Application of reliability analysis for performance assessments in railway infrastructure asset management

by

Nigel Tatenda, Zhuwaki

Thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering in Engineering Management in the Faculty of Engineering at Stellenbosch University

Supervisor: Prof CJ Neels Fourie

Co-Supervisor: Mr Joubert Van Eeden

March 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: Nigel Tatenda, Zhuwaki

Date: March 2017

Abstract

Reliable railway infrastructure systems guarantee the safety of operations and the availability of train services. With an increase in mobility demands, it is increasingly becoming a challenge to deliver railway infrastructure systems with a sustainable functionality that meets the various dependability attributes such as reliability, availability, and maintainability. Decisions related to infrastructure asset management in the railway industry focus on the maintenance, enhancement, and renewal of assets. This is to ensure that the infrastructure assets meet the required level of dependability and quality of service at the lowest life cycle costs. The success of these decisions depends on the effective management of individual assets over their lifetime from the perspective of a whole systems approach. A whole systems approach offers greater advantages over the traditional silo approach which lacks integration and coordination in the maintenance and management of complex cross-functional multi-asset systems. Reliability, when applied to infrastructure asset management, is a mathematical concept associated with dependability in which engineering knowledge is applied to identify and reduce the likelihood or frequency of failures within a system. In addition, it enables a systematic analysis to be performed at various levels of the railway network to quantify the various dependability attributes of individual infrastructure assets and their impact on the overall performance of the infrastructure system.

The objective of this study is to develop a scientific approach to model and evaluate the reliability performance of railway infrastructure systems. This paper presents the development and application of a holistic reliability model for multi-asset systems that can facilitate and improve infrastructure maintenance management processes in railway environments. The model is applied and validated using a practical case study in the context of the Passenger Rail Agency of South Africa (PRASA). The case study applied to PRASA`s Metrorail network concluded that a holistic performance assessment method using reliability analysis can assist in improving the maintenance and management of railway infrastructure assets to guarantee high quality of service.

Keywords: System reliability analysis, Asset management, Railway infrastructure maintenance.

Opsomming

Spoorweg infrastruktuurstelsels waarborg die veiligheid van werksaamhede/bedrywighede asook die beskikbaarheid van treindienste. Met 'n toename in mobiliteitsvereistes raak dit 'n al groter problem/uitdaging om spoorweg infrastruktuur met 'n volhoubaarhieds-funksionaliteit te lewer wat die verskeie afhanklikheidskenmerke, soos betroubaarheid, beskikbaarheid en onderhoudbaarheid. Besluite rakende infrastruktuur batebestuur in die spoorweg-industrie fokus op instandhouding, versterking en vernuwing van bates. Dit is om te verseker dat die infrastruktuur se bates die vereiste vlak van betroubaarheid en kwaliteitsdiens by die laagste moontlike lewensikluskostes handhaaf. Die sukses van hierdie besluite hang af van die effektiewe bestuur van individuele bates tydens hulle leeftyd van die perspektief van die volledige stelsel-aanslag. 'n Volledige stelsel-aanslag bied groter voordele in vergelyking met die tradisionele silo-aanslag waar integriteit en koördinasie ontbreek in die onderhoud en bestuur van komplekse kruis-funksionele multi-bate stelsels. Daarby is dit moontlik om 'n sistemiese analise uit te voer by verskillende vlakke van die spoornetwerk om die verskillende betroubaarheidseienskappe van die individuele infrastruktuur bates en hulle impak op die algehele werksverrigting van die infrastruktuurstelsel te kwantifiseer. Waar dit infrastruktuur batebestuur aangaan, is betroubaarheid 'n wiskundige konsep wat geassosieer word met betroubaarheid in die ingenieurskennis wat toegepas word om die waarksynlikheid en frekwensie van falings binne die stelsel te identifiseer en te verminder. Die doel van hierdie tesis is om 'n wetenskaplike benadering te ontwikkel om die betroubaarheidsnakoming van die spoorweg-infrastruktuurstelsels te modelleer en te evalueer. Hierdie tesis stel die ontwikkeling en toepassing van 'n holistiese betroubaarheidsmodel voor vir 'n multi-bate stelsel wat die infrastruktuur instandhoudingsbestuurprosesse in spoorweg-omgewings kan fasiliteer en verbeter. Die model word toegepas en geldig verklaar deur gebruik te maak van 'n praktiese gevallestudie in die konteks van Passasier Spoor Agentskap van Suid-Afrika (Passenger Rail Agency of South Africa (PRASA)). Die gevallestudie wat toegepas is op PRASA se Metrorail netwerk het tot die gevolgtrekking gekom dat 'n holistiese werksverrigting assesseringsmetode nodig is wat betroubaarheidsanalises gebruik wat kan bydra tot die verbetering van die instandhouding en bestuur van spoorweg-infrastruktuurbates om hoë kwaliteit diens te verseker.

Acknowledgements

Firstly I would like to express my sincere gratitude to my thesis advisor Professor C.J Fourie who has supported me throughout my thesis with his patience, knowledge, and invaluable guidance. He consistently allowed this paper to be my own work by providing me with the room to work in my own way and guiding me in the right the direction whenever he thought I needed it. Without his efforts, this thesis would not have been successful. I would also like to thank my co-supervisor Joubert Van Eeden for his invaluable input in my results and constructive suggestions which contributed immensely to the quality of the work.

I would like to thank the staff at PRASA Western Cape Depot for the support and timeous assistance in providing the necessary information and feedback that has contributed to the success of this thesis. To Robert Venter, I thank you for your support in ensuring I connected with Ayanda Bani, Jaco Cupido, John Mollet, Raymond Maseko and Jaime Mabota from the Engineering services department. Without their passionate participation and input, this thesis could not have been successfully completed.

I am particularity grateful to Pieter Conradie from the PRASA Engineering Research Chair at Stellenbosch University for his continuous encouragement and suggestions throughout the course of my thesis. To Olabanji Asekun, I thank you for the invaluable support in ensuring that my academic experience within the research chair was rewarding and fulfilling.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support. They have been an important and indispensable source of spiritual support throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Contents

List of Figures

List of Tables

List of Abbreviations

AWS	Automatic Warning System
CDF	Cumulative distribution function
CMMS	Computerised maintenance management software
DSM	Design Structure Matrix
EMPAC	Enterprise Maintenance Planning and Control
FMECA	Failure Modes, Effects and Criticality Analysis
FTA	Fault Tree Analysis
HPP	Homogeneous Poisson Process
HRA	Human reliability analysis
IID	Independent and identically distributed
IMS	Integrated Management System
ISM	Interpretative Structural Modelling
LCC	Life cycle cost
LSE	Least square estimator
LTT	Laplace Trend Test
MLE	Maximum Likelihood Estimator
MTBF	Mean Time Between Failures
MTTR	Mean Time To Return
NHPP	Non-Homogeneous Poisson Process
OHTE	Overhead traction equipment
PHA	Preliminary hazard analysis
PM	Performance measurement
PRASA	Passenger Rail Agency of South Africa
RAMS	Reliability, Availability, Maintainability and Safety
RCM	Reliability Centred Maintenance
ROCOF	Rate of occurrence of failures
RP	Renewal Process
RPN	Risk Priority Number
SSIM	Structural Self-interaction Matrix

TPWS   Train Protection Warning System

1 Introduction

1.1 Background

A reliable and sustainable public transport infrastructure sustains the socioeconomic activities of a country and is the backbone of an effective and efficient public transportation system. Rail transport is a significant player in providing public transport in South Africa. The national household transport survey conducted by the Department of Transport of South Africa (DoT SA) reveals that metro workers were more likely to use trains than buses as their main mode of transport [1]. However, railway transport is competing with new modes of urban transit characterised by on-demand transit services and bus rapid transit systems. This is attributed to various factors related to rapid urbanisation, an ageing infrastructure, and increasingly high demands from customers for infrastructure service quality and reliability. To respond to these challenges requires strategies that place railway transport at a competitive edge over other modes of transport. As a result it puts pressure on railway organisations to be innovative in developing well-informed maintenance management strategies for their railway infrastructure assets to guarantee high quality of service. In addition, railway infrastructure assets have high asset value which makes maintenance efforts highly valuable. Therefore, it is important to determine intervention policies in railway infrastructure environments that would achieve the required performance targets at minimum costs [2].

The first of two factors considered to maintain infrastructure quality is the ability to measure the quality of infrastructure on a continuous basis. Secondly there must be criteria to establish the appropriate maintenance and management strategies to restore the infrastructure quality when it falls below acceptable levels. Railway infrastructure assets, however, cover large geographical areas which presents challenges in the maintenance and management of these infrastructure assets. Traditionally, the maintenance and management of railway infrastructure assets consisted of 'blind' periodic inspections on critical maintenance issues based on the knowledge and experience of maintenance staff [3]. This approach is not consistent and cannot continuously capture the performance of infrastructure quality over time. In order to operate a system of high complexity with minimal interruptions, informed decision-making becomes a strategic element in improving the maintenance and management strategies.

Following the success of a reliability centred approach in various industries, developments in the railway industry show that railway organisations are adopting this methodology in their

maintenance and management processes to reduce operational expenditure while maintaining high standards of safety. To inform optimal maintenance interventions and repair policies, systematic evaluations using reliability-centred methods have been applied at different levels of the railway infrastructure system[2], [4]–[11]. Similarly, reliability analysis for modelling the maintenance and management of individual railway infrastructure asset groups have been extensively covered in research [12]–[21]. Carratero et al [3] and Pedegral et al [22] have presented methodologies that combined reliability centred and predictive maintenance techniques to railway systems with the aim of achieving high levels of service quality. These various methodologies demonstrate the application of a reliability centred approach in improving maintenance and management processes. Additionally, a reliability centred approach aids in predicting the technical condition and remaining useful life of railway infrastructure assets allowing appropriate interventions to be implemented [23].

1.2 Research problem

To facilitate effective maintenance and management of infrastructure assets in railway environments, studies have shown that a holistic approach to improving the reliability of railway infrastructure systems simultaneously improves the lifecycle cost performance of infrastructure assets[2], [4], [5]. Reliability models that have been developed and applied in the South African passenger railway industry focus on modelling individual subsystems of the railway system such as rolling stock and infrastructure subsystems [14], [24], [25]. In addition, the current asset management strategy in the South African passenger railway industry does not utilise holistic reliability-based methodologies to support maintenance and management activities. Improving the reliability of one component of a railway system does not contribute toward whole systems improvement. Instead, different behaviours emerge at the interfaces of the different railway infrastructure asset groups due to the different functional and operational characteristics. Improving the decision making process of complex infrastructure systems spread over wide geographical areas requires methods to assess how an intervention on a single asset group impacts other parts of the railway system [26]. Furthermore, identifying high priority components that influence overall system performance provides guidelines for effective system improvement allowing railway organisations to align strategic objectives of the different asset groups towards maintaining the railway network at the expected operational levels.

1.3 Research aim and objectives

The study proposes a holistic systematic analysis to model an evidence-based decision making tool to improve the maintenance and management of railway infrastructure assets using reliability analysis. The holistic systematic analysis addresses the practical application of reliability theory in the passenger railway sector and the joint dependability implications of decision making in railway infrastructure asset management. To achieve the research aim, the objectives of the study seek to:

a) Develop a reliability model to evaluate the reliability performance of railway infrastructure systems;

b) Conduct a case study on the applicability of a holistic reliability-based approach to infrastructure asset management in the Passenger Rail Agency of South Africa (PRASA).

1.4 Scope and limitations

1.4.1 Scope

The scope of the study focused on the maintenance and management of railway infrastructure assets in the South African passenger railway industry. The study will develop a reliability assessment model to evaluate the reliability performance of railway infrastructure assets to assist in predictions for effective and efficient maintenance planning.

1.4.2 Limitations

The research is limited to the reliability performance assessment of railway infrastructure systems. The analysis methods and models only considered the reliability performance of infrastructure assets to reduce the operational expenditure related to maintenance planning and not profit making. The assessment will only focus on identifying critical infrastructure subsystems to assist in railway infrastructure asset management. Application of the model to a case study to verify the applicability of the reliability model in evaluating the performance of railway infrastructure assets is limited to railway lines with sufficient asset failure data.

1.5 Research design and methodology

This thesis is a documentation of applied research, with the objective of developing an evidence-based decision making tool to support railway infrastructure asset management using a reliability centred approach. To meet this objective, both exploratory and descriptive research methodologies were followed. The exploratory research helped in building up the knowledge required to address the research problem by exploring the key issues and variables related to system and component reliability and the effect of maintenance management decisions on the performance of infrastructure systems. Additionally, the exploratory research identified the

3

different infrastructure asset management practices and infrastructure modelling techniques required to build the reliability model that was applied to the case study. The development of the modelling approach and the application of the model to the case study are outcomes of the descriptive research which utilised elements of both qualitative and quantitative research. The quantitative research was utilised to quantify the reliability performance of the infrastructure systems using the appropriate reliability and statistical theory on the collected data. Qualitative research was primarily explanatory and was utilised to present the trends in reliability measuring techniques applied to railway infrastructure asset management. Additionally, the qualitative analysis presented the reliability model and discussed the outcomes of the relationship between the theory and research outcomes. A summary of the methodology is given in Figure 1-1.

