{PM(max)}$, v, ϕ, L, λ_0, and β.

2: set $t_{PM} = 1$.

3: set $SN = 1$, $NF = 0$, $NWS = 0$, $NSF = 0$, $NFPM = 0$, $NFCM = 0$, $CLPM = 0$ and $CLCM = 0$.

4: generate a random number n from Pois($V(t)$) where $V(t) = at^b_{PM}$.

5: generate $X^{'}_s = t_{PM} \times U[0,1]^{\frac{1}{b}}$ where $s = 1, 2 \ldots, n$.

6: sort $X^{'}_s$ from 1 to n and calculate shock gap time $X_s = X^{'}_s - X^{'}_{s-1}$ where $s = 1, 2 \ldots, n$ and $X^{'}_0 = 0$.

7: set $p = 0$, $h = 0$ and $z_{i(j)} = 1$.

8: start at $k = 1$.

9: **repeat**

10: **for all** $i \in \{k, k+1, \ldots, m\}$ **do**

11: simulate m random variate u_i from $U[0,1]$.

12: calculate $t_i = \frac{-\ln(u_i)}{\lambda_{i(j^r)}}$ where $\lambda_{i(j^r)} = \lambda_0 z_{i(j)} + \phi p$.

13: **end for**

14: select the component with least failure time $t_j = min(t_k, t_{k+1}, \ldots, t_m)$.

15: set the next shock gap time $t_s = X_s$ in order from $s = 1$ to $s = n$.

16: select $T_k = min(t_i, t_s)$.

17: if $T_k = t_s$, then $NF = NF + 0$, and $p = p + 1$.

18: if $T_k = t_j$, then $NF = NF + 1$, $h = h + 1$ and recalculate $z_{i(j)}$ in $\lambda_{i(j^r)} = \lambda_0 z_{i(j)} + \phi p$ for proximate neighbours of failed component j.

19: $T_{sys} = \sum_k T_k$.

20: if $T_{sys} \geq t_{PM}$, then $NWS = NWS + 1$, $NFPM = NFPM + 1$, and $CLPM = CLPM + t_{PM}$ and go to step 24.

21: if $T_{sys} < t_{PM}$ and $\sum_{i=1}^{m-h} z_i = L$, set $k = k + 1$ then go to step 10.

22: if $T_{sys} < t_{PM}$ and $\sum_{i=1}^{m-h} z_i < L$, then $NSF = NSF + 1$, $NFCM = NFCM + 1$, and $CLCM = CLCM + t_{PM}$ and go to step 23.

23: if $SN < SRN$ then $SN = SN + 1$ and go to step 4. Else if $SN = SRN$ then go to step 24.

24: calculate $R_{sys}(t_{PM}) = \frac{\sum NWS}{SRN}$.

25: calculate $F_{sys}(t_{PM}) = \frac{\sum NSF}{SRN}$.

26: calculate $E(N_j(T) \mid T < t_{PM}) = \frac{\sum NFCM}{SRN}$.

27: calculate $E(N_j(T) \mid T \geq t_{PM}) = \frac{\sum NFPM}{SRN}$.

28: calculate $E(min(T, t_{PM})) = \frac{\sum(CLCM+CLPM)}{SRN}$.

29: calculate $E(CR(t_{PM})) = \frac{C_T(t_{PM})}{E(min(T,t_{PM}))}$, then go to step 30.

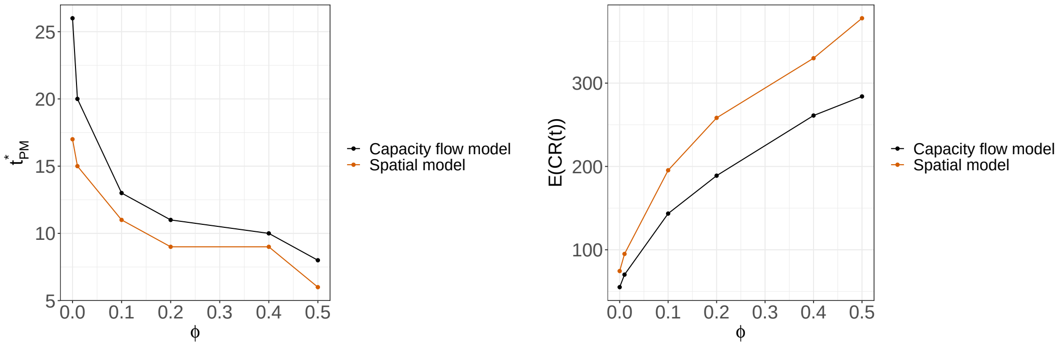
30: set $t_{PM} = t_{PM} + 1$.

31: **until** $t_{PM} < t_{PM(max)}$;

32: **return** $E(CR(t^*_{PM})) = min(E(CR(t_{PM})))$ and t^*_{PM} where $t_{PM} = 1, 2 \ldots, t_{PM(max)}$.

We use algorithm 4 in this section to estimate the system reliability quantities for the two models which were applied in the age replacement model. The results are generated from 1000 simulation runs. Consider when $b = 0.5$, the optimal preventive maintenance time and expected cost rate from the two models are presented in Figure 5.12a and 5.12b respectively. If spatial effect is neglected in the Isosceles system even though it exists and maintenance decisions are based on the capacity flow model due to existence of load-sharing, then the following outcomes are observed. If the system is operating in an environment where it is either not subject to shocks $\phi = 0$ or the impact of shocks on the system is very low $\phi = 0.01$, the optimal time for preventive maintenance is $t_{PM} = 26$ and $t_{PM} = 20$ while the cost rates are \$55.14 and \$70.16 respectively. If the impact of shocks on the system is high, that is from $\phi = 0.1$ to $\phi = 0.5$, the system is recommended for preventive maintenance ranging from $t_{PM} = 13$ to $t_{PM} = 8$ at a high cost rate ranging from \$143.44 to \$284.01. If maintenance decisions for the system are based on the developed spatial model, which captures the spatial effect, then the following outcomes will be observed. If the system is operating in an environment where it is either not subject to shocks $\phi = 0$ or the impact of shocks on the system is very low $\phi = 0.01$, the optimal time for preventive maintenance is $t_{PM} = 17$ and $t_{PM} = 15$ while the cost rates are \$74.46 and \$95.01 respectively. If the impact of shocks on the system is high, that is from $\phi = 0.1$ to $\phi = 0.5$, the system is recommended for preventive maintenance ranging from $t_{PM} = 11$ to $t_{PM} = 6$ at a high cost rate ranging from \$195.39 to \$377.85.

Consider when $b = 1.5$, the optimal preventive maintenance time and expected cost rate from the two models are presented in Figure 5.13a and 5.13b respectively. If

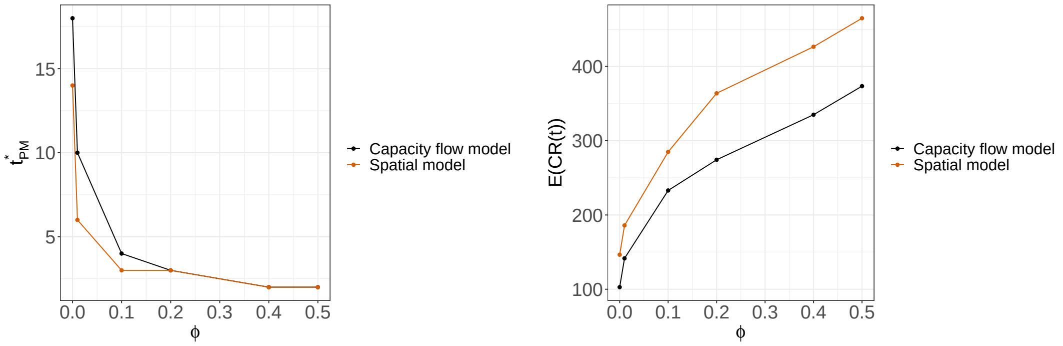


(a) Optimal preventive maintenance time (b) Expected cost rate at t^*_{PM}

Figure 5.12: Comparison of the optimal preventive maintenance time and expected cost rate using the spatial model and capacity flow model considering NHPP shocks with $b = 0.5$

spatial effect is not known to exist in the Isosceles system even though it exists and maintenance decisions are based on the capacity flow model due to existence of load-sharing, then the following outcomes are observed. If the system is operating in an environment where it is either not subject to shocks $\phi = 0$ or the impact of shocks on the system is very low $\phi = 0.01$, the optimal time for preventive maintenance is $t_{PM} = 18$ and $t_{PM} = 10$ while the cost rates are \$102.83 and \$141.60 respectively. If the impact of shocks on the system is high, that is from $\phi = 0.1$ to $\phi = 0.5$, the system is recommended for preventive maintenance ranging from $t_{PM} = 4$ to $t_{PM} = 2$ at a high cost ranging from \$232.99 to \$373.42. If maintenance decisions for the system are based on the developed spatial model, which captures the spatial effect, then the following outcomes will be observed. If the system is operating in an environment where it is either not subject to shocks $\phi = 0$ or the impact of shocks on the system is very low $\phi = 0.01$, the optimal time for preventive maintenance is $t_{PM} = 14$ and $t_{PM} = 6$ while the cost rates are \$146.51 and \$185.96 respectively. If the impact of shocks on the system is high, that is from $\phi = 0.1$ to $\phi = 0.5$, the system is recommended for preventive maintenance ranging from $t_{PM} = 3$ to $t_{PM} = 2$ at a high cost rate ranging from \$284.90 to \$464.88.

Similar to the HPP considered in section 5.4, when $b = 0.5$ and $b = 1.5$ the optimal t_{PM} derived based on the two models decreases as the shock size ϕ are increased,



(a) Optimal preventive maintenance time

(b) Expected cost rate at t^*_{PM}

Figure 5.13: Comparison of the optimal preventive maintenance time and expected cost rate using the spatial model and capacity flow model considering NHPP shocks with $b = 1.5$

indicating that it is better to schedule preventive maintenance faster for a system subject to increased shocks due to the cumulative damage on the system. Furthermore, the cost of maintenance from both models increases as impact of shocks increases.

When shock rate is increasing, that is $b = 1.5$, and the effects of shocks are high, the difference in the optimal preventive maintenance time from both models is not significant compared to when shock rate is decreasing $b = 0.5$. The difference in the optimal preventive maintenance time implies that when shock rate is increasing and the effects of shocks are high, the spatial effect can be ignored and simple equal load-sharing models like the capacity flow model would be suitable for modelling load-sharing and shocks. However, when the effect of shocks on the system is low and spatial effect exists then the model that accounts for spatial effect is suitable.

5.7 Discussion of results and Conclusion

In this chapter, we studied the problem of estimating the reliability of a load-sharing system with spatial dependence, proximity effects and external shock. We then investigated the impact of ignoring spatial dependence effect in maintenance optimization if it exists in a load-sharing system with shocks. We have assumed that a system exists that operates in a way that the load on a failed component is taken up only by its working spatial neighbours in close proximity and each working component is subject to external shocks that damage the components but does not cause immediate failure. The system fails when the sum of loads on all working components is less than system load. Age-replacement model was considered for the optimizing the replacement age of the system. We developed an additive failure rate model to evaluate component reliability. The failure rate model accounts for load-sharing interactions, proximity effects, and external shocks on each component. We developed an algorithm using Monte Carlo method to compute the system's reliability. We also considered a simple case where one external shock arrives before each component failure. We developed a method for computing the system's reliability using modified failure sequence diagram.

In the age-replacement model, preventive maintenance actions are performed on the system to either preventively renew the system at a pre-determined age t_{PM} or correctively maintain the system if the system fails at time T. Component failures were assumed to follow exponential distribution while the effect of shocks were assumed to be modelled by HPP and NHPP.

Numerical examples were used to demonstrate the developed method for the complex case on different configurations of a three-component system and a four-component system. We found first that the degree of component connections (spatial dependence) should be considered in maintenance modelling. The reason is that the level of component connections reflects the level of non-uniform or uniform load-sharing taking place in the system. A system with few component connections is more expensive to maintain compared with those with more connections due to a high level of non-uniform load-sharing. Second, when the shock size is high, the difference in the optimal preventive maintenance time from the developed spatial model and an extension of the capacity flow models is not significant for a system with small components. However, as components size increases, the difference increases, indicating that a model that does not account for spatial effects might be suitable for modelling a system with spatial effects when shock sizes are high and component sizes are low. Lastly, when the shock sizes are low and regardless of component size, a model that does not account for spatial effects when it exists would overestimate the time for preventive maintenance and underestimate the expected cost of maintenance of the system.

Appendix - Chapter 5

A.1 Appendix A

Simulation flowchart for generating the expected cost rate is presented below.

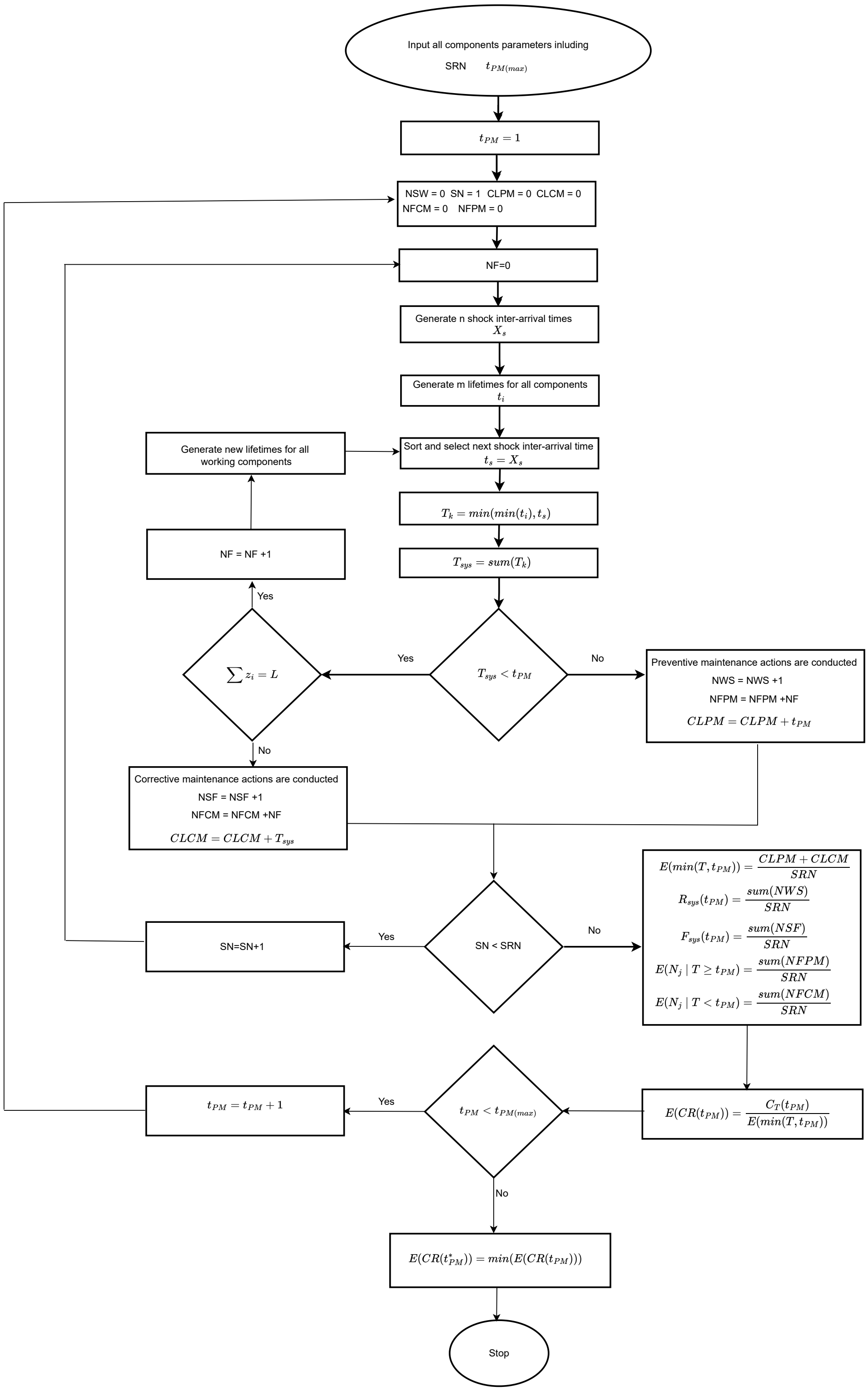


Figure 14: Simulation flowchart for generating the expected cost rate.

Chapter 6

Conclusion and future research

This chapter draws together the various strands of the thesis and summarises its key contributions. This thesis has developed new models to improve the accuracy of reliability prediction in multi-component systems subject to heterogeneity or load-sharing with spatial dependence. This was done by considering the limitations in existing reliability prediction models identified from the literature. These developments are summarised in section 6.1. The chapter will conclude with discussions on future research in section 6.2.

6.1 Summary and Conclusion

6.1.1 Reliability evaluation of a repairable multi-component system considering unit heterogeneity using frailty model

Reliability prediction of repairable systems subject to minimal repair and unobserved heterogeneity have been widely studied in recent years and models that account for unobserved heterogeneity when it exists in failure data have demonstrated improved reliability prediction compared to models that assume homogeneity. However, the literature on repairable systems subject to minimal repair did not previously include:

1. A parametric model for reliability prediction at component level. The ability to predict the occurrence of failure events at an individual unit level can aid maintenance decision making for individual components.

2. An assessment of the impact of misspecification of the random effect distribution with the minimal repair assumption. The choice of the random effect distribution for modelling unobserved heterogeneity effects can affect the validity of intensity function estimators.

3. Consideration of systems whose components become homogeneous overtime.

In chapter 3, we address each of these issues in turn. To address issue 3, an IG frailty model was developed for modelling unobserved heterogeneity in the failure processes of components in a multi-component system. The developed IG frailty model builds on the NHPP based gamma frailty model considered by Asfaw et al. [62], Slimacek and Lindqvist [67]. The IG frailty model, which combines the power law NHPP model and IG distributed frailties, assumes that the relative frailty distribution among surviving components becomes more homogeneous over time. This contrasts with the commonly used gamma frailty models which assume that the relative frailty distribution among surviving components is independent of age. We considered a repairable system whose components become homogeneous over time due to the burn in phase. We developed a parameter estimator for the IG frailty model using a maximum likelihood method. We examined the performance of the IG frailty model estimators using a simulation study and found that the estimator is robust to different parameter values, and that the estimators perform well for large sample sizes.

To address issue 2, the impact of misspecification of the random effect distribution with the minimal repair assumption was investigated. Using a simulation study, a comparison of the power law NHPP with gamma and IG distributed frailties was conducted to investigate the impact of wrongly assuming the random effect distribution for a repairable system subject to minimal repair. We examine the statistical fit of the gamma and IG frailty models as well as their prediction performance. We found that regardless of the degree of heterogeneity or the frequency of failures when early component behaviour is concerned, the probability of selecting a wrong model is low whether for model fit or for prediction purpose. The wrong model is only selected when the sample size is small.

To address issue 1, using empirical Bayes framework, we developed a method for prediction of a component's mean residual life and prediction of the expected number of failures at the component levels. The developed method uses Bayes theorem to update the frailty distribution for each component based on the observed data from the component. Using the developed empirical Bayes framework for component level prediction of expected number of failures and mean residual life based on air conditioner data from a set of airplanes. We found that the IG frailty model outperforms the gamma frailty model for component level prediction. We also compared the two frailty models and a power law NHPP model. Our results indicate that the IG frailty model is better in terms of model fit and outperforms the gamma for system level prediction.

6.1.2 Reliability analysis for load-sharing system with spatial dependence, and proximity effects

Models that account for load-sharing dependence for reliability prediction of systems subject to load-sharing have demonstrated improved reliability prediction compared to models that assume component independence. However, the existing literature on systems subject to load-sharing did not include:

1. A mathematical method to capture the effect of each load change on the failure rate of a working spatial neighbour for system reliability prediction. A model that captures the load-failure rate relationship can aid the assessment of factors that influence the system's reliability.

2. An investigation of the importance and significance of accounting for the spatial effects in load-sharing system reliability prediction. One of the significant questions that this thesis answers in chapter 4 is what is the impact of ignoring spatial dependence if it exists in a load-sharing system?

In chapter 4, we address both of these gaps in the literature. To address gap 1, we developed a model for reliability analysis of load-sharing systems subject to spatial dependence and proximity effect. The model was developed considering homogeneous and heterogeneous components. The capacity flow model was extended to capture spatial dependence and proximity effects in the failure rate function. Closed form expressions for component load function and failure rate were derived. We considered that the system operates in a way that the load on a failed component is taken up only by its working spatial neighbours in proximity. Each component's failure rate was considered changing as a function of extra load taken from a failed spatial neighbour and the load-failure rate relationship was modelled by extending the capacity flow model. The traditional capacity flow model considers that the load of a failed component is shared equally by all the working components. Markov model was used to characterise the deterioration process of the entire system. Sensitivity analysis of the load factor was conducted to examine its effect on reliability estimation. We found that an increased load on the system reduces the system reliability faster. The mathematical formulation for parameter estimation procedure was developed.

To address gap 2, the importance and significance of accounting for the spatial effect was investigated. Using numerical experiment, a comparison of the extended capacity flow model and traditional capacity flow model was conducted to investigate the impact of ignoring spatial dependence if it exists in a load-sharing system. Monte Carlo simulation was used to generate cumulative distribution function of the system and to validate the reliability estimation from the compared models. We found that when heterogeneous components are considered for a four-component system with few component connections, the spatial effect may be ignored and a simpler model such as the model for 3-out-of-4 systems which does not account for spatial effects could be applied. However, when homogeneous components are considered, reliability prediction of a load-sharing system based on capacity flow model can be overestimated or underestimated if spatial effect exists. A similar result was found for a five-component system with complex connections. Overall, we infer that considering homogeneous components, regardless of the level of component connections and system size, if it exists, spatial effect should not be ignored in reliability prediction of a system with load-sharing.

6.1.3 Reliability modelling and preventive maintenance of load-sharing systems with spatial dependence, proximity effects and external shocks

Chapter 5 extends the work in chapter 4 by considering maintenance and external shocks. Like chapter 4, literature on systems subject to load-sharing and shocks did not include:

1. A model for reliability prediction of systems subject to load-sharing, spatial dependence, and shocks.
2. A model for maintenance optimization of load-sharing system with spatial dependence, proximity effect and external shocks.
3. An investigation of the importance and significance of accounting for the spatial dependence in load-sharing system with external shocks. One of the significant questions that this thesis answers in chapter 5 is what is the impact of ignoring spatial dependence if it exists in a load-sharing system with external shocks?

In chapter 5, we address each of these issues in turn. To address issue 1, we developed two methods for reliability analysis of load-sharing systems subject to spatial dependence, proximity effect and external shocks. We considered that the system operates in a way that the load on a failed component is taken up only by its working spatial neighbours in proximity. Each component's failure rate is considered changing as a function of extra load taken from a failed spatial neighbour and external shocks. We assume that external shocks damage the components but does not cause immediate failure. Using Monte Carlo method, we developed an algorithm to model the system's reliability. For the developed model, an additive failure rate model was derived as a function of the load, baseline failure rate and shocks for each component. External shocks on the components were modelled by HPP and NHPP models respectively. Sensitivity analysis of the shock size was conducted to examine its effect on reliability estimation. We found that an increased load on the system reduces the system reliability faster. We also considered a simple case where one external shock arrives before each component failure. We developed a method for computing the system's reliability using modified failure sequence diagram.

To address issue 2, we developed an age-replacement model as an extension of traditional age-replacement model. To address issue 3, we assessed the impact of ignoring spatial dependence in a load-sharing system with external shocks using the developed algorithm. A comparison of the optimal maintenance time and expected cost based on the developed spatial model and an extended capacity flow model was conducted. The extended capacity flow model incorporates the effect of external shocks and considers that the load of a failed component is shared equally by all the working components. When we varied the parameter controlling the shock size, we found that when shock size is high, the difference in the optimal preventive maintenance time from the developed spatial model and an extension of the capacity flow models is not significant for a system with small components. However, as components size increases, the difference increases. Indicating that a model that does not account for spatial effects might be suitable for modelling a system with spatial effects when shock sizes are high and component sizes are low. In addition, when shock sizes are low and regardless of component size, a model that does not account for spatial effects when it exists would overestimate the time for preventive maintenance and underestimate the expected cost of maintenance of the system. Lastly, the number of component connections should be considered in maintenance modelling. The reason is that the level of component connections reflects the level of non-uniform or uniform load-sharing taking place in the system. A system with few component connections is more expensive to maintain compared with those with more connections due to a high level of non-uniform load distribution.