Exploratory research → Research problem and objectives → Quantitative and Qualitative Descriptive research → Relibaility Modelling → Report research findings

Figure 1-1 : Research design and methodology

The research design shown in Figure 1-2 guided the development of a model for reliability-informed decision-making by following an inductive and deductive approach. Generally the inductive and deductive approaches are associated with qualitative and quantitative research respectively. To build a holistic reliability model requires a thorough definition of the system boundaries, a rigorous elicitation of the system data and the integration of that data to create a model. To achieve this a deductive approach was used to generate relationships between system entities and their attributes according to functional and operational requirements derived from logical conclusions based on the existing modelling theories. In addition, the deductive approach was used to build the theoretical frame of reference required for the research through an extensive literature survey and consultations with maintenance experts from PRASA.

The inductive approach focused on the problem solution by applying the developed reliability model to a case study using the developed knowledge base and empirical data. The empirical data consisted of historical asset failure data collected from PRASA Metrorail Information Management System (IMS) and from a series of interviews and consultations with maintenance experts from PRASA Metrorail division. By developing coherent ideas governed by the assumptions which align with the modelling methodology, the inductive and deductive approaches outlined the anticipated outcomes of the reliability model and provided conclusions on the behaviour of the system. In addition, the relationship between the theoretical (model) results and the observed values validated the model for improvements from a reliability-informed perspective.
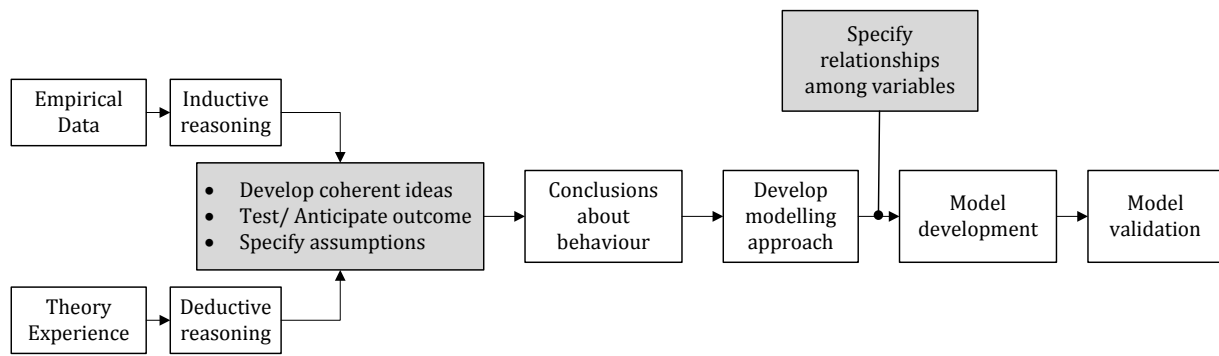
Figure 1-2 : Process of model development and validation [27]

1.6 Structure of thesis

The structure of the thesis shown in Figure 1-3, highlights the key themes that inform the scope of the study. The first section is an introduction which provides a background study to the research problem and highlights the research design and methodology followed by the researcher. The second section of the thesis provides a literature study of transportation systems, highlighting the importance of a healthy transport infrastructure system. This section also describes the railway infrastructure system and presents various asset and performance management systems. The third section provides a literature study of the methodologies employed in modelling the reliability of repairable infrastructure systems. In addition, the reliability model for railway infrastructure systems developed in the third section is applied as a case study in the fourth and final section of the thesis.
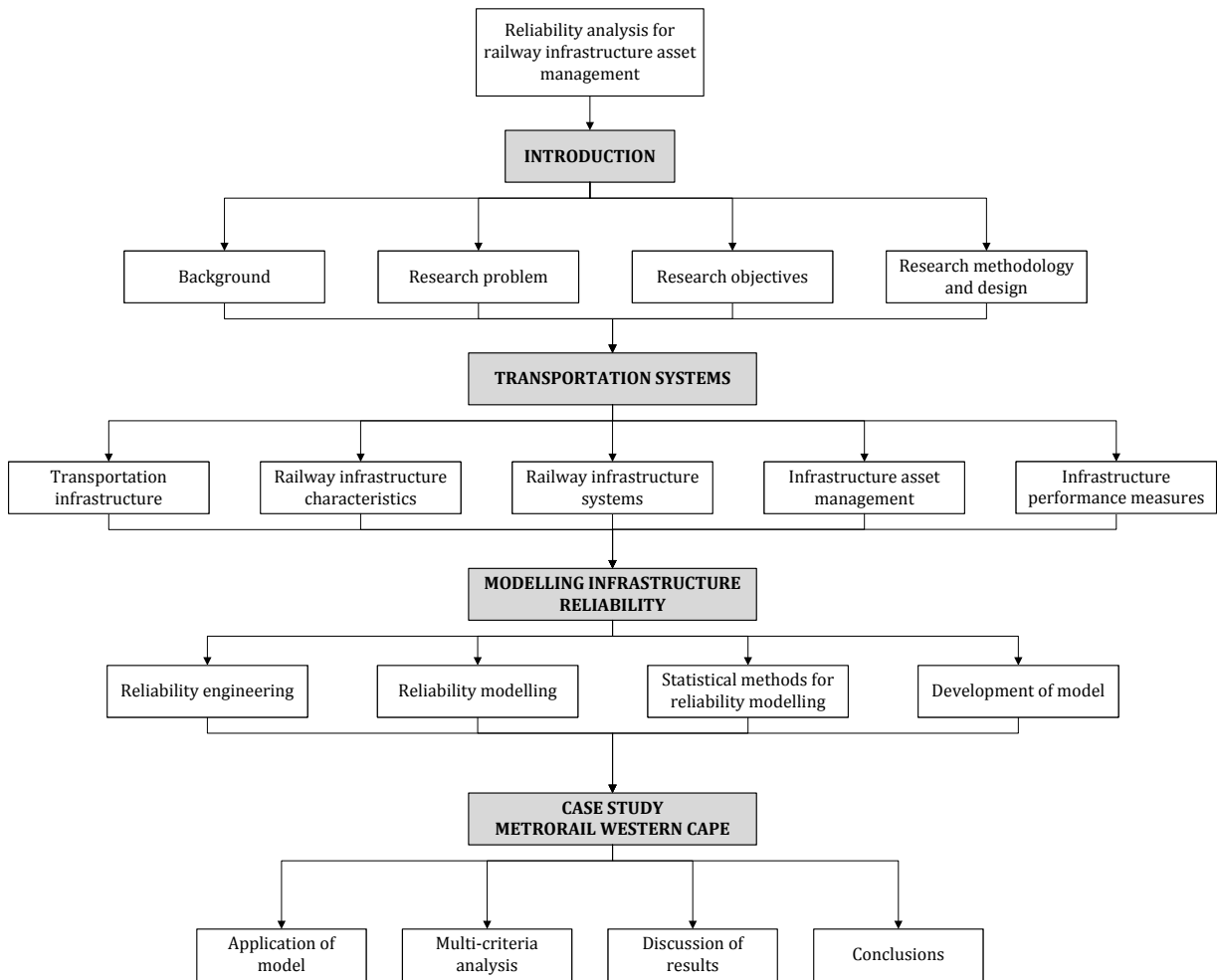
Figure 1-3 : Structure of thesis layout

2 Transportation systems

2.1 Transport infrastructure

A transportation system must guarantee the movement of material objects in time and space. The main function of any transportation process is to move people and goods from one point to another on time, safely and with minimum negative impact on the environment. The different modes of transportation processes have distinct functional, service and operational characteristics which create the core of a mobility system [28]. A mobility system is a collection of civil transport systems that satisfy the needs of a transportation process. The function of a transportation system in meeting the demands of a mobility system depends on several socio-economic factors which are external to the transportation system and its supporting infrastructure.

There is a substantial difference between the different types of civil transport systems. Surface transport systems such as rail and road require infrastructure that spans large geographical areas. Transport infrastructure refers to all the routes and fixed installations that allow for the safe and timeous circulation of traffic. It follows that an unhealthy transport infrastructure is an obstacle to achieving the fundamental goals of a transportation process. There are several challenges to managing transport infrastructure, primarily because once the design and installation is complete it becomes difficult to modify the initial design of the infrastructure assets. Providing a transport infrastructure that is resilient enough to keep up with the increasing mobility needs and resource constraints, depends on maintenance and renewal decisions. Under these circumstances, infrastructure maintenance and management processes should be efficient and effective to guarantee functional and reliable civil transportation systems.

2.1.1 Characteristics of railway infrastructure

A definition of railway infrastructure as given by the European community regulation 2598/1970 comprises routes, tracks, and fixed installations that enable the safe circulation of trains. This definition lists 70 railway infrastructure items ranging from signal systems, power systems, engineering structures (bridges, culverts), and track structures such as turnouts and tunnels. Due to the nature of railway infrastructure system and its complex configuration of multiple components, it is the objective of this study to identify infrastructure components that will form the basis of the modelling framework. To establish the scope of a railway infrastructure system, the elements that characterise the function and structure of the system need to be established. Network Rail's [26] infrastructure asset management strategy classified their assets into ten categories, among them signalling, track, electricals, level crossing and telecoms. Patra [29]

7

mentioned three distinct subsystems when presenting a maintenance decision support model for railway infrastructure; the track system, power system and the signalling system. Apart from the station buildings, marshalling yards and warehouses, the fundamental infrastructure subsystems that primarily enable the movement of a train between two points are signals, electricals, and the permanent way shown in Figure 2-1. A brief discussion of the subsystems and their functions follows.
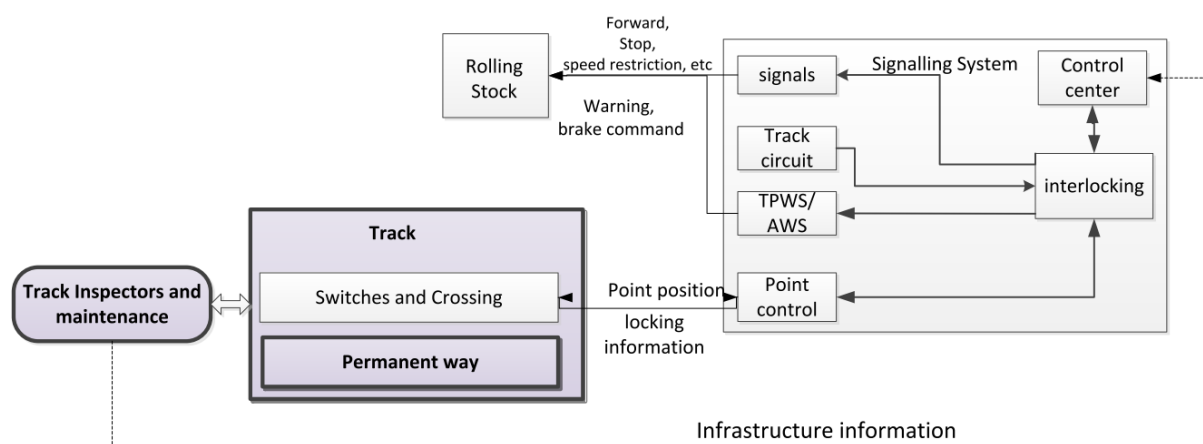
Figure 2-1 : Railway system structure [30]

2.1.1.1 Permanent Way (Perway)

The permanent way is comprised of the superstructure and substructure. Figure 2-2 shows the elements that form the core of the perway subsystem. The superstructure consists of rails, sleepers, rail clippers, and rail pads. The rails are longitudinal steel members that directly guide a train's passage. To resist excessive deflections during operation, the rail must have sufficient stiffness to serve as beams which transfer the concentrated wheel loads to the sleeper supports. The rails fastened to sleepers by rail clippers and rail pads provide damping to reduce the severity of periodic loading caused by the rolling stock. The substructure consists of the ballast, sub-ballast, and formation layer which provides drainage and support to distribute stresses caused by the superstructure. The structural integrity of the track depends on the performance of the ballast hence employing periodic maintenance routines such as ballast tamping maintains high levels of infrastructure performance.
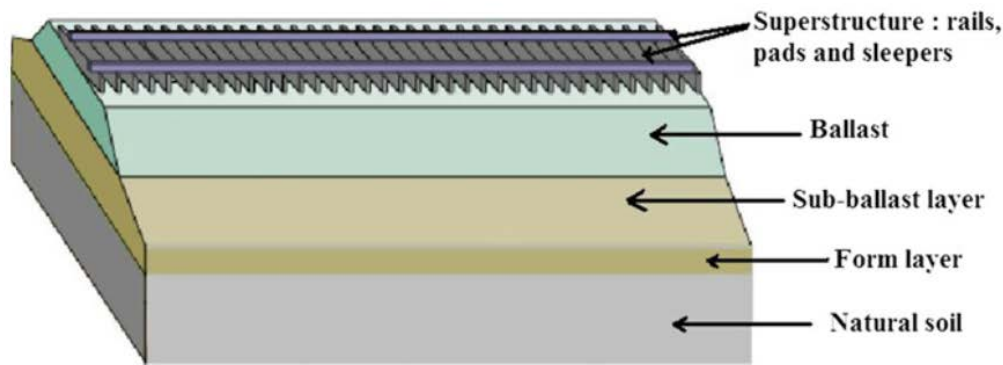
Figure 2-2 : Elements of a railway perway system

2.1.1.2 Signalling

The signalling subsystem is a complex multi-component system comprising hardware and software systems with a primary purpose of traffic control and maintaining traffic regularity. Due to the development of high-speed rail, signalling has become an important technological component in ensuring safety by preventing the occurrence of accidents hence minimising the risk to passengers [17], [31]. The performance of railway signalling systems is determined by the correct functioning of a number of subsystems. The major components of a signalling system include the control centre, track circuit, interlocking system, signals, and point machines. The signal devices which include the signal lamps, track circuits and point machines are controlled by the interlocking system [30]. Figure 2-3 shows the structure of point to point machine. Other important elements of the signalling subsystem include the protection system which contains the Train Protection Warning System (TPWS) and the Automatic Warning System (AWS). The track circuit used to establish the occupation of a railway block by a train can detect broken rails. The control centre manages train scheduling, timetables and assigns speed restrictions (including both temporary and permanent speed restrictions) for the trains. The interlocking system sends the commands to the signals, point machines and the protection system.
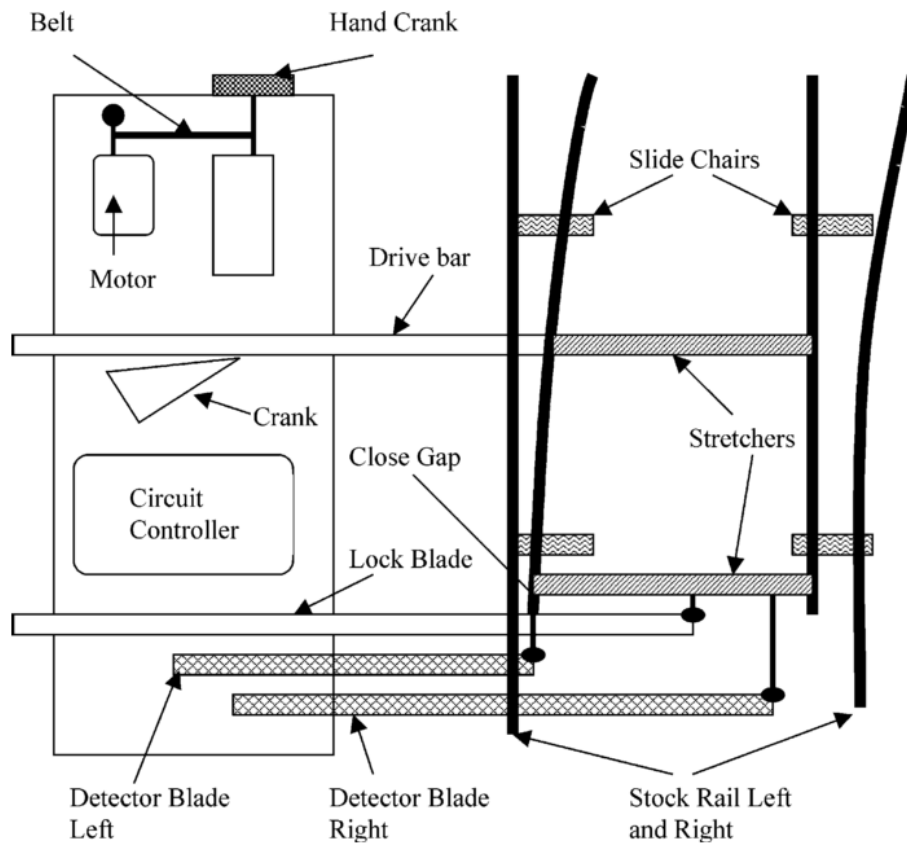
Figure 2-3 : The structure of a point machine [30]

2.1.1.3 Electrical subsystem

The electrical subsystem is an integral component in the electrified railway system. The electrical subsystem consists of all fixed installations that are required to supply traction power to the rolling stock as well as electrical power for the signalling subsystem. The electrical subsystems consist of transmission lines, substations, sectioning points and overhead contact wires. Substations are connected to the primary power utility grid. Electrical power is transmitted via transformers onto the overhead line electrification [32]. Sectioning points located at intermediate locations between substations supply parallel contact lines and provide protection, isolation, and auxiliary supplies. The overhead contact line is equipped with manually or remotely controlled disconnectors which are able to isolate sections or groups of the overhead contact line depending on the operational necessities. Feeder conductors, contact conductors (which make contact with the pantograph), suspension wire ropes, and circuit breakers are other elements of an electrified railway system. Figure 2-4 shows the elements of the electrical subsystems.
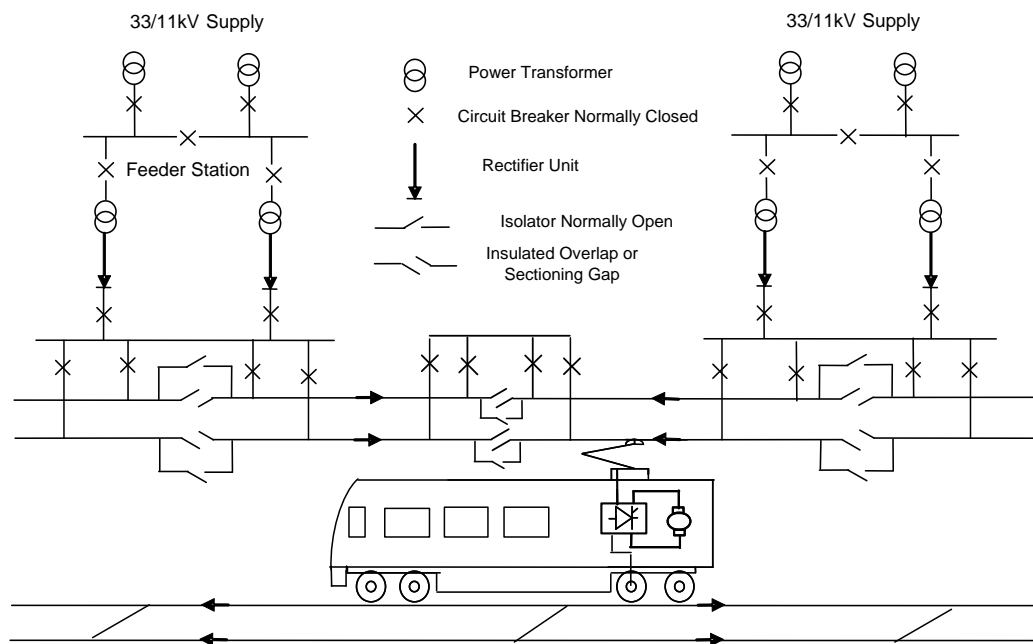
Figure 2-4 : Elements of an electrified railway system

2.2 Infrastructure asset management

The definition of asset management varies with the scope. Literature shows that there are two categories that determine the scope of asset management. The first category defines the scope of the physical assets on which the management processes are applied. The second category defines the decisions and activities that connect the high-level strategies for the asset to the actual work being done on the ground. With these two categories, a formal definition of asset management can be given as the systematic process guiding the acquisition, use, disposal of assets and coordination of activities and practices which enable an organisation to make the most of their service delivery potential in line with the organisational strategic plan. When analysed from a facilities and infrastructure perspective, infrastructure asset management can be seen as a framework that facilitates informed decision-making in maintaining, upgrading and operation of physical assets [33]. Infrastructure asset managers are thus tasked in the operational phase with delivering reliable, available, maintainable and safe infrastructure assets with minimum life cycle costs [2]. A chain of strategic and operational decisions are recognised in such an exercise. From this perspective, it can be established that infrastructure asset management focuses on achieving maximum infrastructure outputs directed at satisfying the expectations and requirements of key stakeholders. Furthermore, infrastructure asset management is concerned with the development of strategies relating to asset selection, inspection and intervention strategies within the constraints of the internal and external factors of an organisation.

Formerly, asset management when applied to infrastructure usually focused on return on investment. It has, however, evolved to introduce new tools and most importantly it now links the use of information for different functions of an organisation. Asset information can be regarded as

11

a fundamental asset on its own as it supports good asset management practices. This is highlighted by Grigg [34] who defines asset management as 'an information-based process' used for life cycle asset management. The gathering of information relating to the performance and the condition of infrastructure assets is an important part of an asset management process. Flintsch & Bryant [35] highlighted that data collection, data management and data integration are essential parts of an asset management framework. Collecting asset information provides an understanding of lifetime characteristics of infrastructure assets. This can assist in quantifying the impact of how planned interventions on an asset group influence other parts of the infrastructure system. An effective asset management system must deliver infrastructure outputs with cost savings without the risk of compromising safety.

The International Union of Railways (UIC) [36] suggested an asset management framework which identifies the key elements of an asset management system. These key elements of the asset management system focus on the core decisions and activities that link strategy to the delivery of the work. To achieve this, there must be mechanisms such as accurate data collection on asset information. This information is used to develop reviewing mechanisms that can monitor and improve the effectiveness of the asset management regime in meeting its objectives. Network Rail [26] emphasised that asset management enables evidence-based decision-making by utilising the knowledge of how assets degrade and fail to maximise the outputs of maintenance and renewal interventions. Federal Highway Administration (FHWA)[35] presented an asset management system with the major elements highlighted in Figure 2-5. These elements which are constrained by the available budget and resource allocations look at the goals and policies of an organisation. An inventory of data enables the continuous monitoring of the asset performance. The evaluation exercise on asset performance informs the short- to long-term plans and project selection criteria that align with the goals and policy of an organisation.
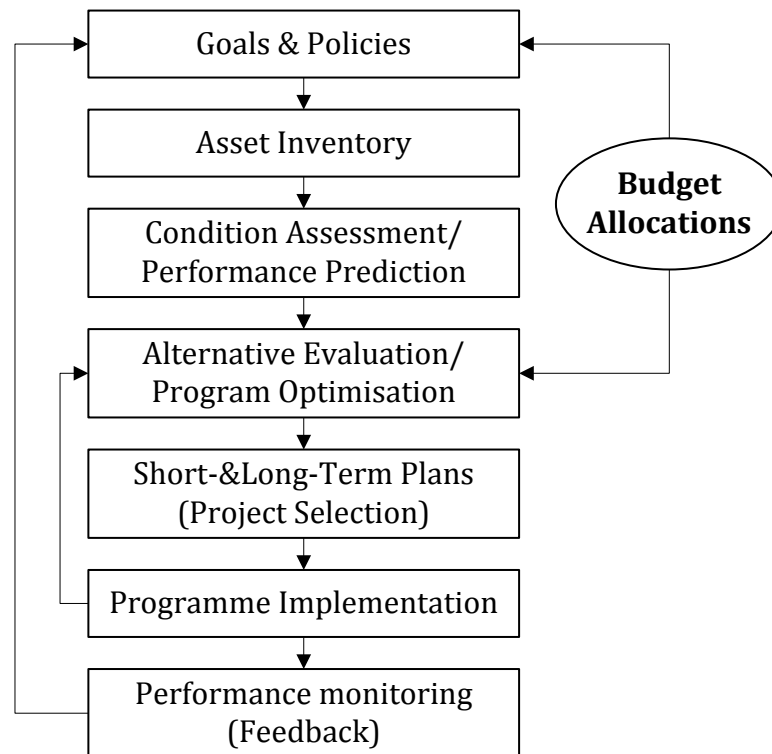
Figure 2-5 : Generic asset management system components [35]

2.2.1 Railway infrastructure maintenance management

2.2.1.1 Maintenance

Maintenance is defined as a combination of all technical, administrative, and managerial actions during the life cycle of an asset intended to retain it, or restore it to a state in which it can perform the required function. Maintenance is primarily needed because of the lack of reliability and loss of quality over time. This means minimal maintenance will result in excessive failure rates and poor performing infrastructure assets. The different impacts of maintenance on the reliability performance of assets is shown in Figure 2-6 .

13
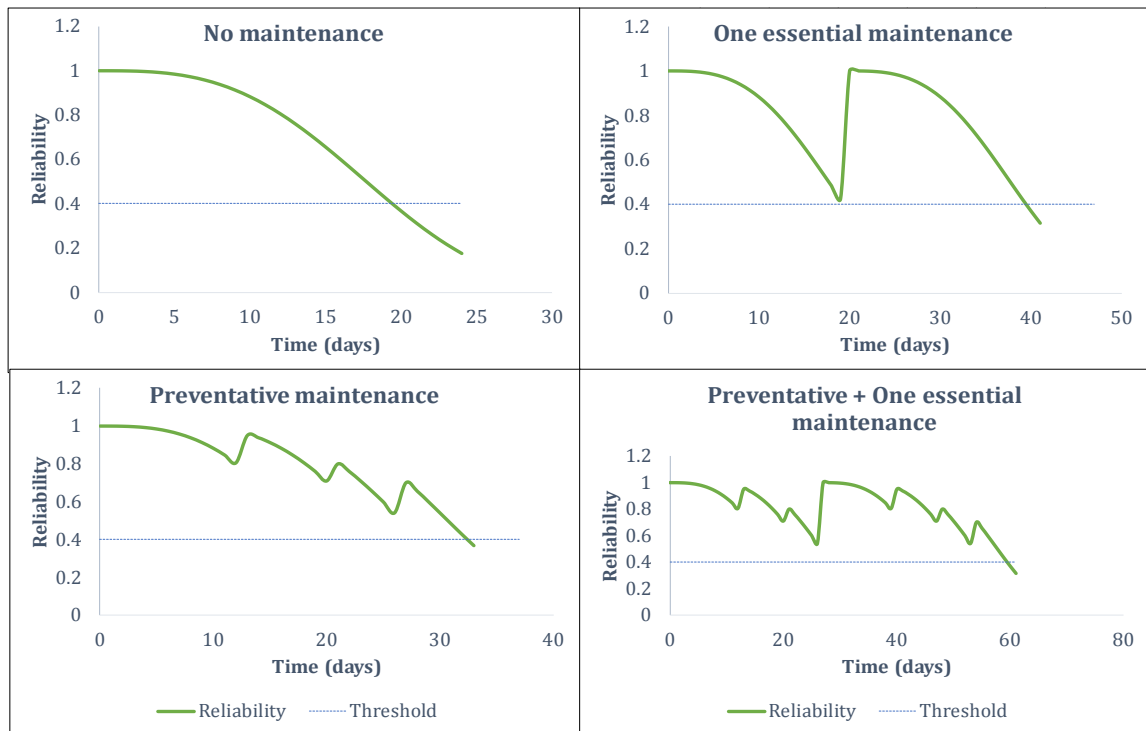
Figure 2-6 : Reliability profiles under different maintenance regimes [37].