6.2 Future Research

In the process of completing the research objectives described in chapter 1, several avenues of research were considered. Some of these were considered superficially, while others were considered in more detail. In section 6.2.1, we describe one of the areas that was explored in detail. In section 6.2.2, we explore areas of research that were only briefly considered but exist as useful avenues to explore.

6.2.1 Reliability analysis for systems with failure interaction, spatial dependence, and proximity effects

In chapter 4, we developed a model for reliability prediction of a load-sharing system subject to spatial dependence and proximity effects. However, there is no model yet in the literature for reliability prediction of systems subject to local failure interaction considering proximity effects. Consider a system in which one component performs a crucial function (Dominant component) and others performs less crucial function (Secondary components). The failure behaviour of such a system is such that whenever any secondary component has failed, its failure causes an increase in the deterioration rate of the dominant component. In addition, the deterioration rate of the dominant component is only influenced by the failed state of the secondary components that are its spatial neighbours.

We present an example to illustrate this situation. An example of such system is a water distribution system, in which the repair of failed distribution pipes causes pressures that could increase the deterioration in a nearby main pipe while a service pipe's state has no effect on the main pipe's deterioration [233].

Others such as [227] and [233] have studied these systems. However, we explored this topic in two ways. First, whilst previous works assume type I interaction, i.e, that the failure of the influencing component leads to the immediate failure of the affected component, we could assume that the failure of the influencing component only increases the deterioration rate of the affected component. Unlike [227] that assumes the failure of any secondary component will affect the dominant component, we assume that the failure of only nearby secondary components will affect the dominant component. As nearby secondary components in the system will affect the deterioration of the primary component, a proportional hazard model (PHM) could be used to describe the relationship between the influencing components and the influenced component. Markov model could be used to describe the deterioration process of the entire system. To illustrate, here we will formulate the failure rate model for the dominant component before describing future work.

Consider a multi-component system with local failure interaction. The system consists of one primary and $m+n$ secondary components where n secondary component are spatial neighbours of the primary component and m components are non neighbours. The failure of the j^{th} secondary components $j = 1, 2, \ldots, m + n$, does not make the system stop, but can reduce the system's performance while the failure of the primary component can cause the system to stop functioning.

The failure interaction between components of this system is as follows. Failure of the j^{th} secondary component, $j = 1, 2, \ldots, m+n$, has no effect on the behaviour of the other secondary component. However, failure of a nearby secondary component acts as a shock to the primary component without inducing an instantaneous failure, but increasing its failure rate. We assume that the failure of the dominant and secondary component occurs according to a HPP with constant failure rate.

In addition to the above assumptions, we make the following assumptions:

1. The j^{th} secondary component starts as good as new, i.e., the initial age of this component is zero.
2. We assume that when a secondary component fails, shocks from it increase the failure rate of the primary component within a distance threshold α by ϑ_{ij} percent.
3. ϑ_{ij} is assumed to be time-invariant.

If we assume that shocks from the j^{th} secondary components would increase the failure rate of the primary component by ϑ_{ij} percent and each ω_{ij} is a function of the distance between secondary component j^{th} (influencing component) and primary component (affected component).

Given the number of failures of each secondary component, and the percentage of shock ϑ_{ij} transferred from them to a surviving primary component, the conditional failure rate for the primary component p is given by:

$$\begin{aligned} h_p^{con}(t) &= h_p(t \mid N_1(t) = k_1, N_2(t) = k_2, \ldots, N_{m+n}(t) = k_{m+n}) \\ &= h_p^0(t)(1+\vartheta_{p1})^{k_1}(1+\vartheta_{p2})^{k_2} \ldots (1+\vartheta_{p(m+n)})^{k_{(m+n)}} \end{aligned}$$

$$h_p^{con}(t) = h_p^0(t) \prod_{j=1}^{m+n} (1+\vartheta_{pj})^{k_j} \quad , \tag{6.1}$$

where $h_p^{con}(t)$ is the conditional failure rate of the primary component. $h_p^0(t)$ is the baseline failure rate for the primary component. k_j is the number of failures of the j^{th} secondary component. ϑ_{pj} represents the degree of the effect a secondary component has on the primary component and $0 \leq \vartheta_{pj} < 1$. If $\vartheta_{pj} = 0$, then the failure of a secondary component has no effect on the primary component.

For simplicity, we assume that the secondary components are independent of each other. The probability of having $N_j(t)$ failures of the j^{th} secondary component at time t is:

$$Pr(N_j(t)) = \frac{h_j(t)^{k_j} exp^{-h_j(t)}}{k_j!}. \tag{6.2}$$

The expected failure rate for the primary component at a time t is given by:

$$\begin{aligned} h_p(t) = h_p^{con}&(t \mid N_1(t) = k_1, \ldots, N_{m+n}(t) = k_{m+n}) Pr(N_1(t) = k_1), \ldots, \\ &Pr(N_{(m+n)}(t) = k_{(m+n)}) \end{aligned}$$

$$h_p(t) = \sum_{k_1}^{\infty} \ldots \sum_{k_{(m+n)}}^{\infty} h_p^{con} Pr(N_1(t)) \ldots Pr(N_{(m+n)}(t)). \tag{6.3}$$

Thus the failure rate model for the primary component is

$$h_p(t) = h_p^0(t) exp^{\left(\vartheta_{p1}h_1^0(t)+\vartheta_{p2}h_2^0(t)+,\ldots,+\vartheta_{p(m+n)}h_{(m+n)}^0(t)\right)} \quad , \tag{6.4}$$

where $h_j^0(t)$ for $j = 1, \ldots, m+n$ is the baseline failure rate of the secondary components.

Since all the m secondary components are not close enough to the primary component to affect, $\vartheta_{pj} = 0$ for each m^{th} component thus the failure rate of the primary component is only affected by all n neighbour secondary components given by:

$$h_p(t) = h_p^0(t)exp^{\left(\vartheta_{p1}h_1^0(t)+\vartheta_{p2}h_2^0(t)+,\ldots,+\vartheta_{p(n)}h_{(n)}^0(t)\right)} \quad , \tag{6.5}$$

where Eq.(6.5) indicates that the failure rate of the primary component is greater than its own independent failure rate because $\vartheta_{pn} > 0$.

The failure distribution function of the primary component is given by

$$F_p(t) = 1 - exp^{\left\{\int_0^t h_p(x)dx\right\}}. \tag{6.6}$$

Since we assume the components are spatially dispersed, ϑ_{pj} could be modelled by a distance function. We assume that the distance between all secondary components and the primary component is an important criterion of spatial influence between the components. If we assume that there is a threshold distance beyond which there is no direct spatial influence between the primary and secondary components then we can apply the constant proximity model introduced in section 4.2.2 of chapter 4 to derive ϑ_{pj} by:

$$\vartheta_{pj} = \begin{cases} c, & 0 \leq d_{pj} \leq \alpha \\ 0, & \quad d_{pj} > \alpha \end{cases} \quad , \tag{6.7}$$

where c is a predefined constant value and its the same for all pair of primary and n secondary components within the threshold distance α while all other m components outside the threshold distance have $\vartheta_{pj} = 0$. In this work, we assume that all n secondary components have a predefined constant value $\vartheta_{pj} = c$.

If we assume that the each secondary component has varied effect on the primary component, we could model the influence by an exponential model introduced in section 4.2.2 of chapter 4. The exponential model assumes a diminishing effect, in which ϑ_{pj} reduces as the distance between the primary and secondary component increases up to a threshold α (see section 4.2.2 of chapter 4 for more details on the exponential model).

In order to describe the behaviour of the load-sharing system with proximity and spatial dependence, we will apply the aforementioned failure rate model in a Markov model. The failure rate model will be used to account for the effect of the influencing secondary components on the transition rates of the primary component in the Markov model. As introduced in chapter 4, Markov models are frequently used in reliability and maintainability work where events, such as the failure or repair of a component, can occur at any point in time. The model assumes that the conditional distribution of a future state is independent of the past states of the process i.e., the behaviour of the system in each state is memory less [145]. Thus, the sojourn time of each state is exponentially distributed and the transition probability to each state is independent of the process history. For any given system, a Markov model consists of a list of the possible states of that system, the possible transition paths between those states, and the rate parameters of those transitions.

Let λ denote the rate parameter of the transition from State 1 to State 2 and $S_i(t)$ be the probability of the system being in State i at time t. An advantage of Markov models is that they are simple to generate even though they require a complicated mathematical approach. Consider a voting system with one primary component, two influencing secondary components and one non-influencing secondary components. Using Figure 6.1, the state transition for the system will be described. We start by defining each state $S(t)$ of the system. When the system is at state S_0, all components in the system are working (functional). When the system is in state S_7-S_{13}, the primary component has failed as a result the system no longer functions. States S_1-S_6 is the system state when any of the secondary components have failed. When the system is in state S_6, we assume that the system has partially failed, as such the system's performance is reduced.

The system's evolution is determined by the transitions among states. We shall discuss the state transitions of the system in the following way. Denote the transition rate of the influencing and non-influencing secondary components by λ_1^* and λ_1 respectively.

The transition from state S_0 to S_1 means that one of the influencing secondary components have failed with transition rate $2\lambda_1^*$ while transition from state S_0 to S_2 means that the functional non-influencing component has now failed with transition rate λ_1. The transition from state S_1 to S_3 or S_4 implies that a second influencing or

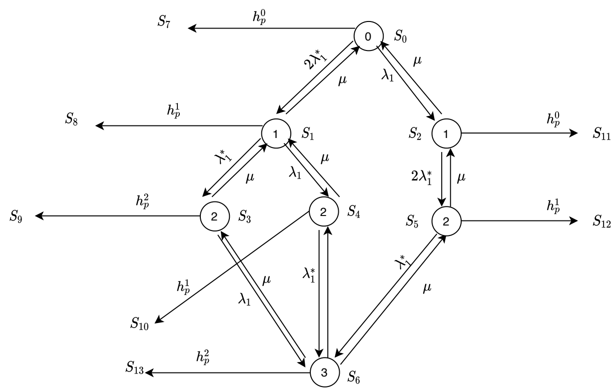


Figure 6.1: State transition for a system with one primary component, two influencing secondary components and one non-influencing secondary components

non-influencing component has failed with transition rate λ_1^* or λ_1. This procedure is continued until the last working secondary components have failed.

When r influencing secondary component is in failed state, i.e states, the transition rate of the primary component is then given by h_p^r where r is the number of failed influencing (secondary) components. For example, transition from state S_1 to S_8 means that the primary component has failed with failure rate h_p^1.

In order to fully develop this modelling, we could do the following three things:

- We could consider a more complex spatial model than the discrete threshold considered here.
- Here, we have only considered a single primary component. Further work would be required to expand this to m primary components.
- Numerical analysis is required to evaluate the accuracy of the model.

However, we have laid the ground work for future analysis to build upon.

6.2.2 Areas for further research

Here, we consider future research around the contributions of each chapter.

6.2.2.1 Future Research in Chapter 3

In chapter 3 we studied reliability prediction for a system subject to unobserved heterogeneity and developed a method based on empirical Bayes for prediction at the component level. However, the effect of covariates was not captured in the modelling. One area to consider for further work would be to extend the developed prediction method to include covariate effect. A component level prediction method that captures both observed and unobserved factors of heterogeneity has not been studied in the literature for system subject to minimal repair. A method that combines both observed and unobserved covariates would improve the accuracy of reliability prediction if covariates were recorded for each component in the system. Also, the developed prediction framework could be extended considering systems subject to imperfect repairs. The impact of misspecification of a random effect distribution could be investigated considering systems subject to imperfect repairs.

6.2.2.2 Future Research in Chapter 4

We studied reliability prediction for a load-sharing system with spatial dependent components and proximity effect in chapter 4. The components in this study only have two states. For further work, a system with multi-state components can be considered. In addition, the impact of ignoring spatial effect can be investigated for systems with multi-state components. Considering systems with binary state components, the impact of ignoring spatial effect can also be further explored for a large system with heterogeneous components. Also, it would be valuable to develop an efficient method suitable for reliability prediction of systems with large components. Furthermore, a more elaborate proximity model could be considered.