From a basic approach, maintenance is conducted on infrastructure in either a reactive or a proactive manner. Proactive maintenance takes place at regular intervals or in many cases it follows certain criteria to restore the desired functionality. Reactive maintenance refers to the maintenance actions taken only after a system fails to meet its desired functionality. Maintenance activities can be performed either as preventative maintenance or as corrective maintenance as seen in Figure 2-7. Preventative maintenance takes place at predetermined intervals or according to specific criteria. Additionally, preventative maintenance reduces the probability of failure and degradation in a system. Corrective maintenance is carried out after a fault has been detected and can be classified as deferred or immediate. Immediate maintenance is carried out as soon as a system failure is detected whereas deferred maintenance is not immediate but is postponed either due to strategic reasons or external uncontrollable factors [38].

14
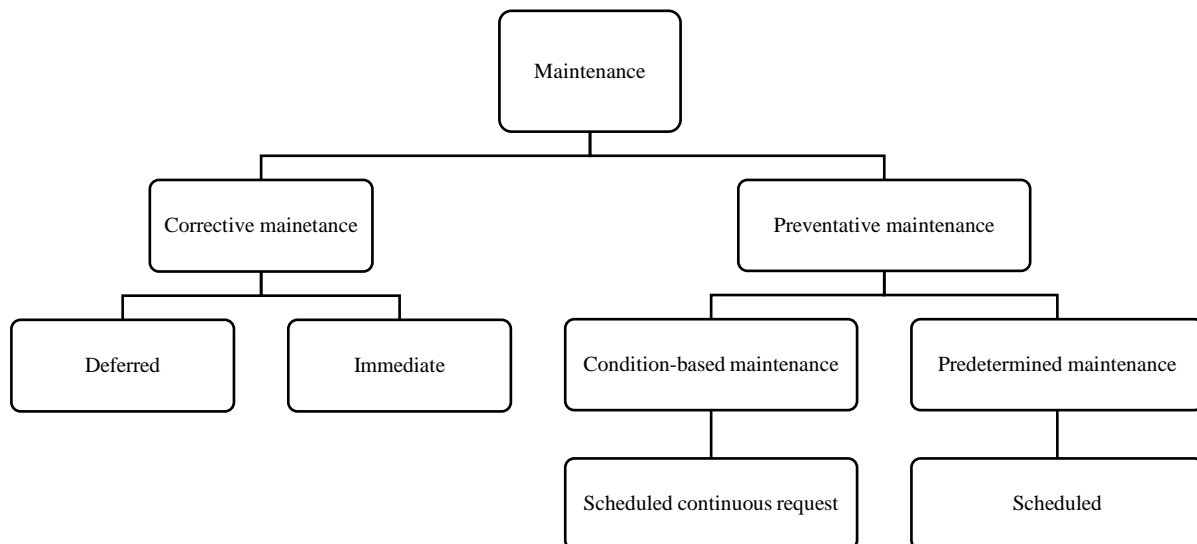
Figure 2-7 : Classification of maintenance processes[39]

2.2.1.2 Maintenance management

Maintenance management supports the planning and scheduling of the maintenance and capital improvement activities. Muyengwa and Marowa [40] highlighted that maintenance management and reliability are associated with an organisation's competitiveness and must be awarded adequate attention in the organisation's strategic plan. Maintenance management thus becomes an important component of infrastructure asset management. Maintenance management's sole purpose is to maximise system availability at minimum costs by reducing the probability of equipment or system breakdowns [41]. From an overall approach, the management of any maintenance process is described as the management of available maintenance resources such as capital, material, personnel, and information to guarantee the desired result in terms of high physical asset integrity. Managing unexpected inputs, undesirable outputs, system anomalies, or unwanted events follows a course of action and series of stages that must be followed to describe and implement the correct strategies. To achieve this entails the setting up of goals and strategies, planning, execution, analysis and continuous improvement of the process through evaluations. Figure 2-8 shows the general maintenance management process for Rete Ferroviaria Italiana (RFI) [5]. This maintenance management strategy is based on the implementation of maintenance planning and the control cycle requires maintenance plans to be customised for the different cluster of railway assets that are subject to different operating conditions.
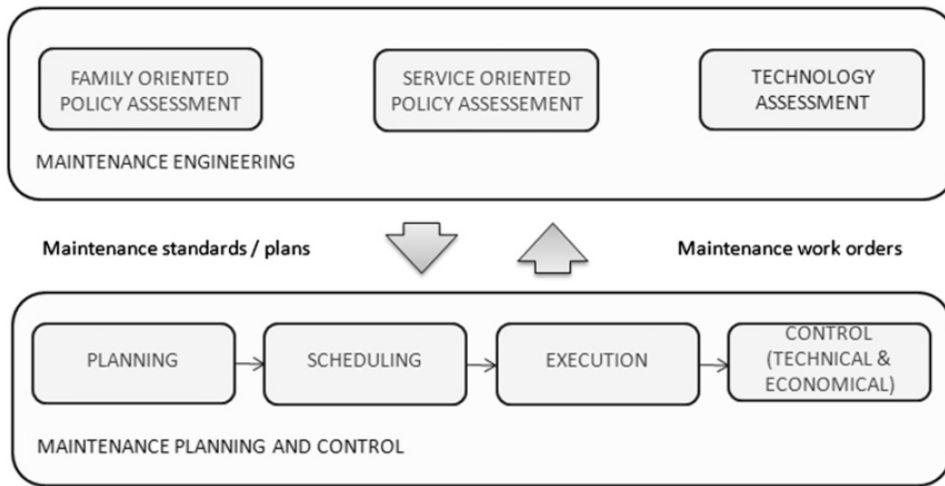
Figure 2-8 : General maintenance management process for RFI [5].

An effective maintenance management strategy ensures the successful management of costs and quality and their relationship to asset performance. Figure 2-9 shows the relationship between maintenance management, asset performance, and asset maintenance. To manage performance it needs to be measured, hence performance indicators are utilised to reflect the performance of complex systems. Quality indicators for asset performance are interpreted through cost and system effectiveness; these indicators act as decision tools for the different interventions specific to asset maintenance [42]. To assess if the maintenance management process supports the overall objectives of the organisation, performance measurement systems are adopted to generate useful information on the condition of infrastructure assets [41]. Infrastructure performance measurement systems will be discussed in section 2.3.2.
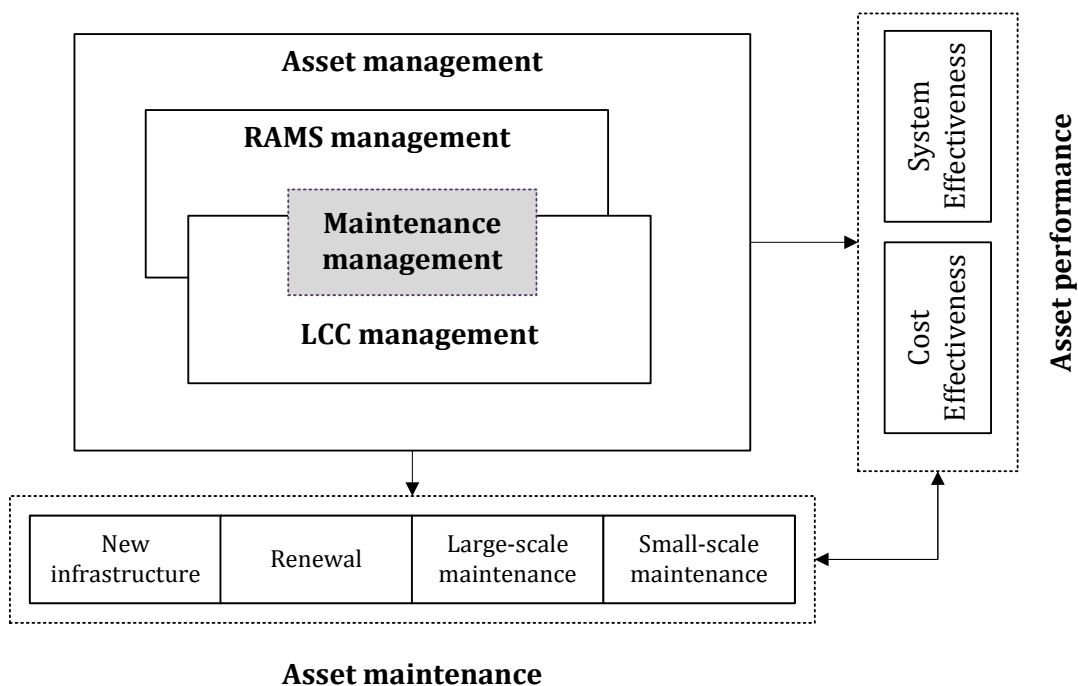
Figure 2-9 : Factors influencing maintenance management

16

2.2.2 Reliability centred maintenance

Reliability Centred Maintenance (RCM) has its origins in the airline industry and can be defined as a systematic approach to systems functionality, failures of the functionality, causes and effects of failure and infrastructure affected by failures [22]. The RCM approach takes into account the consequences of failures by classifying them into safety and environmental, operational (delays), non-operational and hidden failure consequences. This classification of failure consequences can then be used to create a strategic framework for maintenance intervention strategies for infrastructure systems. Essentially an RCM approach seeks to balance high corrective maintenance costs with those of programmed maintenance interventions (preventative or predictive). Figure 2-10 shows the principle objective of the RCM philosophy. The objective seeks to integrate preventative, predictive maintenance, condition monitoring and run-to-failure techniques to improve system dependability with minimum maintenance intervention. To achieve this objective the RCM firstly seeks to enhance the safety and reliability of systems by highlighting and establishing the system's most important functions. This implies that an RCM approach is concerned with a loss of function. Secondly, the aim of the RCM approach is not to prevent failures from happening but rather to prevent and reduce the consequences of failures on the performance of the system. Lastly, RCM is capable of reducing maintenance expenditure by either adding or removing maintenance interventions that are unnecessary to improving system functionality.
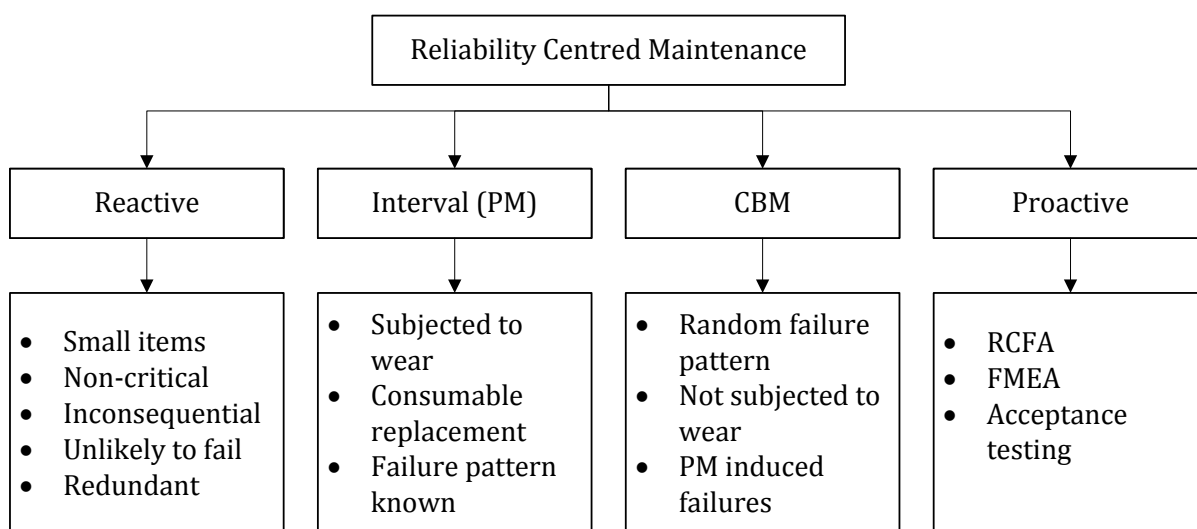
Figure 2-10 : Components of reliability centred maintenance program [43]

Applying the RCM methodology to railway infrastructure systems as part of the RAIL project, Carretero et al [3] developed an RCM framework that could be applied to railway infrastructure maintenance. This framework was later adopted by the Spanish railway company (RENFE) and the German railway company (DB A.G.). Jidayi [24] highlighted the benefits of applying an RCM approach to railway infrastructure maintenance management which included improvement in system reliability, availability and, most importantly, a reduction in the life cycle costs of railway

infrastructure related to safety. Gonzalez et al [9] explicitly modelled the uncertainty that characterises the deterioration rate of railway infrastructure and developed an optimal maintenance and repair policy for a railway network using an RCM methodology.