6.2.2.3 Future Research in Chapter 5

We studied reliability modelling and preventive maintenance of a load-sharing system with spatial dependent components, proximity effect, and external shocks. We considered an optimal age-replacement policy for the load-sharing system in chapter 5. For future work, an optimal inspection policy can be considered for the system and the policy can be extended to the case where the maintenance action required for a failed component is also a decision variable. In this case, at an inspection time, one could decide the best course of action for each failed component (i.e replacement or minimal repair), and also determine the next optimal inspection time. Also, one could consider non-periodic inspection policy.

Bibliography

[1] J. H. Saleh and K. Marais, "Highlights from the early (and pre-) history of reliability engineering," *Reliability engineering & system safety*, vol. 91, no. 2, pp. 249–256, 2006.

[2] E. Zio, "Reliability engineering: Old problems and new challenges," *Reliability Engineering and System Safety*, vol. 94, no. 2, pp. 125–141, 2009.

[3] D. Lea, V. Bull, S. S. Webb, and R. Duncan, *Oxford learner's dictionary of academic English.* Oxford University Press, 2014.

[4] J. McLinn, "A short history of reliability," *The Journal of the Reliability Information Analysis Center*, pp. 8–15, 2011.

[5] C. E. Ebeling, *An introduction to reliability and maintainability engineering.* Tata McGraw-Hill Education, 2004.

[6] I. Bazovsky, *Reliability theory and practice.* Courier Corporation, 2004.

[7] S. S. Bhamare, O. P. Yadav, and A. Rathore, "Evolution of reliability engineering discipline over the last six decades: a comprehensive review," *International Journal of Reliability and Safety*, vol. 1, no. 4, pp. 377–410, 2007.

[8] M. Revie, *Evaluation of bayes linear modelling to support reliability assessment during procurement.* PhD thesis, Strathclyde, 2008.

[9] L. Condra, *Reliability improvement with design of experiment.* Crc Press, 2001.

[10] "Product excellence using 6 sigma (peuss)." `https://www.scribd.com/document/248666611/Introduction-to-Reliability-WARWICK-MANUFACTURING-GROUP`, note = Accessed: 5-3-2020,.

[11] W. G. Ireson, C. F. Coombs Jr, and R. Y. Moss, *Handbook of reliability engineering and management 2/E.* McGraw Hill Professional, 1996.

[12] L. Walls, J. Quigley, and J. Marshall, "Modeling to support reliability enhancement during product development with applications in the uk aerospace industry," *IEEE Transactions on Engineering Management*, vol. 53, no. 2, pp. 263–274, 2006.

[13] S. Taghipour and M. L. Kassaei, "Periodic inspection optimization of a k-out-of-n load-sharing system," *IEEE Transactions on Reliability*, vol. 64, no. 3, pp. 1116–1127, 2015.

[14] A. Coppola, "Reliability engineering of electronic equipment a historical perspective," *IEEE Transactions on Reliability*, vol. 33, no. 1, pp. 29–35, 1984.

[15] F. Asgari and M. R. Lotfi, "Interactive failures and its effect on the predicted reliability," *Jurnal UMP Social Sciences and Technology Management Vol*, vol. 3, no. 3, 2015.

[16] Y. Sun, *Reliability prediction of complex repairable systems: an engineering approach.* PhD thesis, Queensland University of Technology, 2006.

[17] J. Faulin, A. A. Juan, S. S. M. Alsina, and J. E. Ramirez-Marquez, *Simulation methods for reliability and availability of complex systems.* Springer Science & Business Media, 2010.

[18] Z. Vintr and D. Valis, "A tool for decision making in k-out-of-n system maintenance," in *Applied Mechanics and Materials*, vol. 110, pp. 5257–5264, Trans Tech Publ, 2012.

[19] P. O'Connor and A. Kleyner, *Practical reliability engineering.* John Wiley & Sons, 2012.

[20] T. Economou, *Bayesian modelling of recurrent pipe failures in urban water systems using non-homogeneous Poisson processes with latent structure.* PhD thesis, 2010.

[21] V. M. S. Hussain and V. Naikan, *Point Process Based Maintenance Modeling for Repairable Systems: A Review.* Citeseer, 2010.

[22] X. Zhang, R. Xiu, C. Chen, and Y. Liu, "Reliability assessment for multi-state systems under a dynamic load-sharing strategy," in *2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE)*, pp. 367–371, IEEE, 2013.

[23] B. Liu, J. Wu, and M. Xie, "Cost analysis for multi-component system with failure interaction under renewing free-replacement warranty," *European Journal of Operational Research*, vol. 243, no. 3, pp. 874–882, 2015.

[24] V. Slimacek and B. H. Lindqvist, "Nonhomogeneous Poisson process with nonparametric frailty and covariates," *Reliability Engineering & System Safety*, vol. 167, pp. 75–83, 2017.

[25] J. Ansell and L. Walls, "Dependency analysis in reliability studies," *IMA Journal of Management Mathematics*, vol. 3, no. 4, pp. 333–348, 1991.

[26] L. Thomas, "A survey of maintenance and replacement models for maintainability and reliability of multi-item systems," *Reliability Engineering*, vol. 16, pp. 297–309, jan 1986.

[27] 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*, vol. 261, no. 2, pp. 405–420, 2017.

[28] M. Marseguerra, E. Zio, and L. Podofillini, "Condition-based maintenance optimization by means of genetic algorithms and monte carlo simulation," *Reliability Engineering & System Safety*, vol. 77, no. 2, pp. 151–165, 2002.

[29] 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*, vol. 195, pp. 319–327, 2018.

[30] Z. Zhang, S. Wu, S. Lee, and J. Ni, "Modified iterative aggregation procedure for maintenance optimisation of multi-component systems with failure interaction," *International Journal of Systems Science*, vol. 45, no. 12, pp. 2480–2489, 2014.

[31] Z. Zhang, S. Wu, B. Li, and S. Lee, "(n, n) type maintenance policy for multi-component systems with failure interactions," *International Journal of Systems Science*, vol. 46, no. 6, pp. 1051–1064, 2015.

[32] A. V. Suprasad, M. B. Krishna, and P. Hoang, "Tampered failure rate load-sharing systems: Status and perspectives," in *Handbook of performability engineering*, pp. 291–308, Springer, 2008.

[33] W. Y. Yun, G. R. Kim, and H. Yamamoto, "Economic design of a load-sharing consecutive k-out-of-n: F system," *IIE Transactions*, vol. 44, no. 1, pp. 55–67, 2012.

[34] P. Do, P. Scarf, and B. Iung, "Condition-based maintenance for a two-component system with dependencies," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 946–951, 2015.

[35] Z. Liang and A. K. Parlikad, "A tiered modelling approach for condition-based maintenance of industrial assets with load sharing interaction and fault propagation," *IMA Journal of Management Mathematics*, vol. 26, no. 2, pp. 125–144, 2015.

[36] S. V. Amari, K. B. Misra, and H. Pham, "Reliability analysis of tampered failure rate load-sharing k-out-of-n: G systems," in *Proc. 12th ISSAT Int. Conf. on Reliability and Quality in Design, Honolulu, Hawaii*, pp. 30–35, Citeseer, 2006.

[37] S. V. Amari and R. Bergman, "Reliability analysis of k-out-of-n load-sharing systems," in *2008 Annual Reliability and Maintainability Symposium*, pp. 440–445, IEEE, 2008.

[38] L. Wang, J. Zhang, W. Chen, and X. Jia, "Reliability evaluation of a load-sharing parallel system with failure dependence," *Communications in Statistics-Simulation and Computation*, vol. 45, no. 9, pp. 3094–3113, 2016.

[39] J. Zhang, Y. Zhao, and X. Ma, "Reliability modeling methods for load-sharing k-out-of-n system subject to discrete external load," *Reliability Engineering and System Safety*, vol. 193, p. 106603, 2020.

[40] H. Kim and P. H. Kvam, "Reliability estimation based on system data with an unknown load share rule," *Lifetime Data Analysis*, vol. 10, no. 1, pp. 83–94, 2004.

[41] C. Park, "Parameter estimation for the reliability of load-sharing systems," *IIE Transactions (Institute of Industrial Engineers)*, vol. 42, no. 10, pp. 753–765, 2010.

[42] N. Zhang, M. Fouladirad, and A. Barros, "Maintenance analysis of a two-component load-sharing system," *Reliability Engineering & System Safety*, vol. 167, pp. 67–74, 2017.

[43] X. Zhao, B. Liu, and Y. Liu, “Reliability Modeling and Analysis of Load-Sharing Systems with Continuously Degrading Components,” *IEEE Transactions on Reliability*, vol. 67, no. 3, pp. 1096–1110, 2018.

[44] G. Asha, A. V. Raja, and N. Ravishanker, “Reliability modelling incorporating load share and frailty,” *Applied Stochastic Models in Business and Industry*, vol. 34, no. 2, pp. 206–223, 2018.

[45] D. Wang, C. Jiang, and C. Park, “Reliability analysis of load-sharing systems with memory,” *Lifetime Data Analysis*, vol. 25, no. 2, pp. 341–360, 2019.

[46] J. Zhang, Y. Zhao, and X. Ma, “A new reliability analysis method for load-sharing k-out-of-n: F system based on load-strength model,” *Reliability Engineering & System Safety*, vol. 182, pp. 152–165, 2019.

[47] C. P. De Paula, L. B. Visnadi, and H. F. de Castro, “Multi-objective optimization in redundant system considering load sharing,” *Reliability Engineering & System Safety*, vol. 181, pp. 17–27, 2019.

[48] M. Franco, J. M. Vivo, and D. Kundu, “A generalized Freund bivariate model for a two-component load sharing system,” *Reliability Engineering and System Safety*, vol. 203, p. 107096, 2020.

[49] M. Sharifi, S. Taghipour, and A. Abhari, “Inspection interval optimization for a k-out-of-n load sharing system under a hybrid mixed redundancy strategy,” *Reliability Engineering & System Safety*, vol. 213, p. 107681, 2021.

[50] J. Guo, Y. Shen, Z. Lu, H. Che, Z. Liu, and S. Zeng, “Reliability modeling for consecutive k-out-of-n: F systems with local load-sharing components subject to dependent degradation and shock processes,” *Quality and Reliability Engineering International*, vol. 36, no. 5, pp. 1553–1569, 2020.

[51] M. F. Goodchild, “Geographical information science,” *International journal of geographical information systems*, vol. 6, no. 1, pp. 31–45, 1992.

[52] L. Wang, Q. Yang, and Y. Tian, “Reliability analysis of 6-component star markov repairable system with spatial dependence,” *Mathematical Problems in Engineering*, vol. 2017, 2017.

[53] L. Wang and S. Si, “Reliability analysis of circular markov repairable systems with spatial dependence,” *Journal of northwestern polytechnical university*, vol. 32, no. 6, pp. 923–928, 2014.

[54] L. Wang, Y. Tian, and Z. Pei, “Reliability analysis of 6-component lattice load-sharing Markov repairable system with spatial dependence,” *International Journal of Performability Engineering*, vol. 13, no. 3, pp. 279–287, 2017.

[55] Z. Zhang, X. Si, C. Hu, and X. Kong, “Degradation modeling–based remaining useful life estimation: A review on approaches for systems with heterogeneity,” *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 229, no. 4, pp. 343–355, 2015.

[56] J. Lin and M. AspLund, “Comparison study of heavy haul locomotive wheels’ running surfaces wearing,” *Eksploatacja i Niezawodność*, vol. 16, no. 2, pp. 276–287, 2014.

[57] J. Lin and M. Asplund, "Bayesian semi-parametric analysis for locomotive wheel degradation using gamma frailties," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 229, no. 3, pp. 237–247, 2015.

[58] J. Lin, J. Pulido, and M. Asplund, "Reliability analysis for preventive maintenance based on classical and Bayesian semi-parametric degradation approaches using locomotive wheel-sets as a case study," *Reliability Engineering & System Safety*, vol. 134, no. 3, pp. 143–156, 2015.

[59] A. Ghomghale, M. Ataei, R. Khalokakaie, A. Barabadi, and A. Nouri Qarahasanlou, "The application of frailty model in remaining useful life estimation (case study: Sungun copper mine's loading system)," *Journal of Modeling in Engineering*, vol. 18, no. 62, pp. 129–142, 2020.