2.3 Infrastructure performance measures

The railway system, being a transportation process, must achieve a required quality of service at any given time. The infrastructure system must meet the expectations of the defined level of service which invariably depend on the different elements and operations of the railway system. To assess if the infrastructure meets these expectations, the performance of the infrastructure must be measured and can be expressed as a function of effectiveness, reliability and costs[44].

$$Infrastructure\ performance\ = F(effectiveness, reliability, cost)$$

Infrastructure that reliably meets or exceeds the quality of service expectations at low cost is performing well. From the perspective of an organisation, the reliability of infrastructure is the likelihood that infrastructure effectiveness will be guaranteed over an extended period. On the other hand, from the perspective of the customer, reliability is the probability that a service will be available at least at the specific times during the design life of the infrastructure system. Infrastructure performance captures the ability to move goods, people, and a variety of other services that support economic and social activities. In this regard, infrastructure is a means to an end. The effectiveness, efficiency, and reliability of its contribution to these other ends must essentially be the measures of infrastructure performance.

Performance measurement is the process of using a tool or a procedure to evaluate an efficiency parameter for a system. On the surface, performance measurement in infrastructure may seem straightforward but in reality, it is influenced by a number of factors. A well-designed performance measurement (PM) system is a management and improvement tool that can be utilised as a basis for decision-making by the strategic, operational and tactical levels of management [45]. Performance measures must thus be based on the criteria that correspond to the desired outcome of an infrastructure asset strategy. This section introduces a discussion on the connection between performance measurement and reliability. Thereafter, a discussion of infrastructure performance measurement systems will be introduced.

2.3.1 Performance measures and reliability

Measuring is a management tool which facilitates and supports effective decision-making. In and of itself, it does not determine performance but can facilitate good management. The term measurement entails an approach that is rigorous, systematic, and quantifiable. There are two distinct approaches to measuring performance; quantitative and qualitative. A quantitative approach produces data that provides insight on facts and figures and employs the use of statistical data analysis, whereas qualitative methods seek to explain, understand, and evaluate the causes of an outcome. Stenstrom [46] highlighted that it is not possible to measure everything with only qualitative and quantitative methods. The qualitative and quantitative techniques are both required in order to create a measurement system that is as complete as possible. Qualitative measurement methods can be used to check conformity with quantitative techniques.

The performance of an asset is a result of an execution of various programs that have an ultimate goal of improving its performance. These programs include asset management interventions, maintenance and performance measurement models that can be used to evaluate the impact of the intervention processes. Infrastructure asset management is an information-based process. As such, the most common approach in developing these programs utilises empirical evidence (quantitative data) collected during the investigation of failures. The performance of an asset can be outlined by four distinct elements which are:

- Capability – The ability to perform the intended function on a system basis;

- Reliability – The ability to start and continue to operate;

- Efficiency – The ability to effectively and easily meet its objectives;

- Availability – The ability to quickly become operational following a failure.

From these distinct elements, it can be observed that capability and efficiency are measures that are determined and influenced by the design and construction of the infrastructure asset. Essentially, capability and efficiency reflect the levels to which an infrastructure asset is designed and built. Reliability, on the other hand, is related to the operation of a component and is influenced by its ability to remain operational. In some cases, an asset can achieve high reliability levels but fail to achieve high performance. This occurs usually when the asset fails to meet design objectives. On the other hand, reliability and availability are the building blocks that ensure high asset performance. A conceptual hierarchy for an integrated approach to improving performance by way of focusing on reliability and availability is presented as in Figure 2-11. From the hierarchy, the role of reliability and availability analysis is put into context. Evidently, it can be seen that the performance of an asset can be improved through a continuous reliability improvement programme and can further increase the design life cycle of the infrastructure assets.

19
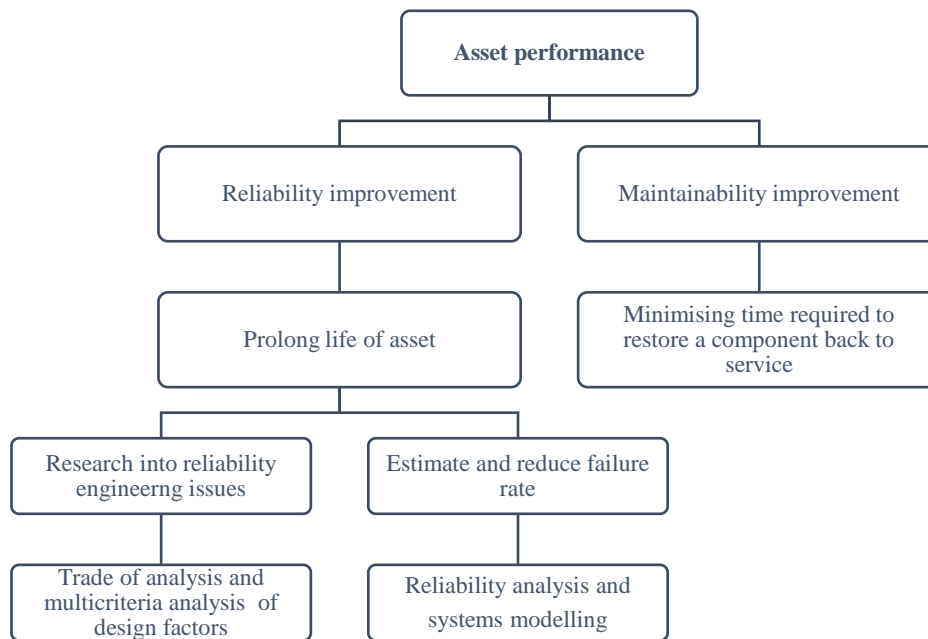
Figure 2-11 : Conceptual hierarchy for achieving high performance

2.3.2 Infrastructure performance measurement systems

Railway infrastructure assets are capital-intensive and have a long lifespan, hence the operation and maintenance requires sustainable long-term strategies. There are several stakeholders in railway operations, and as with many cases where there are multiple stakeholders, there are scenarios where the stakeholders have conflicting requirements. These can complicate the assessment and monitoring of railway infrastructure performance. The development and integration of performance measurement methods are critical to ensuring a successful performance measurement framework. A successful performance measurement system must be robust to withstand the demands that arise from organisational changes, technological developments and policy shifts.

Developing sustainable strategic plans for large complex geographically spread-out technical systems involves the collection of information, setting goals, changing the goals to specific objectives and setting up activities that enable the achievement of these objectives. The impact of the interventions on railway infrastructure assets needs to be quantified to establish their performance against the operational objectives. To achieve this, the infrastructure assets' performance is monitored and steered according to the objective of the organisational asset management strategy. Stenstrom [46] conducted a study to review railway infrastructure performance indicators that are used by researchers and professionals in the field of railway infrastructure asset management. The indicators are classified as managerial and infrastructure condition indicators as shown in Figure 2-12. Managerial indicators provide insight into the overall system-level performance while condition monitoring indicators are at the component or subsystem level. Managerial indicators are obtained from computer systems like computerised

maintenance management software (CMMS) whereas infrastructure condition indicators are extracted by sensors and other inspection methods applicable to the railway industry. Brinkman [47] interviewed ProRail's stakeholders and discovered that the most important infrastructure performance indicators are affordability, availability, reliability and safety. Therefore, cost and quality indicators form the basis of railway infrastructure management.
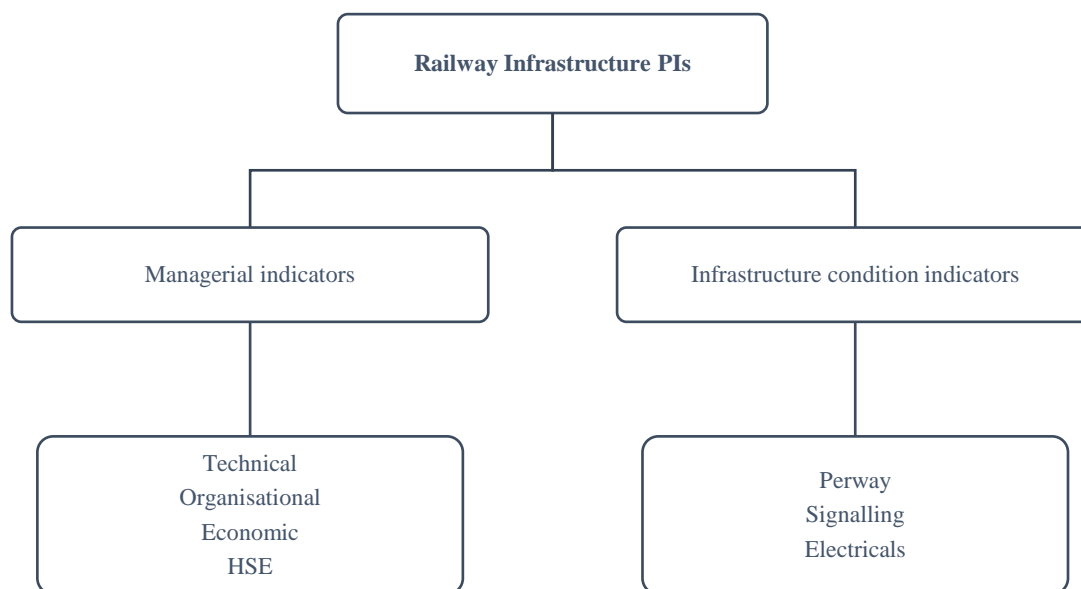
Figure 2-12 : Generic structure of railway infrastructure PIs [46]

Railway infrastructure performance indicators such as reliability, availability, maintainability, and safety are utilised for monitoring and steering the performance of railway infrastructure assets. Stenstrom [11] developed a model to monitor and analyse the operation and maintenance performance of railway infrastructure. The model recommended that performance measurement strategies need to be dynamic and versatile. To make critical decisions the performance indicators must be traced back to the root of the problem. Railway infrastructure managers place threshold values on their indicators to indicate when an intervention is required. If this approach is not used accurately, aggregated data and threshold values can make an infrastructure system reactive. To counter such a scenario, composite indicators can be used to simplify the performance measurement process because they summarise the overall performance of complex assets into a single number which is easy to interpret for decision-makers. A composite indicator called the infrastructure index was proposed by Famurewa et al [7]. This indicator was constructed based on failure frequency, train delays, and active repair time (MTTR).

An essential characteristic of performance management for railway infrastructure is the development of systematic analysis at various levels of the railway network. Patra [42] presented this by proposing an integrated approach to railway infrastructure asset management which incorporates RAMS management and life cycle costs (LCC). A systematic analysis is the core of any

21

continuous improvement program in railway operations [48]. A discussion of RAMS and its influence on infrastructure reliability will be given in the following section.

2.3.2.1 Reliability Availability Maintainability and Safety – RAMS

The concept of measuring the performance of systems is embodied in the European Standard EN50126 which requires RAMS targets to be established at an early stage in railway projects [49]. To identify these RAMS targets thoroughly, some rationale of how to achieve them has to be developed. Defining the Reliability, Availability, Maintainability, and Safety (RAMS) parameters for the entire railway system assists railway managers in executing their duties within affordable maintenance and logistical costs. RAMS analysis is a systematic analysis that can be used to quantify and categorise capacity constraints as well as improve the impact of infrastructure intervention strategies that enhance reliability. Furthermore, RAMS techniques enable reliability engineers to forecast failures from collected field data. RAMS in railways is described as an engineering discipline that comprises a set of activities that integrates reliability, availability, maintainability and safety characteristics. This set of activities that encompasses different fields of expertise is linked to the study of failure, maintenance, and availability of systems. The focus of this paper is to look at the aspect of RAMS which is reliability, within the context of railway infrastructure management. To develop a sound reliability model will require a brief look at the variables that influence reliability within the RAMS framework.