[60] R. Barabadi, M. Ataei, R. Khalokakaie, and A. Nouri Qarahasanlou, "Spare-part management in a heterogeneous environment," *Plos one*, vol. 16, no. 3, p. e0247650, 2021.

[61] R. Zaki, A. Barabadi, J. Barabady, and A. Nouri Qarahasanlou, "Observed and unobserved heterogeneity in failure data analysis," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 236, no. 1, pp. 194–207, 2022.

[62] Z. G. Asfaw and B. H. Lindqvist, "Unobserved heterogeneity in the power law nonhomogeneous poisson process," *Reliability Engineering & System Safety*, vol. 134, pp. 59–65, 2015.

[63] B. H. Lindqvist, G. Elvebakk, and K. Heggland, "The trend-renewal process for statistical analysis of repairable systems," *Technometrics*, vol. 45, no. 1, pp. 31–44, 2003.

[64] X. Wang, "Wiener processes with random effects for degradation data," *Journal of Multivariate Analysis*, vol. 101, no. 2, pp. 340–351, 2010.

[65] C.-Y. Peng, "Inverse gaussian processes with random effects and explanatory variables for degradation data," *Technometrics*, vol. 57, no. 1, pp. 100–111, 2015.

[66] A. D'Andrea, C. C. Feitosa, and V. Tomazella, "Frailty modeling for repairable systems with minimum repair : An application to dump truck data of a brazilian mining company," 2017.

[67] V. Slimacek and B. Lindqvist, "Reliability of wind turbines modeled by a poisson process with covariates, unobserved heterogeneity and seasonality," *Wind energy*, vol. 19, no. 11, pp. 1991–2002, 2016.

[68] H. Yin, X. Yang, and R. Peng, "Generalized accelerated failure time frailty model for systems subject to imperfect preventive maintenance," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[69] A. Deep, D. Veeramani, and S. Zhou, "Event prediction for individual unit based on recurrent event data collected in teleservice systems," *IEEE Transactions on Reliability*, vol. 69, no. 1, pp. 216–227, 2019.

[70] S. Jahani, S. Zhou, D. Veeramani, and J. Schmidt, "Multioutput gaussian process modulated poisson processes for event prediction," *IEEE Transactions on Reliability*, vol. 70, no. 4, pp. 1569–1580, 2021.

[71] O. Aalen, O. Borgan, and H. Gjessing, *Survival and event history analysis: a process point of view.* Springer Science & Business Media, 2008.

[72] J. F. Lawless, "Regression Methods for Poisson Process Data," *Journal of the American Statistical Association*, vol. 82, pp. 808–815, sep 1987.

[73] L. Hsu, M. Gorfine, and K. Malone, "On robustness of marginal regression coefficient estimates and hazard functions in multivariate survival analysis of family data when the frailty distribution is mis-specified," *Statistics in Medicine*, vol. 26, no. 25, pp. 4657–4678, 2007.

[74] V. Slimacek and B. H. Lindqvist, "Nonhomogeneous poisson process with nonparametric frailty," *Reliability Engineering & System Safety*, vol. 149, pp. 14–23, 2016.

[75] J. Lawless and M. Crowder, "Covariates and Random Effects in a Gamma Process Model with Application to Degradation and Failure," *Lifetime Data Analysis*, vol. 10, no. 3, pp. 213–227, 2004.

[76] *Condition-based maintenance policies for multi-component systems considering stochastic dependences.* PhD thesis, 2016.

[77] H. Ascher and H. Feingold, *Repairable systems reliability: modeling, inference, misconceptions and their causes.* M. Dekker New York, 1984.

[78] P. Economou and C. Caroni, "Graphical Tests for the Assumption of Gamma and Inverse Gaussian Frailty Distributions," *Lifetime Data Analysis*, vol. 11, no. 4, pp. 565–582, 2005.

[79] W. Nelson, *Accelerated testing: statistical models, test plans and data analyses.* John Wiley Sons, 2009.

[80] K. Murphy, C. Carter, and S. Brown, "The exponential distribution: the good, the bad and the ugly. A practical guide to its implementation," in *Annual Reliability and Maintainability Symposium. 2002 Proceedings (Cat. No.02CH37318)*, pp. 550–555, IEEE, 2002.

[81] S. Comert, H. Yadavari, M. Altun, and E. Erturk, "Reliability prediction of electronic boards by analyzing field return data," in *European Safety and Reliability Symposiums*, 2014.

[82] C. Kalaiselvan and L. B. Rao, "Comparison of reliability techniques of parametric and non-parametric method," *Engineering Science and Technology, an International Journal*, vol. 19, no. 2, pp. 691–699, 2016.

[83] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, no. 2, pp. 187–202, 1972.

[84] M.-Y. You and G. Meng, "Residual life prediction of repairable systems subject to imperfect preventive maintenance using extended proportional hazards model," *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, vol. 226, no. 1, pp. 50–63, 2012.

[85] N. Gorjian, L. Ma, M. Mittinty, P. Yarlagadda, and Y. Sun, "A review on reliability models with covariates," in *Engineering Asset Lifecycle Management*, pp. 385–397, Springer, 2010.

[86] I. Nafisah, M. Shrahili, N. Alotaibi, and P. Scarf, "Virtual series-system models of imperfect repair," *Reliability Engineering & System Safety*, vol. 188, pp. 604–613, 2019.

[87] H. E. Ascher, "Evaluation of Repairable System Reliability Using the "Bad-As-Old" Concept," *IEEE Transactions on Reliability*, vol. R-17, pp. 103–110, jun 1968.

[88] L. H. Crow, "Reliability analysis for complex, repairable systems," tech. rep., Army Material Systems Analysis Activity Aberdeen Proving Ground MD, 1975.

[89] L. H. Crow, "Confidence interval procedures for the weibull process with applications to reliability growth," *Technometrics*, vol. 24, no. 1, pp. 67–72, 1982.

[90] J. F. Lawless, "Regression methods for poisson process data," *Journal of the American Statistical Association*, vol. 82, no. 399, pp. 808–815, 1987.

[91] T. L. Landers, S. Jiang, and J. R. Peek, "Semi-parametric pwp model robustness for log-linear increasing rates of occurrence of failures," *Reliability Engineering & System Safety*, vol. 73, no. 2, pp. 145–153, 2001.

[92] R. L. Prentice, B. J. Williams, and A. V. Peterson, "On the regression analysis of multivariate failure time data," *Biometrika*, vol. 68, no. 2, pp. 373–379, 1981.

[93] G. Pulcini, "Modeling the failure data of a repairable equipment with bathtub type failure intensity," *Reliability Engineering & System Safety*, vol. 71, no. 2, pp. 209–218, 2001.

[94] K. J. Ryan, "Some flexible families of intensities for non-homogeneous poisson process models and their bayes inference," *Quality and Reliability Engineering International*, vol. 19, no. 2, pp. 171–181, 2003.

[95] V. V. Krivtsov, "Practical extensions to nhpp application in repairable system reliability analysis," *Reliability Engineering & System Safety*, vol. 92, no. 5, pp. 560–562, 2007.

[96] C. Guarnaccia, J. Quartieri, J. M. Barrios, and E. R. Rodrigues, "Modeling environmental noise exceedances using non-homogeneous poisson processes," *The Journal of the Acoustical Society of America*, vol. 136, no. 4, pp. 1631–1639, 2014.

[97] R. B. Karaömer and O. Chouseinoglou, "Comparison of non-homogeneous poisson process software reliability models in web applications," *AJIT-e: Bilişim Teknolojileri Online Dergisi*, vol. 7, no. 24, pp. 7–28, 2016.

[98] M. Guida, R. Calabria, and G. Pulcini, "Bayes inference for a non-homogeneous poisson process with power intensity law (reliability)," *IEEE Transactions on Reliability*, vol. 38, no. 5, pp. 603–609, 1989.

[99] S. E. Rigdon and A. P. Basu, "The power law process: a model for the reliability of repairable systems," *Journal of Quality Technology*, vol. 21, no. 4, pp. 251–260, 1989.

[100] J. O. Berger and R. L. Wolpert, "The likelihood principle," IMS, 1988.

[101] J. Fernández-Villaverde and J. F. Rubio-Ramírez, "Estimating nonlinear dynamic equilibrium economies: a likelihood approach," *Available at SSRN 486083*, 2004.

[102] N. Reid, "Likelihood inference," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 5, pp. 517–525, 2010.

[103] A. A. Abiodun, "Application Of Shared Gamma And Inverse-Gaussian Frailty Models To Cancer Data," *Global Journal of Mathematical Sciences*, vol. 7, no. 1, pp. 35–39, 2008.

[104] A. Monfort, "A reappraisal of misspecified econometric models," *Econometric Theory*, vol. 12, no. 4, pp. 597–619, 1996.

[105] H. White, *Estimation, inference and specification analysis.* No. 22, Cambridge university press, 1996.

[106] J. Fernández-Villaverde and J. F. Rubio-Ramırez, "Comparing dynamic equilibrium models to data: a bayesian approach," *Journal of Econometrics*, vol. 123, no. 1, pp. 153–187, 2004.

[107] Y.-G. Zhao and A. H. Ang, "System reliability assessment by method of moments," *Journal of Structural Engineering*, vol. 129, no. 10, pp. 1341–1349, 2003.

[108] S. C. Wang, "Principles of statistical inference: likelihood and the bayesian paradigm," *The Paleontological Society Papers*, vol. 16, pp. 1–18, 2010.

[109] G. Casella and R. L. Berger, *Statistical inference.* Cengage Learning, 2021.

[110] J. Quigley, T. Bedford, and L. Walls, "Empirical bayes estimates of development reliability for one shot devices," in *Safety and Reliability*, vol. 29, pp. 35–46, Taylor & Francis, 2009.

[111] L. Lee and S. K. Lee, "Some results on inference for the weibull process," *Technometrics*, vol. 20, no. 1, pp. 41–45, 1978.

[112] S. K. Ghosh, P. Mukhopadhyay, and J.-C. J. Lu, "Bayesian analysis of zero-inflated regression models," *Journal of Statistical planning and Inference*, vol. 136, no. 4, pp. 1360–1375, 2006.

[113] S. Schmidt and V. Wieland, "Chapter 22 - the new keynesian approach to dynamic general equilibrium modeling: Models, methods and macroeconomic policy evaluation," in *Handbook of Computable General Equilibrium Modeling SET, Vols. 1A and 1B* (P. B. Dixon and D. W. Jorgenson, eds.), vol. 1 of *Handbook of Computable General Equilibrium Modeling*, pp. 1439–1512, Elsevier, 2013.

[114] A. Gelman, "Objections to bayesian statistics," *Bayesian Analysis*, vol. 3, no. 3, pp. 445–449, 2008.

[115] G. B. Sciences and B.-m. S. Company, "individuals with multiple cancers," vol. 27, no. 12, pp. 2127–2144, 2009.

[116] M. P. Almeida, R. S. Paixão, P. L. Ramos, V. Tomazella, F. Louzada, and R. S. Ehlers, "Bayesian non-parametric frailty model for dependent competing risks in a repairable systems framework," *Reliability Engineering & System Safety*, vol. 204, p. 107145, 2020.

[117] D. D. Hanagal and A. Pandey, "Correlated inverse gaussian frailty models for bivariate survival data," *Communications in Statistics-Theory and Methods*, vol. 49, no. 4, pp. 845–863, 2020.

[118] T. Bedford, J. Quigley, and L. Walls, "Expert elicitation for reliable system design," *Statistical Science*, pp. 428–450, 2006.

[119] H. Nie, *Statistical analysis in Empirical Bayes and in causal inference.* PhD thesis, University of Pennsylvania, 2011.

[120] E. J. Green and W. E. Strawderman, "The use of bayes/empirical bayes estimation in individual tree volume equation development," *Forest Science*, vol. 31, no. 4, pp. 975–990, 1985.

[121] L. D. Brown, "In-season prediction of batting averages: A field test of empirical bayes and bayes methodologies," *The Annals of Applied Statistics*, vol. 2, no. 1, pp. 113–152, 2008.

[122] W. Vesely, S. Uryasev, and P. Samanta, "Failure of emergency diesel generators: a population analysis using empirical bayes methods," *Reliability Engineering & System Safety*, vol. 46, no. 3, pp. 221–229, 1994.