2.3.2.2 Interrelation of RAMS

Studying the RAMS framework establishes that safety and availability are considered to be outputs of any RAMS analysis. As a result, conflicts between safety and availability requirements present obstacles to achieving a dependable system [42]. Infrastructure managers can achieve high service safety and availability targets by meeting all reliability and maintainability requirements and by effectively controlling the short- and long-term maintenance operation activities. Figure 2-13 highlights the important relationships between RAMS elements and their relationship with maintenance support. Maintenance support is the ability of the maintenance department to provide the required resources for executing tasks under the given maintenance policy. The safety of a system is considered a subset of reliability in cases where the severity and risk of the failure consequences are taken into account. Safety depends on the maintainability of the system components expressed as the ease of performing maintenance procedures to restore a system into a safe operating mode. Availability is influenced by reliability in terms of the probability of occurrence of each failure mode and time to detect, locate, and restore the failure mode respectively. All failures adversely affect the reliability of a system whereas, on the other hand, specific failures will have an adverse effect on the safety characteristics of the system [42].
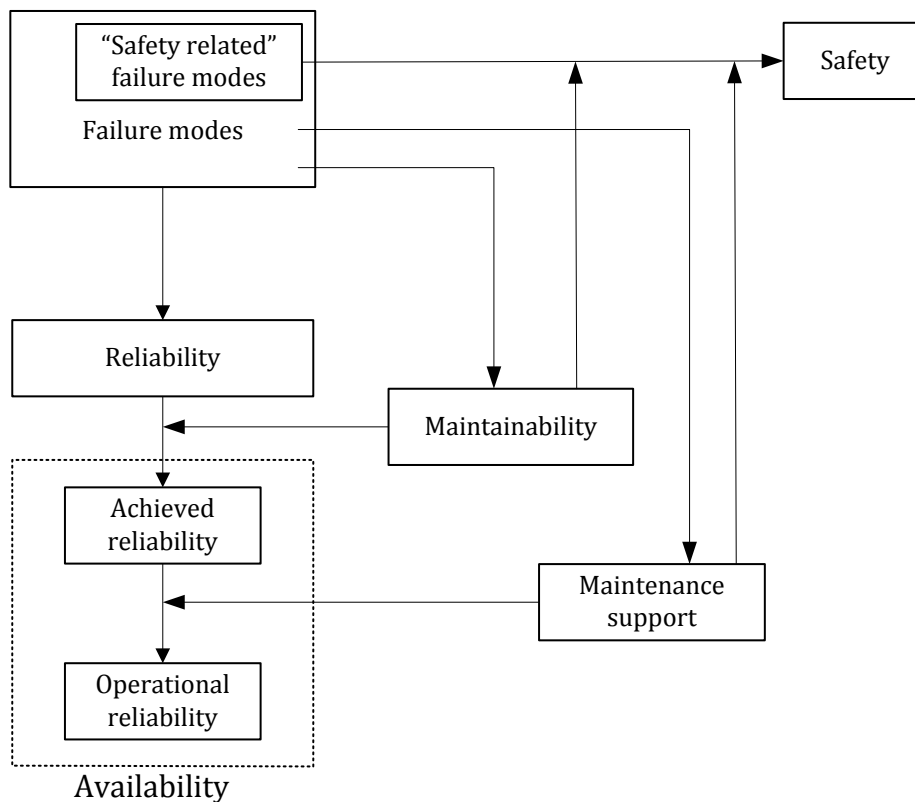
Figure 2-13 : Interrelationship of RAMS elements[42]

In order to achieve a dependable system, the external factors that influence RAMS parameters need to be identified. In railway systems, RAMS is influenced by three conditions: 1) the system; 2) maintenance conditions, and 3) operating conditions. The system conditions are sources of failures that are introduced internally in the system throughout its life cycle, whereas operating and maintenance conditions are sources of failures that are introduced during the operations and maintenance interventions on the system. These three sources of failure can interact with each other through the internal and external factors of the system and their causes need to be assessed and managed throughout the life cycle of the system. Figure 2-14 shows a simplified approach to performing a RAMS analysis which incorporates life cycle costs (LCC) according to the EN50126. A RAMS analysis is a measurement framework that utilises failure information to develop probability distributions representing a system's ability to perform the intended functions. RAMS techniques can be employed to predict failures in railway infrastructure systems and have been applied extensively to develop measurement systems for railway infrastructure maintenance management [12], [42], [50], [51].
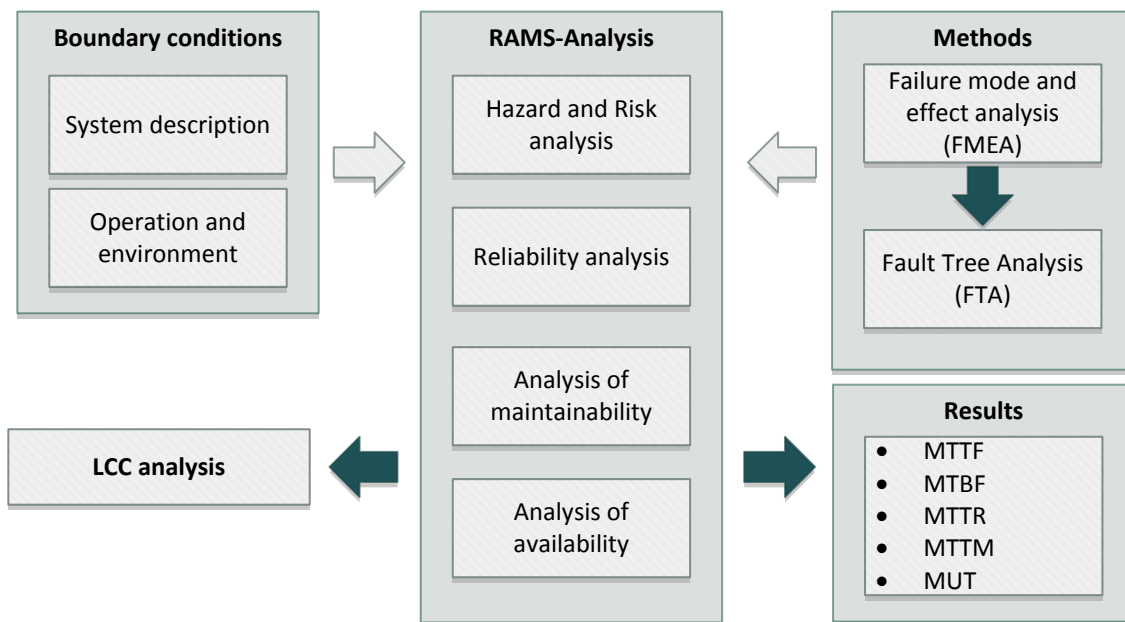
Figure 2-14 : Simplified RAMS analysis according to EN50126

2.3.3 Modelling railway performance

The central concept in systems and maintenance engineering is dependability. This is a collective term used to describe availability and the factors influencing it such as reliability, maintainability, and safety. Using the dependability approach, it is then possible to establish the input and output factors that influence railway infrastructure performance by considering the factors that influence infrastructure availability. Stenstrom [11] proposed that reliability, maintainability, supportability and maintenance interventions can be considered inputs with failure frequency, train delay, punctuality and mean repair time as outputs, as illustrated in Figure 2-15. Supportability depends on the execution and planning of maintenance interventions within an organisation, as input parameters such as preventative maintenance and train timetable scheduling influence the output parameters such as failure frequency and capacity utilisation respectively. The INNOTRACK project, Patra [42], Jidayi [24], Nawabi et al [52] and Famurewa [53] identified several indicators related to RAMS and life cycle costs for railway infrastructure. Among these indicators are the following:

- Failure frequency;

- Train delays due to infrastructure failures;

- Mean Time To Return (MTTR);

- Mean Time To Failures (MTTF);

- Mean Time Between Failures (MTBF).

24

Reliability

Maintainability (M)

Supportability (S)

Preventative maintenance(PM)

Train timetable (TTT)

Railway Infrastructure

Failure frequency	$= F(R,PM,TTT)$
Logistic time	$= F(S)$
Repair time	$= F(M)$
Train delay	$= F(failures,LT,RT) = f(R,PM,TTT,S,M)$
Railway availability	$= F(failures,LT,RT) = f(R,PM,TTT,S,M)$
Train punctuality	$= F(failures,LT,RT) = f(R,PM,TTT,S,M)$
Train cancellation	$= F(failures,LT,RT) = f(R,PM,TTT,S,M)$
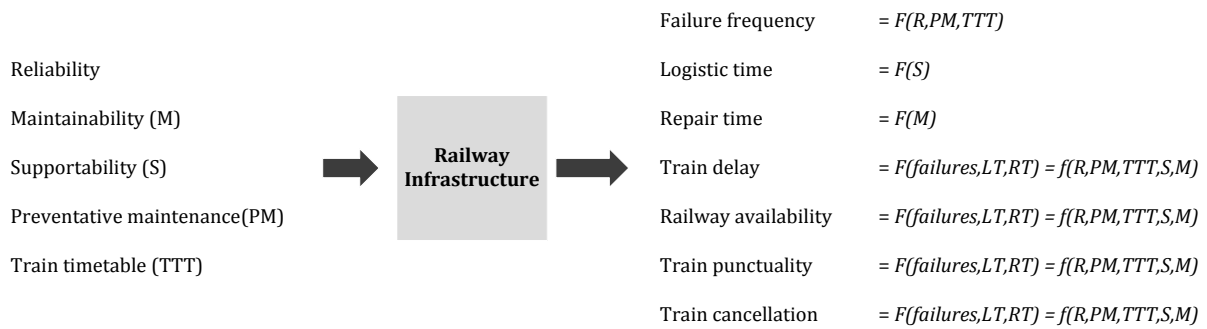
Figure 2-15 : Input and output factors of infrastructure performance [11]

The main objective of known modelling work in infrastructure reliability evaluations is to assist management by predicting the consequences of alternative decisions. A challenge to transport infrastructure managers is how to effectively measure reliability. Reliability of transportation systems is perceived in terms of travel time reliability from a passenger point of view and system availability from that of the operator [28]. Restel [54] investigated the impact of infrastructure type on the reliability of railway transportation systems; the correlation between infrastructure type and the frequency of failures and failure consequences was highlighted. Reliability theory utilises failure data in modelling and quantifying system reliability, hence with Restel's [54] findings and Stenstrom's [11] influencing factors for infrastructure availability, it is possible to map the occurrence of failures and their consequences to measure system reliability.

2.4 Section summary

This section provided a background to transportation systems and the importance of healthy infrastructure systems towards ensuring that railway systems meet their desired level of service. The methodologies employed in asset management of infrastructure systems was presented, and in addition, the performance measurement methods for transport infrastructure systems were introduced.

3 Railway infrastructure systems

The preceding section provided background on the transportation systems and characterised the different properties of railway infrastructure systems. The strategic and management issues related to infrastructure maintenance management were also highlighted. This section presents a systems perspective and the fundamental concepts of systems thinking that will enable the successful modelling of railway infrastructure systems for reliability evaluations. The section will further examine the procedures that are required in performing a dependability analysis for reliability modelling of infrastructure systems.

3.1 Systems perspective

It has been highlighted that the railway infrastructure system consists of various multiple components of varying complexity. This characteristic enables infrastructures to be viewed as systems. A system is a distinct deterministic entity comprising an interconnected and/or interacting collection of discrete components that takes in resources from its environment to process them to produce an output [33]. Infrastructure systems are a collection of assets and subsystems, which individually and collectively perform a required function. Using a systems approach the infrastructure system can be viewed as an open system consisting of interacting components arranged in a hierarchical and decomposable structure. This means the internal and external factors that influence the system can be established by studying the parameters that characterise railway infrastructure systems. The parameters that characterise railway infrastructure systems are the function, the structure, and the history of the system. Analysing the railway infrastructure system reveals that it can be described to consist of operational subsystems called domains of infrastructure. The function and structure consists of these domains made up of maintenance components of varying technological properties and complex functional configurations extending between several geographical locations. The domains are coupled with two driving systems: the first driving system controls the operations of the system while the second driving system controls the structure of the network and its infrastructure. To coordinate and guarantee the effectiveness of the two driving systems, strategic decisions need to be employed to ensure that the infrastructure system meets the expected performance requirements and to achieve this a systematic analysis of the factors that influence infrastructure performance is required.