[123] J. K. Vaurio, "Uncertainties and quantification of common cause failure rates and probabilities for system analyses," *Reliability Engineering & System Safety*, vol. 90, no. 2-3, pp. 186–195, 2005.

[124] J. Quigley, G. Hardman, T. Bedford, and L. Walls, "Merging expert and empirical data for rare event frequency estimation: Pool homogenisation for empirical bayes models," *Reliability Engineering & System Safety*, vol. 96, no. 6, pp. 687–695, 2011.

[125] J. Yang, "Probability modelling of supplier development investment decisions under uncertainty," 2020.

[126] C. J. Lu and W. O. Meeker, "Using degradation measures to estimate a time-to-failure distribution," *Technometrics*, vol. 35, no. 2, pp. 161–174, 1993.

[127] S. V. Amari and G. Dill, "Redundancy optimization problem with warm-standby redundancy," in *2010 Proceedings-annual reliability and maintainability symposium (RAMS)*, pp. 1–6, IEEE, 2010.

[128] L. Hu, D. Yue, and J. Li, "Availability analysis and design optimization for a repairable series-parallel system with failure dependencies," *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 10, pp. 6693–6705, 2012.

[129] C. L. Hwang, F. A. Tillman, and M. H. Lee, "System-Reliability Evaluation Techniques for Complex/Large Systems: A Review," *IEEE Transactions on Reliability*, vol. 30, no. 5, pp. 416–423, 1981.

[130] G. Levitin, "A universal generating function approach for the analysis of multi-state systems with dependent elements," *Reliability Engineering and System Safety*, 2004.

[131] Y. Gu and J. Li, "Multi-state system reliability: A new and systematic review," in *Procedia Engineering*, 2012.

[132] M. Takeshi, "A monte carlo simulation method for system reliability analysis," *Nucl. Saf. Simul*, vol. 4, no. 1, pp. 44–52, 2013.

[133] E. Zio, M. Marella, and L. Podofillini, "A monte carlo simulation approach to the availability assessment of multi-state systems with operational dependencies," *Reliability Engineering & System Safety*, vol. 92, no. 7, pp. 871–882, 2007.

[134] E. Zio, L. Podofillini, and G. Levitin, "Estimation of the importance measures of multi-state elements by monte carlo simulation," *Reliability Engineering & System Safety*, vol. 86, no. 3, pp. 191–204, 2004.

[135] Y.-F. Li and E. Zio, "A multi-state model for the reliability assessment of a distributed generation system via universal generating function," *Reliability Engineering & System Safety*, vol. 106, pp. 28–36, 2012.

[136] P. Gao, L. Xie, W. Hu, C. Liu, and J. Feng, "Dynamic fuzzy reliability analysis of multistate systems based on universal generating function," *Mathematical Problems in Engineering*, vol. 2018, 2018.

[137] P. Wang, L. Goel, Y. Ding, L. P. Chang, and M. Andrew, "Reliability-based long term hydro/thermal reserve allocation of power systems with high wind power penetration," in *2009 IEEE Power & Energy Society General Meeting*, pp. 1–7, IEEE, 2009.

[138] Y. Ding, P. Wang, L. Goel, P. C. Loh, and Q. Wu, "Long-term reserve expansion of power systems with high wind power penetration using universal generating function methods," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 766–774, 2010.

[139] G. Levitin, L. Xing, H. Ben-Haim, and Y. Dai, "Multi-state systems with selective propagated failures and imperfect individual and group protections," *Reliability Engineering & System Safety*, vol. 96, no. 12, pp. 1657–1666, 2011.

[140] S. Ghorbani, *Reliability analysis for systems subject to degradation and shocks.* PhD thesis, Rutgers University-Graduate School-New Brunswick, 2014.

[141] E. Zio, "Some Challenges and Opportunities in Reliability Engineering," *IEEE Transactions on Reliability*, vol. 65, no. 4, pp. 1769–1782, 2016.

[142] T. M. Welte, J. Vatn, and J. Heggset, "Markov state model for optimization of maintenance and renewal of hydro power components," in *2006 International Conference on Probabilistic Methods Applied to Power Systems*, pp. 1–7, IEEE, 2006.

[143] Y.-H. Lin, Y.-F. Li, and E. Zio, "Reliability assessment of systems subject to dependent degradation processes and random shocks," *IIE transactions*, vol. 48, no. 11, pp. 1072–1085, 2016.

[144] J. P. Kharoufeh and S. M. Cox, "Stochastic models for degradation-based reliability," *IIE Transactions*, vol. 37, no. 6, pp. 533–542, 2005.

[145] S. Ambani, L. Li, and J. Ni, "Condition-based maintenance decision-making for multiple machine systems," *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, vol. 131, no. 3, pp. 031009:1–031009:9, 2009.

[146] A. Birolini, *Reliability engineering: theory and practice*. Springer Science & Business Media, 2013.

[147] B. Natvig and A. Streller, "The steady-state behaviour of multistate monotone systems," *Journal of applied probability*, vol. 21, no. 4, pp. 826–835, 1984.

[148] J. Xue and K. Yang, "Dynamic reliability analysis of coherent multistate systems," *IEEE transactions on Reliability*, vol. 44, no. 4, pp. 683–688, 1995.

[149] E. L. Abner, R. J. Charnigo, and R. J. Kryscio, "Markov chains and semi-markov models in time-to-event analysis," *Journal of biometrics & biostatistics*, no. e001, p. 19522, 2013.

[150] J. Shao and L. R. Lamberson, "Modeling a shared-load k-out-of-n: G system," *IEEE Transactions on Reliability*, vol. 40, no. 2, pp. 205–209, 1991.

[151] M. Sharifi, A. E. Memariani, and A. R. Nour, "Real time study of a k-out-of-n system: n identical elements with increasing failure rates," *Iranian Journal of Operations Research*, 2009.

[152] M. A. U. H. Broek, R. H. Teunter, B. de Jonge, and J. Veldman, "Joint condition-based maintenance and load-sharing optimization for two-unit systems with economic dependency," *European Journal of Operational Research*, 2021.

[153] S. Bloch-Mercier, "A preventive maintenance policy with sequential checking procedure for a markov deteriorating system," *European Journal of Operational Research*, vol. 142, no. 3, pp. 548–576, 2002.

[154] A. J. Endharta and Y. M. Ko, "Economic design and maintenance of a circular k-out-of-n: G balanced system with load-sharing units," *IEEE Transactions on Reliability*, vol. 69, no. 4, pp. 1465–1479, 2020.

[155] B. H. Lindqvist, "Statistical modeling and analysis of repairable systems," in *Statistical and probabilistic models in reliability*, pp. 3–25, Springer, 1999.

[156] A. Wieke, "Frailty models in survival analysis [habilitation]," *Halle-Wittenberg: Medizinische Fakultät der Martin-Luther-Universität*, 2007.

[157] R. Jiang, M. Zuo, and H.-X. Li, "Weibull and inverse weibull mixture models allowing negative weights," *Reliability Engineering & System Safety*, vol. 66, no. 3, pp. 227–234, 1999.

[158] A. W. Marshall and I. Olkin, *Life distributions*, vol. 13. Springer, 2007.

[159] C. A. Vallejos, *Incorporating unobserved heterogeneity and multiple event types in survival models: A Bayesian approach.* PhD thesis, University of Warwick, 2014.

[160] J. P. Sy and J. M. Taylor, "Estimation in a cox proportional hazards cure model," *Biometrics*, vol. 56, no. 1, pp. 227–236, 2000.

[161] D. L. Price and A. K. Manatunga, "Modelling survival data with a cured fraction using frailty models," *Statistics in medicine*, vol. 20, no. 9-10, pp. 1515–1527, 2001.

[162] L. Wu, Y. Zou, and D. Lord, "Comparison of sichel and negative binomial models in hot spot identification," *Transportation Research Record*, vol. 2460, no. 1, pp. 107–116, 2014.

[163] J. W. Vaupel, K. G. Manton, and E. Stallard, "The impact of heterogeneity in individual frailty on the dynamics of mortality," *Demography*, vol. 16, no. 3, pp. 439–454, 1979.

[164] A. Wienke, *Frailty models in survival analysis.* CRC Press, 2011.

[165] D. D. Hanagal, *Modeling survival data using frailty models.* Springer, 2011.

[166] L. Zha, D. Lord, and Y. Zou, "The poisson inverse gaussian (pig) generalized linear regression model for analyzing motor vehicle crash data," *Journal of Transportation Safety & Security*, vol. 8, no. 1, pp. 18–35, 2016.

[167] Y. Zou, D. Lord, Y. Zhang, and Y. Peng, "Comparison of sichel and negative binomial models in estimating empirical bayes estimates," *Transportation research record*, vol. 2392, no. 1, pp. 11–21, 2013.

[168] E. Hauer, D. W. Harwood, F. M. Council, and M. S. Griffith, "Estimating safety by the empirical bayes method: a tutorial," *Transportation Research Record*, vol. 1784, no. 1, pp. 126–131, 2002.

[169] W. S. Gachau, "Frailty models with applications in medical research : observed and simulated data," 2014.

[170] C. E. McCulloch and J. M. Neuhaus, "Generalized linear mixed models," in *International Encyclopedia of the Social Behavioral Sciences (Second Edition)* (J. D. Wright, ed.), pp. 845–852, Oxford: Elsevier, second edition ed., 2015.

[171] T. A. Balan and H. Putter, "A tutorial on frailty models," *Statistical methods in medical research*, vol. 29, no. 11, pp. 3424–3454, 2020.

[172] S. Kheiri, A. Kimber, and M. R. Meshkani, "Bayesian analysis of an inverse gaussian correlated frailty model," *Computational statistics & data analysis*, vol. 51, no. 11, pp. 5317–5326, 2007.

[173] D. D. Hanagal and A. D. Dabade, "Bayesian estimation of parameters and comparison of shared gamma frailty models," *Communications in Statistics-Simulation and Computation*, vol. 42, no. 4, pp. 910–931, 2013.

[174] P. Hougaard, "Frailty models for survival data," *Lifetime Data Analysis*, vol. 1, no. 3, pp. 255–273, 1995.

[175] L. Duchateau and P. Janssen, *The frailty model.* Springer Science & Business Media, 2007.

[176] A. Wienke, K. G. Arbeev, I. Locatelli, and A. I. Yashin, "A comparison of different bivariate correlated frailty models and estimation strategies," *Mathematical Biosciences*, vol. 198, pp. 1–13, nov 2005.

[177] D. Darmofal, "Bayesian spatial survival models for political event processes," *American Journal of Political Science*, vol. 53, no. 1, pp. 241–257, 2009.

[178] M. R. Ghadimi, M. Mahmoodi, K. Mohammad, M. Rasouli, H. Zeraati, and A. Fotouhi, "Factors affecting survival of patients with oesophageal cancer: a study using inverse gaussian frailty models," *Singapore medical journal*, vol. 53, no. 5, p. 336, 2012.

[179] J. Cui, A. Forbes, A. Kirby, I. Marschner, J. Simes, M. West, and A. Tonkin, "Parametric conditional frailty models for recurrent cardiovascular events in the lipid study," *Clinical Trials*, vol. 5, no. 6, pp. 565–574, 2008.

[180] J. M. Box-Steffensmeier and S. De Boef, "Repeated events survival models: the conditional frailty model," *Statistics in medicine*, vol. 25, no. 20, pp. 3518–3533, 2006.

[181] P. Hougaard, "Summary," pp. 483–496, 2000.

[182] A. Wienke, N. V. Holm, A. Skytthe, and A. I. Yashin, "The heritability of mortality due to heart diseases: a correlated frailty model applied to danish twins," *Twin Research and Human Genetics*, vol. 4, no. 4, pp. 266–274, 2001.

[183] L. J. Bain and F. Wright, "The negative binomial process with applications to reliability," *Journal of Quality Technology*, vol. 14, no. 2, pp. 60–66, 1982.

[184] M. Engelhardt and L. J. Bain, "Statistical analysis of a compound power-law model for repairable systems," *IEEE transactions on reliability*, vol. 36, no. 4, pp. 392–396, 1987.

[185] B. H. Lindqvist, G. Elvebakk, and K. Heggland, "The trend-renewal process for statistical analysis of repairable systems," *Technometrics*, vol. 45, no. 1, pp. 31–44, 2003.

[186] X. Liu, J. Vatn, Y. Dijoux, and H. Toftaker, "Unobserved heterogeneity in stable imperfect repair models," *Reliability Engineering and System Safety*, vol. 203, p. 107039, 2020.

[187] Y. Li and L. Ryan, "Modeling spatial survival data using semiparametric frailty models," *Biometrics*, vol. 58, no. 2, pp. 287–297, 2002.

[188] B. P. Carlin, S. Banerjee, *et al.*, "Hierarchical multivariate car models for spatio-temporally correlated survival data," *Bayesian statistics*, vol. 7, no. 7, pp. 45–63, 2003.

[189] L. Zhao and T. E. Hanson, "Spatially dependent polya tree modeling for survival data," *Biometrics*, vol. 67, no. 2, pp. 391–403, 2011.

[190] P. Ojiambo and E. Kang, "Modeling spatial frailties in survival analysis of cucurbit downy mildew epidemics," *Phytopathology*, vol. 103, no. 3, pp. 216–227, 2013.

[191] S. Banerjee, B. P. Carlin, and A. E. Gelfand, *Hierarchical modeling and analysis for spatial data.* CRC press, 2014.

[192] S. A. Thamrin, Amran, A. K. Jaya, S. Rahmi, and Ansariadi, "Bayesian inference for spatial parametric proportional hazards model using spatsurv r," in *AIP Conference Proceedings*, vol. 1827, p. 020015, AIP Publishing LLC, 2017.

[193] K. J., O. Ngesa, and G. Orwa, "Misspecification of frailty random effects in a clustered survival data," *Journal of Statistical and Econometric Methods*, vol. 8, no. 2, pp. 1–25, 2019.

[194] D. G. Clayton, "A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence," *Biometrika*, vol. 65, no. 1, pp. 141–151, 1978.

[195] P. Hougaard, "A class of multivanate failure time distributions," *Biometrika*, vol. 73, no. 3, pp. 671–678, 1986.

[196] D. Oakes, "Bivariate survival models induced by frailties," *Journal of the American Statistical Association*, vol. 84, no. 406, pp. 487–493, 1989.

[197] U. Abdulkarimova, *Frailty Models for Modelling Heterogeneity.* PhD thesis, 2013.

[198] T. Lancaster, "Econometric methods for the duration of unemployment," *Econometrica: Journal of the Econometric Society*, pp. 939–956, 1979.

[199] O. O. Aalen, "Two examples of modelling heterogeneity in survival analysis," *Scandinavian journal of statistics*, pp. 19–25, 1987.

[200] R. Ellermann, P. Sullo, and J. M. Tien, "An alternative approach to modeling recidivism using quantile residual life functions," *Operations research*, vol. 40, no. 3, pp. 485–504, 1992.

[201] P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding, *Statistical models based on counting processes.* Springer Science & Business Media, 2012.

[202] P. Hougaard, "Life table methods for heterogeneous populations: distributions describing the heterogeneity," *Biometrika*, vol. 71, no. 1, pp. 75–83, 1984.

[203] M. Gacula Jr and J. Kubala, "Statistical models for shelf life failures," *Journal of Food Science*, vol. 40, no. 2, pp. 404–409, 1975.

[204] H. Vilmann, S. Kirkeby, and D. Kronborg, “Histomorphometrical analysis of the influence of soft diet on masticatory muscle development in the muscular dystrophic mouse,” *Archives of oral Biology*, vol. 35, no. 1, pp. 37–42, 1990.

[205] O. E. Barndorff-Nielsen, “A note on electrical networks and the inverse gaussian distribution,” *Advances in Applied Probability*, vol. 26, no. 1, pp. 63–67, 1994.

[206] V. Seshadri, “Histomorphometry,” in *The Inverse Gaussian Distribution*, pp. 194–197, Springer, 1999.

[207] R. Chhikara and J. Folks, “The inverse gaussian distribution as a lifetime model,” *Technometrics*, vol. 19, no. 4, pp. 461–468, 1977.

[208] J. H. Cha and M. Finkelstein, “Stochastic intensity for minimal repairs in heterogeneous populations,” *Journal of applied probability*, vol. 48, no. 3, pp. 868–876, 2011.

[209] J. H. Cha and M. Finkelstein, “Stochastic analysis of preventive maintenance in heterogeneous populations,” *Operations research letters*, vol. 40, no. 5, pp. 416–421, 2012.

[210] J. H. Cha, “Optimal replacement of heterogeneous items with minimal repairs,” *IEEE Transactions on Reliability*, vol. 65, no. 2, pp. 593–603, 2016.

[211] R. Dekker, R. E. Wildeman, and F. A. Van der Duyn Schouten, “A review of multi-component maintenance models with economic dependence,” *Mathematical methods of operations research*, vol. 45, no. 3, pp. 411–435, 1997.

[212] F. Camci, “System maintenance scheduling with prognostics information using genetic algorithm,” *IEEE Transactions on reliability*, vol. 58, no. 3, pp. 539–552, 2009.

[213] T. Nowakowski and S. Werbińka, “On problems of multicomponent system maintenance modelling,” *International Journal of Automation and Computing*, vol. 6, no. 4, p. 364, 2009.

[214] R. P. Nicolai and R. Dekker, “Optimal maintenance of multi-component systems: a review,” *Complex system maintenance handbook*, pp. 263–286, 2008.

[215] D. Murthy and D. Nguyen, “Study of two-component system with failure interaction,” *Naval Research Logistics Quarterly*, vol. 32, no. 2, pp. 239–247, 1985.

[216] D. Murthy and D. Nguyen, “Study of a multi-component system with failure interaction,” *European Journal of Operational Research*, vol. 21, no. 3, pp. 330–338, 1985.

[217] R. P. Nicolai and R. Dekker, “Optimal maintenance of multi-component systems: a review,” *Complex system maintenance handbook*, pp. 263–286, 2008.

[218] L. Bian and N. Gebraeel, “Stochastic modeling and real-time prognostics for multi-component systems with degradation rate interactions,” *Iie Transactions*, vol. 46, no. 5, pp. 470–482, 2014.

[219] Y. Sun, L. Ma, J. Mathew, and S. Zhang, "An analytical model for interactive failures," *Reliability Engineering & System Safety*, vol. 91, no. 5, pp. 495–504, 2006.

[220] C.-K. Sung, S.-H. Sheu, T.-S. Hsu, and Y.-C. Chen, "Extended optimal replacement policy for a two-unit system with failure rate interaction and external shocks," *International Journal of Systems Science*, vol. 44, no. 5, pp. 877–888, 2013.

[221] T. Nakagawa and D. Murthy, "Optimal replacement policies for a two-unit system with failure interactions," *RAIRO-Operations Research*, vol. 27, no. 4, pp. 427–438, 1993.

[222] T. Satow and S. Osaki, "Optimal replacement policies for a two-unit system with shock damage interaction," *Computers & Mathematics with Applications*, vol. 46, no. 7, pp. 1129–1138, 2003.

[223] M.-T. Lai and Y.-C. Chen, "Optimal periodic replacement policy for a two-unit system with failure rate interaction," *The international journal of advanced manufacturing technology*, vol. 29, no. 3-4, pp. 367–371, 2006.

[224] M.-T. Lai, "A discrete replacement model for a two-unit parallel system subject to failure rate interaction," *Quality and Quantity*, vol. 43, no. 3, pp. 471–479, 2009.

[225] J.-P. Jhang and S.-H. Sheu, "Optimal age and block replacement policies for a multi-component system with failure interaction," *International Journal of Systems Science*, vol. 31, no. 5, pp. 593–603, 2000.

[226] M.-T. Lai, "Periodical replacement model for a multi-unit system subject to failure rate interaction," *Quality & quantity*, vol. 41, no. 3, pp. 401–411, 2007.

[227] Z. Li, Q. Gao, S. Liu, and G. Liu, "Reliability analysis for a multi-unit system with failure rate interaction," in *2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE)*, pp. 395–398, IEEE, 2013.

[228] T. J.-M. Meango and M.-S. Ouali, "Failure interaction models for multicomponent systems: a comparative study," *SN Applied Sciences*, vol. 1, no. 1, pp. 1–25, 2019.

[229] Y. Sun and L. Ma, "Estimate of interactive coefficients," in *2009 8th International Conference on Reliability, Maintainability and Safety*, pp. 310–314, IEEE, 2009.

[230] R. I. Zequeira and C. Bérenguer, "Periodic imperfect preventive maintenance with two categories of competing failure modes," *Reliability Engineering & System Safety*, vol. 91, no. 4, pp. 460–468, 2006.

[231] I. Castro, "A model of imperfect preventive maintenance with dependent failure modes," *European Journal of Operational Research*, vol. 196, no. 1, pp. 217–224, 2009.

[232] H. Fan, C. Hu, M. Chen, and D. Zhou, "Cooperative predictive maintenance of repairable systems with dependent failure modes and resource constraint," *IEEE Transactions on Reliability*, vol. 60, no. 1, pp. 144–157, 2011.

[233] G. Levitin and L. Xing, “Reliability and performance of multi-state systems with propagated failures having selective effect,” *Reliability Engineering & System Safety*, vol. 95, no. 6, pp. 655–661, 2010.

[234] M.-T. Lai and Y.-C. Chen, “Optimal replacement period of a two-unit system with failure rate interaction and external shocks,” *International Journal of Systems Science*, vol. 39, no. 1, pp. 71–79, 2008.

[235] H. R. Golmakani and H. Moakedi, “Periodic inspection optimization model for a multi-component repairable system with failure interaction,” *The International Journal of Advanced Manufacturing Technology*, vol. 61, no. 1, pp. 295–302, 2012.

[236] T. Zhao, X. Qi, and J. Liu, “Interactive failure analysis of gyroscope based on net failure mode and effect analysis,” in *2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE)*, pp. 246–249, IEEE, 2013.

[237] J. Wang and M. Li, “Redundancy allocation for reliability design of engineering systems with failure interactions,” in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 46322, p. V02BT03A042, American Society of Mechanical Engineers, 2014.

[238] C. Hou, C. Chen, J. Wang, and K. Shi, “Reliability analysis for software cluster systems based on proportional hazard model,” in *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, vol. 1, pp. 32–41, IEEE, 2016.

[239] B. Liu, M. Xie, and W. Kuo, “Reliability modeling and preventive maintenance of load-sharing systemswith degrading components,” *IIE Transactions*, vol. 48, no. 8, pp. 699–709, 2016.

[240] H. E. Daniels, “The statistical theory of the strength of bundles of threads. i,” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 183, no. 995, pp. 405–435, 1945.

[241] P. K. Porwal, I. J. Beyerlein, and S. L. Phoenix, “Statistical strength of a twisted fiber bundle: an extension of daniels equal-load-sharing parallel bundle theory,” *Journal of Mechanics of Materials and Structures*, vol. 1, no. 8, pp. 1425–1447, 2006.

[242] G. Levitin and S. V. Amari, “Optimal load distribution in series–parallel systems,” *Reliability Engineering & System Safety*, vol. 94, no. 2, pp. 254–260, 2009.

[243] E. M. Scheuer, “Reliability of an m-out of-n system when component failure induces higher failure rates in survivors,” *IEEE Transactions on reliability*, vol. 37, no. 1, pp. 73–74, 1988.

[244] R. C. Gupta, “Reliability of ak out of n system of components sharing a common environment,” *Applied mathematics letters*, vol. 15, no. 7, pp. 837–844, 2002.

[245] S. Tang and F. Wang, “Reliability analysis for a repairable parallel system with time-varying failure rates,” *Applied Mathematics-A Journal of Chinese Universities*, vol. 20, no. 1, pp. 85–90, 2005.

[246] T. Alfok, C. V. Duy, and M. Baykal-Gursoy, "Two load sharing processors with failures," *Computers & operations research*, vol. 25, no. 3, pp. 183–189, 1998.

[247] H. Liu, "Reliability of a load-sharing k-out-of-n: G system: non-iid components with arbitrary distributions," *IEEE Transactions on Reliability*, vol. 47, no. 3, pp. 279–284, 1998.

[248] T. Yinghui and Z. Jing, "New model for load-sharing k-out-of-n: G system with different components," *Journal of Systems Engineering and Electronics*, vol. 19, no. 4, pp. 748–842, 2008.