3.2 System analysis

A system analysis is a process orientated towards the acquisition and orderly investigation and processing of information specific to the system and relevant to a decision or a given goal. The end product of the process is a model related to the attributes of system dependability such as reliability. The selection of a suitable analysis method is based on available data, dependability assessment and system engineering requirements [53]. Fleming et al [55] presented a systematic procedure which highlights the basic steps in performing a system analysis as shown in Figure 3-1. System analysis typically involves the establishing the objectives and constraints and alternative courses of action. The analysis is performed by investigating the likelihood of impacts in terms of the objective of the analysis.
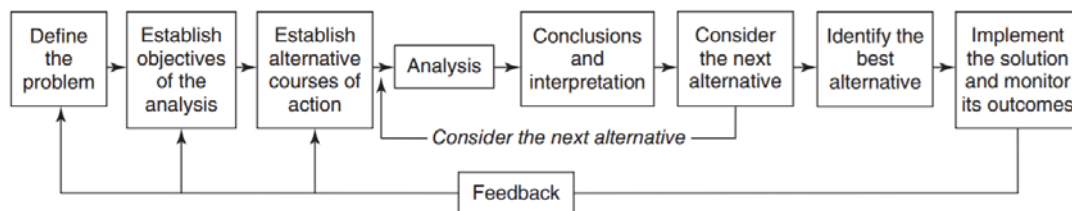
Figure 3-1 : Basic steps in a system analysis

In a study of maintenance analysis for enhanced infrastructure capacity Famurewa [48] presented a systematic analysis approach to develop an effective decision support programmed for effective infrastructure performance shown in Figure 3-2. From a technical point utilising multi-criteria criticality analysis of the different routes and lines will involve the aggregation of different indicators using multicriteria aggregation techniques. To provide a thorough analysis of the dependability of a system at the specific indenture level two approaches are identified; these are inductive and deductive approaches [56]. An inductive approach is one in which the reasoning proceeds from the most specific to the most general. Failure modes and effect analysis (FMEA) and Consequence tree methods are examples of inductive approaches. These methods analyse system failure by closely studying the effects and consequences of failures on the system itself and or on its environment. A deductive approach reasoning proceeds from the most general to the most specific. Fault Tree Analysis is an example of a deductive approach. A discussion of these methods is given in section 4.2.
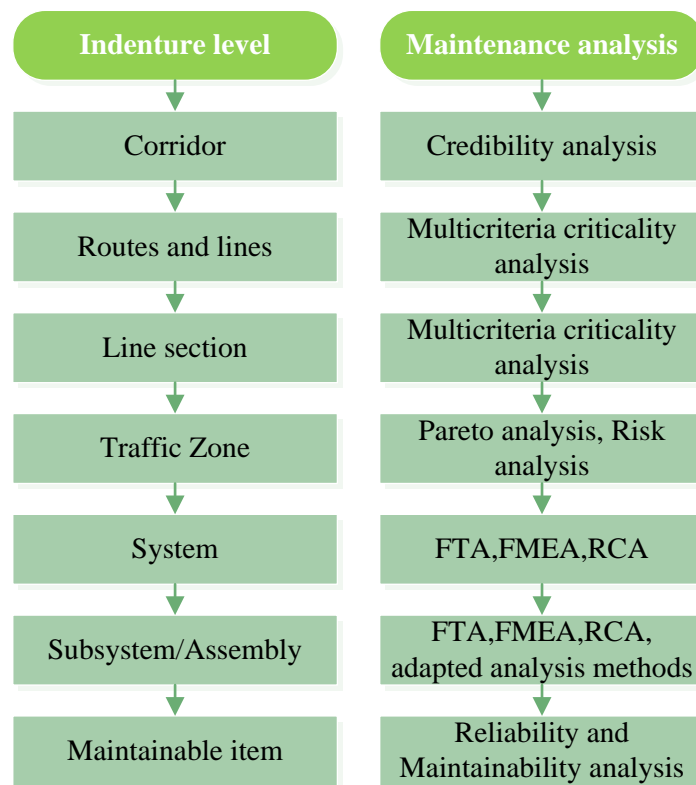
Figure 3-2 : Indenture levels for maintenance analysis for continuous improvement[53]

3.3 Systems modelling

Different modelling paradigms have been established in literature and are summarised in Figure 3-3 as time-driven and event-driven [57]. The system dynamics approach is a time-driven paradigm which involves iterative evaluations of a system of ordinary differential equations. Models developed from this approach require that the state of the system varies with time. Additionally system dynamic models are applicable in scenarios where the number of components in a system is large. For these scenarios, the system is modelled as a stream of continuous interconnected quantities of information in feedback loops. With event-driven modelling, the state of the system only changes when an event from a set of possible events occurs. Event-driven modelling focuses on the occurrence of an event describing the evolution of a system as a sequence of events. The event-driven approach simulates the simultaneous operation and interactions of multiple agents with the goal of recreating and/or predicting the appearance of a complex phenomenon. Two different modelling approaches can be employed in the event-driven paradigm. Event-driven modelling can be performed using agent-based or discrete events approach. Agent-based models, unlike discrete events, have continuous states and they use more sophisticated decision rules.
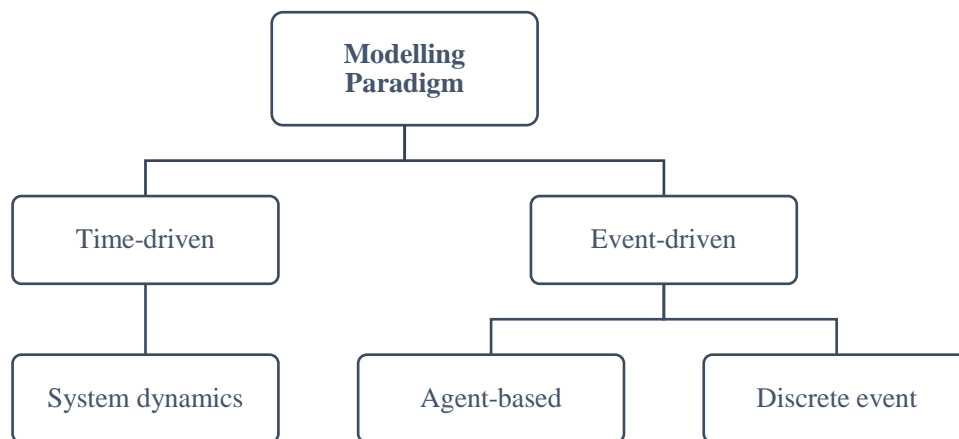
Figure 3-3 : Modelling paradigms

3.4 System dependencies

A railway network is an example of a complex system. A complex system can be defined as a system which has a structure of multiple units which work together to perform a particular function. Complex systems have different types of interactions between the constituent assets which arise from the design of the system and the intended function. This implies that reliability models for complex systems should not assume that lifetime or time to failure distributions of a systems component are statistically independent. Valenzuela [57] identified three major types of interactions in systems, which are stochastic, structural and economic dependencies. These interactions influence the operating environment of infrastructure systems. Stochastic dependence occurs when the condition of an individual asset influences the lifetime distribution of other assets within the system. Structural dependence occurs where components structurally form a part, so that the maintenance of a failed component requires or results in the dismantling of working components. This dependence can be illustrated in a railway infrastructure environment. Regular maintenance on the track and ballast may lower the track so that no contact occurs between the pantograph and the rolling stock's contact wire. In a multi-unit repairable system, the economic dependence between components of the system is said to occur if the cost of performing maintenance on the group of components is different from the cost of performing the same type of maintenance individually [57].

The methods of fault identification and criticality ranking require decomposing a complex system into subsystems, noting the relationships between the different subsystems and finally determining the internal and external factors that impact a system's performance. These physical interactions between the different subsystems need to be identified, described, and summarised in a dependency matrix. In a study of critical infrastructure interdependency modelling, Pederson

et al [58] utilised a dependency matrix to show the dependencies between critical infrastructure networks and their relative impact. In railway systems, many different fault states can occur during operation. To assist infrastructure managers and railway undertakings with their safety management systems, Andreas et al [59] developed a cause-consequence fault state matrix to describe the complex dependencies between different fault states in railway systems.

The design structure matrix (DSM) is an analysis tool for modelling and can be used for purposes of decomposition and integration of subsystems. A DSM shown in Figure 3-4 presents the relationships between the different system components in a compact, visual, and analytical format. System components are represented by the shaded elements along the diagonal and off-diagonal marks signify the dependency of one component on another. When the matrix is read across a row it reveals what other elements in the row it provides to. On the other hand, reading down a column reveals what other elements in the column an element depends on. In other words, reading down a column reveals the input source and reading across the row indicates output sinks [60].
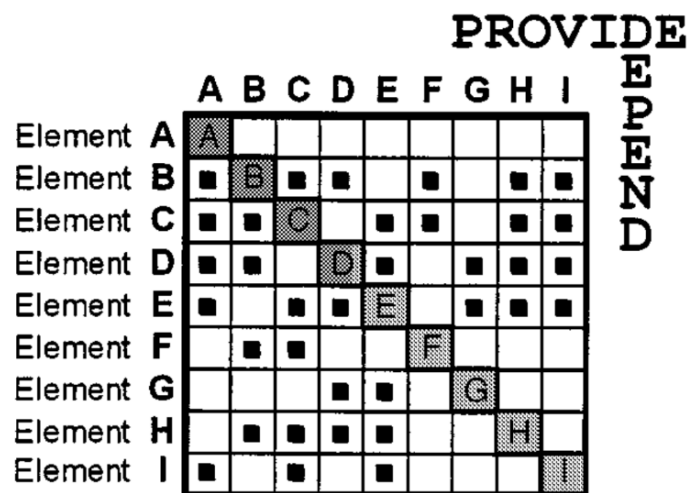
Figure 3-4 : Design Structure Matrix (DSM) Example

Interpretative Structural Modelling (ISM) is a method for analysing and identifying complex relationships by breaking down a complicated system between the various systems elements into a clear hierarchical structure. Singh and Gupta [61] identified critical infrastructure sectors and their dependencies using the ISM and structural self-interacting matrix (SSIM) to develop hierarchical relationships among the system elements. The SSIM defines the nature of relationships between components in a system by establishing whether a relationship exists between two infrastructures *i* and *j* and further determines the direction of association given that a relationship exists. Figure 3-5 shows an example of an SSIM with 8 elements, the symbols V, A, X and O show the type of relation that exists between the elements.

V – Infrastructure j depends on infrastructure I

A – Infrastructure I depends on infrastructure j

X – Infrastructure I and j are interdependent

O – Infrastructure I and j are unrelated

From the SSIM a reachability matrix is developed which is then partitioned into different levels upon which ISM is used to build a structural model. ISM has been used to evaluate the service quality of railway passenger trains to guide the improvement process of railway service quality for passenger trains [62]. These different approaches can be used to identify the dependencies in modelling the reliability of railway infrastructure systems. The DSM approach presents a straightforward methodology in comparison to the SSIM. An increase in the number of variables to a problem or issue increases the complexity of the ISM methodology [63]. The DSM will be used in the study to highlight the infrastructure dependencies in railway infrastructure environments.

elements	8	7	6	5	4	3	2
1	O	A	A	O	A	A	A
2	O	O	A	O	X	A	
3	V	V	V	V	X		
4	V	X	X	V			
5	O	O	O				
6	O	A					
7	O						

Figure 3-5 : An example of a Structural Self-interaction Matrix (SSIM)

3.5 Dependability analysis

The principal stages that are distinguishable in any dependability analysis when developing a model are summarised in Figure 3-6 as functional, qualitative, quantitative and validation criteria. The functional and technical analysis involve collecting data, defining technical characteristics and the main functions of a system together with the external limitations. A qualitative analysis defines the objectives of the dependability analysis and establishes the scope of study regarding the dependability attributes required from the analysis such as reliability, availability, maintainability, or safety. The resolution level which describes the level of components and the degree of required information must be specified and highlighted in the qualitative stage for the system under analysis. The primary objective of a qualitative analysis is to establish all the failure mechanisms and failure combinations which affect the dependability of a system. The events that are likely to occur in the system and its environment such as failures and faults of system components become the elements of the reliability model. As a result information on the failure modes, their causes, and related dependability data must be made available to enable the

presentation of failures and faults (along with their combinations) of the components of the system which are detrimental to one of the dependability attributes (reliability).

Functional analysis	Qualitative analysis	Quantitative analysis	Validation criteria and conclusions
• Data colection. • Functional and technical characteristics.	• Objectives of the dependabilty study. • Qualitative analysis methods. • Resolution levels • Modelling of system dependability.	• Dependability measure • Sensitivity studies • Validation	• Identifying critical components • validate results

Figure 3-6 : Dependability procedures

Quantitative analysis is concerned with characterising the system dependability with measures such as probability. The probabilities can be obtained from mathematical statistical modelling which utilises probability failure distributions derived from information collected during elementary events within the system. A quantitative analysis identifies the strong and weak points of the system, the critical components, and the level of dependability that the system carries. Information of a quantitative nature apart from dependability data includes operating time, characteristics of preventative and corrective maintenance, and the statistical data about severe environmental conditions. There is some degree of uncertainty that comes with collecting failure data of a system. Validating the developed model integrates the outcomes of the quantitative and qualitative analysis. This process will draw conclusions and establish the failures and the combinations that influence the dependability of the system as well as identifying the most critical components and the most important functions of a system.

3.6 Section summary

This section presented a systems approach to modelling railway infrastructure systems. The modelling paradigms available to model railway infrastructure systems were presented and methods to model the dependencies in infrastructure systems were provided. Moreover, a general approach to performing a dependability analysis for the reliability of infrastructure systems was highlighted.

4 Reliability theory

To develop a substantive reliability model requires a study of reliability theory and the different modelling methodologies that can be employed in modelling railway infrastructure systems. This section presents the concepts involved in reliability modelling and the methodologies to study the failure processes in railway infrastructure systems. Repairable systems theory applicable to railway infrastructure systems is presented together with the appropriate statistical theory required to develop reliability models for railway infrastructure systems.