[249] L. Huang and Q. Xu, "Lifetime reliability for load-sharing redundant systems with arbitrary failure distributions," *IEEE Transactions on Reliability*, vol. 59, no. 2, pp. 319–330, 2010.

[250] M. Jain and R. Gupta, "Load sharing M-out of-N: G system with non-identical components subject to common cause failure," *International Journal of Mathematics in Operational Research*, vol. 4, no. 5, pp. 586–605, 2012.

[251] I. Maatouk, E. Châtelet, and N. Chebbo, "Reliability of multi-states system with load sharing and propagation failure dependence," in *2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pp. 42–46, IEEE, 2011.

[252] Z. Ye, M. Revie, and L. Walls, "A load sharing system reliability model with managed component degradation," *IEEE transactions on reliability*, vol. 63, no. 3, pp. 721–730, 2014.

[253] C. Yang, S. Zeng, and J. Guo, "Reliability analysis of load-sharing-out-of-system considering component degradation," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[254] B. Singh, K. K. Sharma, and A. Kumar, "A classical and Bayesian estimation of a k-components load-sharing parallel system," *Computational Statistics and Data Analysis*, vol. 52, no. 12, pp. 5175–5185, 2008.

[255] B. Singh and P. K. Gupta, "Load-sharing system model and its application to the real data set," *Mathematics and Computers in Simulation*, vol. 82, no. 9, pp. 1615–1629, 2012.

[256] M. Ibnabdeljalil and W. A. Curtin, "Strength and reliability of notched fiber-reinforced composites," *Acta Materialia*, vol. 45, no. 9, pp. 3641–3652, 1997.

[257] S. D. Durham, J. D. Lynch, W. J. Padgett, T. J. Horan, W. J. Owen, and J. Surles, "Localized load-sharing rules and Markov-Weibull fibers: A comparison of microcomposite failure data with Monte Carlo simulations," *Journal of Composite Materials*, vol. 31, no. 18, pp. 1856–1882, 1997.

[258] L. Wang, Y. Tian, Y. Wang, and B. Liu, "A 6-component circular multi-state Markov repairable system with spatial dependence," in *2015 First International Conference on Reliability Systems Engineering (ICRSE)*, pp. 1–5, IEEE, 2015.

[259] J. Guo, Z. Liu, H. Che, and S. Zeng, "Reliability model of consecutive $(2, k)$-out-of-$(2, n)$: F systems with local load-sharing," *IEEE Access*, vol. 6, pp. 8178–8188, 2018.

[260] G. Levitin and M. Finkelstein, "Optimal loading of elements in series systems exposed to external shocks," *Reliability Engineering & System Safety*, vol. 192, p. 105924, 2019.

[261] H. Che, S. Zeng, and J. Guo, "Reliability analysis of load-sharing systems subject to dependent degradation processes and random shocks," *IEEE Access*, vol. 5, pp. 23395–23404, 2017.

[262] R. Ahmadi, "Reliability and maintenance modeling for a load-sharing k-out-of-n system subject to hidden failures," *Computers & Industrial Engineering*, vol. 150, p. 106894, 2020.

[263] M. Arabzadeh Jamali and H. Pham, "Opportunistic maintenance model for load sharing k-out-of-n systems with perfect pm and minimal repairs," *Quality Engineering*, pp. 1–10, 2022.

[264] C. Shekhar, A. Kumar, and S. Varshney, "Load sharing redundant repairable systems with switching and reboot delay," *Reliability Engineering & System Safety*, vol. 193, p. 106656, 2020.

[265] J. E. Freund, "A bivariate extension of the exponential distribution," *Journal of the American Statistical Association*, vol. 56, no. 296, pp. 971–977, 1961.

[266] P. Pozsgai, W. Neher, and B. Bertsche, "Models to consider load-sharing in reliability calculation and simulation of systems consisting of mechanical components," in *Annual Reliability and Maintainability Symposium, 2003.*, pp. 493–499, IEEE, 2003.

[267] H.-H. Lin, K.-H. Chen, and R.-T. Wang, "A multivariant exponential shared-load model," *IEEE Transactions on Reliability*, vol. 42, no. 1, pp. 165–171, 1993.

[268] R. Mohammad, A. Kalam, and S. V. Amari, "Reliability evaluation of phased-mission systems with load-sharing components," in *2012 Proceedings Annual Reliability and Maintainability Symposium*, pp. 1–6, IEEE, 2012.

[269] R. Mohammad, A. Kalam, and S. V. Amari, "Reliability of load-sharing systems subject to proportional hazards model," in *2013 Proceedings Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1–5, IEEE, 2013.

[270] Y. Kleiner and B. Rajani, "Comprehensive review of structural deterioration of water mains: statistical models," *Urban water*, vol. 3, no. 3, pp. 131–150, 2001.

[271] A. Scheidegger, J. P. Leitao, and L. Scholten, "Statistical failure models for water distribution pipes–a review from a unified perspective," *Water research*, vol. 83, pp. 237–247, 2015.

[272] R. S. Chhikara and L. Folks, *The inverse Gaussian distribution : theory, methodology, and applications.* M. Dekker, 1989.

[273] W. O. Onchere, *Frailty Models Applications In Pension Schemes.* PhD thesis, University of Nairobi, 2013.

[274] L. H. M. Morita, V. L. Tomazella, and F. Louzada-Neto, “Accelerated lifetime modelling with frailty in a non-homogeneous poisson process for analysis of recurrent events data,” *Quality Technology & Quantitative Management*, vol. 15, no. 2, pp. 209–229, 2018.

[275] L. S. Piancastelli, W. Barreto-Souza, and V. D. Mayrink, “Generalized inverse-gaussian frailty models with application to target neuroblastoma data,” *Annals of the Institute of Statistical Mathematics*, vol. 73, no. 5, pp. 979–1010, 2021.

[276] D. D. Hanagal and R. Sharma, “Analysis of bivariate survival data using shared inverse Gaussian frailty model,” *Communications in Statistics-Theory and Methods*, vol. 44, no. 7, pp. 1351–1380, 2015.

[277] L. H. Morita, V. L. Tomazella, N. Balakrishnan, P. L. Ramos, P. H. Ferreira, and F. Louzada, “Inverse gaussian process model with frailty term in reliability analysis,” *Quality and Reliability Engineering International*, vol. 37, no. 2, pp. 763–784, 2021.

[278] J. Boxall, A. O’Hagan, S. Pooladsaz, A. Saul, and D. Unwin, “Estimation of burst rates in water distribution mains,” in *Proceedings of the Institution of Civil Engineers-Water Management*, vol. 160, pp. 73–82, Thomas Telford Ltd, 2007.

[279] E. Levlin, “Water and waste pipes,” *Water Resources Engineering-Royal Institute of Technology*, 2004.

[280] V. Tzatchkov, M. Hansen, and H. Ramírez, “A practical approach for prioritizing the replacement of water pipes in mexico city,” *MRS Online Proceedings Library (OPL)*, vol. 1278, 2010.

[281] H. Xiao, D. Shi, Y. Ding, and R. Peng, “Optimal loading and protection of multi-state systems considering performance sharing mechanism,” *Reliability Engineering & System Safety*, vol. 149, pp. 88–95, 2016.

[282] D. F. Percy and B. M. Alkali, “Generalized proportional intensities models for repairable systems,” *IMA Journal of Management Mathematics*, vol. 17, no. 2, pp. 171–185, 2006.

[283] B. Tessema, S. Ayalew, and K. Mohammed, “Modeling the determinants of time-to-age at first marriage in ethiopian women: A comparison of various parametric shared frailty models,” *Sci J Public Health*, vol. 3, no. 5, pp. 707–718, 2015.

[284] N. Ata Tutkun and P. Marthin, “A comparative study with bootstrap resampling technique to uncover behavior of unconditional hazards and survival functions for gamma and inverse gaussian frailty models,” *Mathematical Sciences*, vol. 15, no. 1, pp. 99–109, 2021.

[285] M. Shoukri, M. Asyali, R. VanDorp, and D. Kelton, “The poisson inverse gaussian regression model in the analysis of clustered counts data,” *Journal of Data Science*, vol. 2, no. 1, pp. 17–32, 2004.

[286] M. Zhang and M. Revie, “Model selection with application to gamma process and inverse gaussian process,” *CRC/Taylor & Francis Group*, 2016.

[287] M. J. Brewer, A. Butler, and S. L. Cooksley, "The relative performance of aic, aicc and bic in the presence of unobserved heterogeneity," *Methods in Ecology and Evolution*, vol. 7, no. 6, pp. 679–692, 2016.

[288] S. Izumi and M. Ohtaki, "Aspects of the armitage–doll gamma frailty model for cancer incidence data," *Environmetrics: The official journal of the International Environmetrics Society*, vol. 15, no. 3, pp. 209–218, 2004.

[289] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4. Springer, 2006.

[290] A. Banbeta, D. Seyoum, T. Belachew, B. Birlie, and Y. Getachew, "Modeling time-to-cure from severe acute malnutrition: application of various parametric frailty models," *Archives of Public Health*, vol. 73, no. 1, pp. 1–8, 2015.

[291] S. A. Adham and A. A. AlAhmadi, "Gamma and inverse gaussian frailty models: A comparative study," *Journal of Mathematics and Statistics Invention (IJMSI) EISSN*, pp. 2321–4767, 2016.

[292] P. Loungani, "How accurate are private sector forecasts? cross-country evidence from consensus forecasts of output growth," *International journal of forecasting*, vol. 17, no. 3, pp. 419–432, 2001.

[293] B. H. Lindqvist, "On the statistical modeling and analysis of repairable systems," *Statistical science*, vol. 21, no. 4, pp. 532–551, 2006.

[294] D. D. Boos, "Introduction to the bootstrap world," *Statistical science*, vol. 18, no. 2, pp. 168–174, 2003.

[295] B. P. Carlin and T. A. Louis, "Empirical bayes: Past, present and future," *Journal of the American Statistical Association*, vol. 95, no. 452, pp. 1286–1289, 2000.

[296] N. Gebraeel, A. Elwany, and J. Pan, "Residual life predictions in the absence of prior degradation knowledge," *IEEE Transactions on Reliability*, vol. 58, no. 1, pp. 106–117, 2009.

[297] M. Munda, F. Rotolo, and C. Legrand, "parfm: Parametric frailty models in r," *Journal of Statistical Software*, vol. 51, pp. 1–20, 2012.

[298] Y. Chen, "On the four types of weight functions for spatial contiguity matrix," *Letters in Spatial and Resource Sciences*, vol. 5, no. 2, pp. 65–72, 2012.

[299] A. D. Polyanin, *Handbook of linear partial differential equations for engineers and scientists*. Chapman and hall/crc, 2001.

[300] B. Biswas, S. Chatterjee, S. Mukherjee, and S. Pal, "A discussion on euler method: A review," *Electronic Journal of Mathematical Analysis and Applications*, vol. 1, no. 2, pp. 2090–2792, 2013.

[301] B. N. Sur and T. Sarkar, "Numerical method of reliability evaluation of a stand-by redundant system," *Microelectronics Reliability*, vol. 36, no. 5, pp. 693–696, 1996.

[302] M. K. Sharma and K. K. Bansal, "Numerical Solution for Reliability and Availability Analysis of a Three Unit Standby System," *International Journal of Engineering and Technology (IJET) Vol*, vol. 3, no. 3, pp. 1804–1808, 2017.

[303] W. Wang, F. Zhao, R. Peng, and L. Guo, "A preventive maintenance model for linear consecutive k-out-of-n: F systems with dependent components," in *Engineering Asset Management-Systems, Professional Practices and Certification*, pp. 439–449, Springer, 2015.

[304] H. Liu, "Reliability of a load-sharing k-out-of-n: G system: non-iid components with arbitrary distributions," *IEEE Transactions on Reliability*, vol. 47, no. 3, pp. 279–284, 1998.

[305] Y. Li, F. P. Coolen, C. Zhu, and J. Tan, "Reliability assessment of the hydraulic system of wind turbines based on load-sharing using survival signature," *Renewable Energy*, vol. 153, pp. 766–776, 2020.

[306] S. M. Ross, J. J. Kelly, R. J. Sullivan, W. J. Perry, D. Mercer, R. M. Davis, T. D. Washburn, E. V. Sager, J. B. Boyce, and V. L. Bristow, *Stochastic processes*, vol. 2. Wiley New York, 1996.