4.1 Reliability engineering

Reliability engineering has evolved to be an integral part of engineering and engineering design as it involves techniques and procedures that analyse the performance of systems and the underlying causes of system failure [64]. To achieve high levels of reliability in railway infrastructure it is important to balance between reliability, availability and cost-effectiveness [12]. The need to balance these attributes has seen a widespread application of reliability evaluations in performance measurement. Generally, reliability engineering has been used in several applications such as maintenance improvements, life cycle cost analysis (LCC), capital equipment replacement and economic evaluation analysis. This presents divergent definitions of reliability depending on the context in which it is applied. Fundamentally, reliability is used as a measure of a system's success in providing its function properly throughout its design life. Elsayed [65] defined reliability as the probability that a product will operate or provide a service properly for a specified period of time. Similarly, Modarres et al [66] described reliability as an item's ability to successfully perform an intended function. The prediction of failures is inherently a probabilistic problem; accordingly, in engineering analysis, reliability evaluation is thus a probabilistic process. Lewis [67] supported this by defining reliability as the probability that a system will perform its intended function for a specified period of time under a given set of conditions. What emerges from this definition as expressed by Lewis [67] and Conradie [25] is that a strict definition of reliability accounts for four distinct aspects which are *probability*, *function*, *time* and *operating conditions.*

The goal in a reliability analysis is to obtain an understanding of the system's likely behaviour by calculating the different performance measures. The performance measures are often presented as indices to aggregate information on the frequency of failure scenarios and their respective consequences. Quantitative reliability assessments emphasise the importance of estimating probabilities of failures. The probabilities can be used as a measure to estimate the effect of a component's performance towards a system's unreliability. Reliability systems analysis follows a stochastic approach where the objective is to obtain failure information for the entire system

33

based on the failure information of the systems components as shown in Figure 4-1. The quantitative assessments are then used to inform asset management decisions [68].
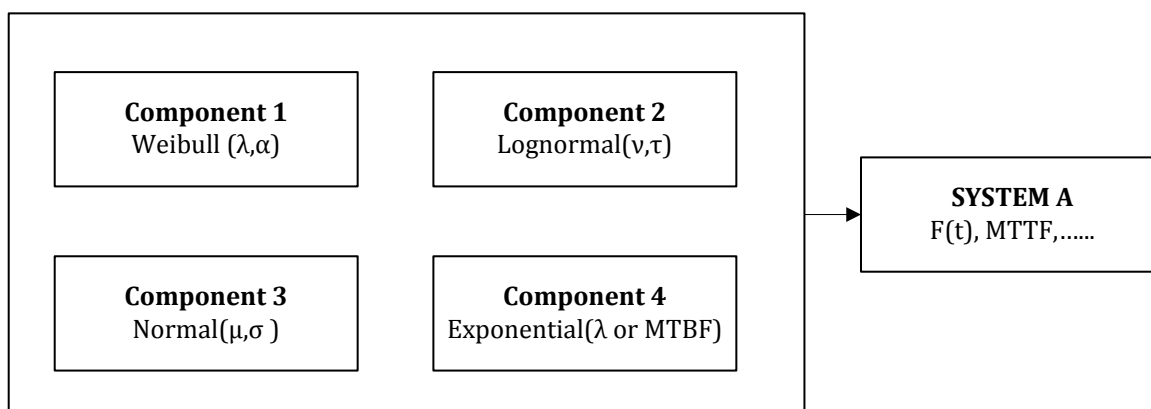
Figure 4-1 : Modelling component to system failure[50]

Reliability, when applied to infrastructure asset management, can be defined as a mathematical concept associated with dependability in which engineering knowledge is applied to identify and reduce the likelihood or frequency of failures within a system. Reliability is an attribute of dependability when performing a predictive analysis of a system. The end product of that process is a model related to the attributes of system dependability. To successfully apply reliability theory to railway infrastructure systems, a description of the expected functions of the system, the associated boundary conditions, failure frequency and the intervention and inspection strategies must be given [52]. Table 4-1 shows the typical guidelines to follow when performing reliability assessments.

Table 4-1 : Steps in a reliability assessment [69]

Step name	Description	Result
1. System configuration definition	Determine the basic functional blocks for the infrastructure system and dependencies among components	List of functional blocks, function , input , output, etc.
2. Data collection	Collection of necessary reliability and maintenance data	Reliability and maintenance data
3. Model building	Continuous time stochastic simulation model	Application of reliability modelling techniques
4. Simulation	Simulation scenarios and experiment design	Scenario listings and application of model
5. Results and analysis	Simulation results calculation	Results of parameters and reliability functions of interests

4.1.1 Reliability modelling

Infrastructure system failures occur because of individual asset failures. Railway infrastructure systems exhibit a high level of asset interdependence. This means that individual asset failure not only results in total system failure but rather triggers the failure of other assets within the same system (secondary failures). To develop a reliability model that captures all possible scenarios the subsystems, structures and activities that play a role in the initiation and propagation or arrest of failures must be identified and understood. This is achieved by utilising different levels of abstraction, a typical one being a high-level definition represented by a functional block diagram. A functional block diagram illustrates the operational, interrelationship and interdependence of the functional components of a system [66]. A hierarchical relationship which decomposes the system into subsystems and components can be logically derived from a functional block diagram with the process objective being the correct functioning of the system as shown in Figure 4-2. Functional hierarchies are developed from functional block diagrams by using deductive and systematic means.
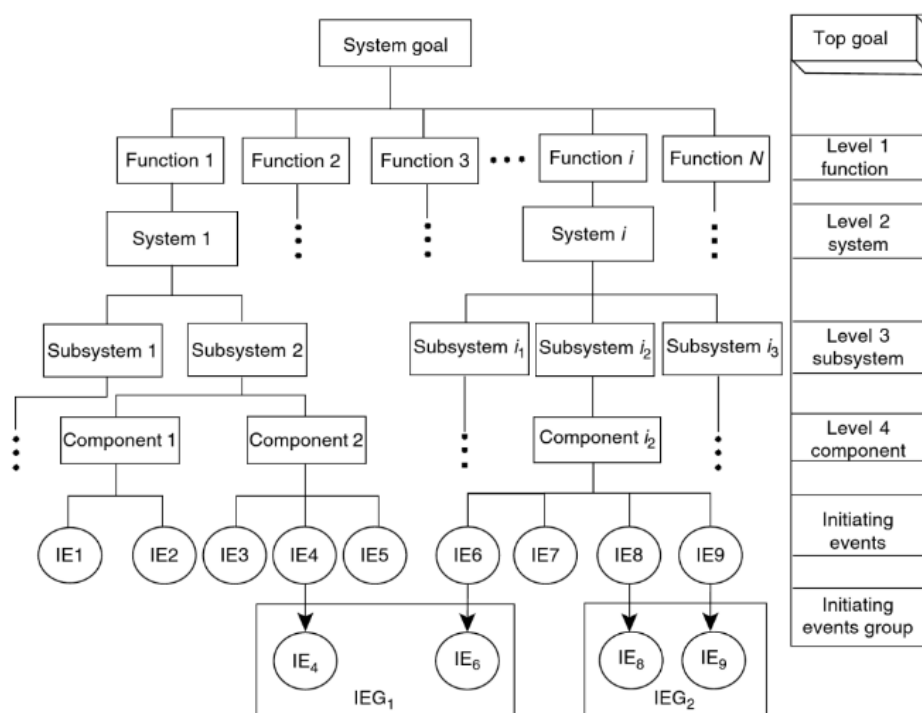
Figure 4-2 : Functional diagram (adapted from Risk Analysis in Engineering: 2006) [51]

Representing the functional relationship between individual assets in an infrastructure system enables the application of different techniques within the RAMS analysis framework that can be utilised to study failure effect and criticality in railway infrastructure systems [42]. Reliability block diagrams are among one of the simplest techniques to represent the logical configuration of a system. Reliability block diagrams are derived from functional diagrams and they enable a system to be seen as a function, which makes it possible to describe the system with a structure

function. A structure function is used to map the state of the components to that of the system. A basic characteristic of all functional systems is coherence. A system can be described to be coherent if all components that constitute it are relevant and if its structure function is monotone [57]. Two main classes exist that combine system components into a structure; a series structure and a parallel structure. Complex configurations use a combination of both series and parallel structures. A series structure only functions if and only if all *n* components in that configuration are functioning, whereas for a parallel structure, the system can function if one out of the *n* components is functioning [70]. The configuration of a series and a parallel system are shown in Figure 4-3.

Series system
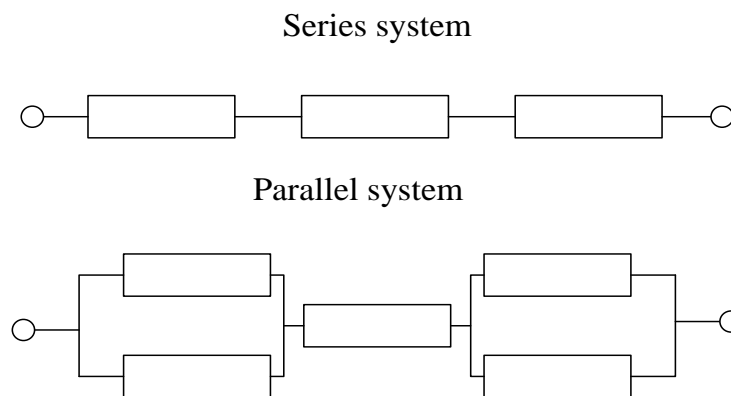
Parallel system

Figure 4-3 : Reliability block diagram showing the two main classes of configuring systems

The equations that are used to evaluate the system reliability of series and parallel configurations are given in equation 4.1 and 4.2 respectively

$$R_s(t) = R_1(t) \cdot R_2(t) ..... R_N(t) = \prod_{i=1}^{n} R_i(t) \tag{4.1}$$

$$R_s(t) = R_1(t) \cdot R_2(t) ..... R_N(t) = 1 - \prod_{i=1}^{n} R_i(t) \tag{4.2}$$

At the heart of any prediction, the problem is to select a suitable model structure. A model structure is a parameterised family of candidate models of some sort, within which the search for a model is conducted. A basic rule in estimation is not to estimate what you already know. In other words, one should utilise prior knowledge and physical insight about the system when selecting the model structure [71]. The decision as to whether to take the black-box or white-box approach is determined by the correct use of reliability engineering theory. Valenzuela [57] highlighted a white-box versus black-box dichotomy where the distinction is based on whether the failure process of a system is modelled with or without the explicit recognition of individual components that comprise the system. A component refers to the elementary building block of a white-box

system model. These correspond to the lower level entities if the models are developed hierarchically. Black-box models are constructed by correlating input measurables with output observables where parameters of various models are estimated. In reliability modelling, the primary goal is the most accurate replication of data, which makes a black-box modelling approach useful.

A model structure was presented by Rama and Andrews [2] in developing a holistic approach to infrastructure asset management. The model structure utilised a modelling approach that supported a multi-asset system by developing a framework to support informed decision-making in railway infrastructure asset management. Figure 4-4 shows a generic framework for modelling infrastructure life cycle costs (LCCs) railway infrastructure assets with two elements, the infrastructure state model, and the cost model. Using the infrastructure state model and the cost model, performance parameters can be estimated by studying the effects of changes in individual assets and how those changes are cascaded to the rest of the infrastructure system.
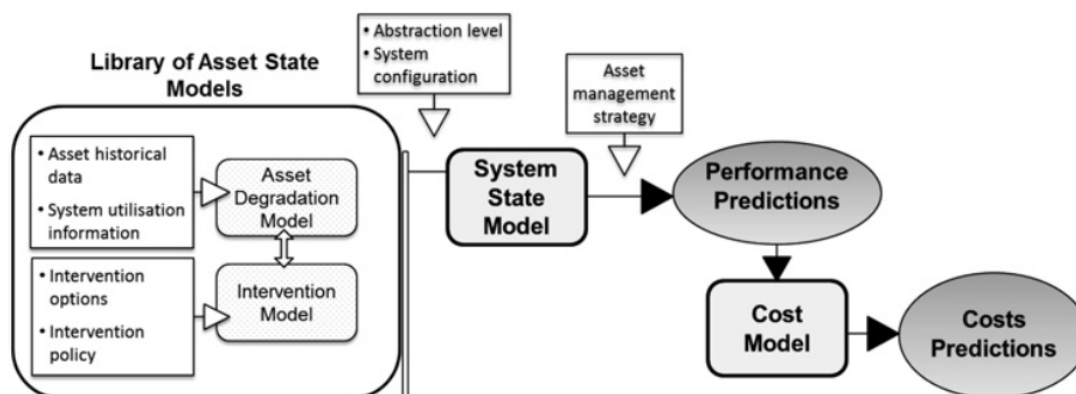
Figure 4-4 : Framework for decision support in infrastructure asset management[2]

In a similar approach Macchi [5] et al applied a reliability-based approach to maintenance improvement by proposing a family-based approach that identifies and groups items into families with the same reliability targets. Starting from this documentation, a railway system model is built by understanding the reliability logics as a result of interpretation of the trains flowing through the system. Using the reliability block diagram logics at each infrastructure indenture level as shown in Figure 4-5 the railway system model is built using generic operational states. The three generic states are a normal operating state, degraded state and downtime state.

MODELLING LEVEL	ENTITIES CONTAINED IN THE MODEL	RELIABILITY LOGICS
RAILWAY LINE MODELLING	• Hub stations • Tracks and stations along the line • Alternative transport routes	• Series logic • MSS logic
RAILWAY TRACK MODELLING	• Run tracks (even and odd tracks) • Run and non-run tracks (in case of stations) • Families of railway items (connecting adjacent tracks)	• Series logic • MSS logic • Parallel logic
RAILWAY ITEM MODELLING	• Families of real railway items (placed within the track	• Series logic

Figure 4-5 : Family-based approach to modelling reliability[5]

The reliability modelling approaches that have been presented prove that several analytical methods can be applied to evaluate the reliability of the railway infrastructure systems. Holistic models that have been presented accounted for the functional and operational characteristics of the infrastructure assets. These models, however, do not consider the common role of humans who execute the different processes required for effective asset management. Felice and Petrillo [72] proposed a methodological approach to improving railway transportation systems' reliability based on FMECA and human reliability analysis (HRA). This integrated approach seeks to consider the inherent complexity of human influence in improving system reliability. HRA provides a comprehensive logical analysis of factors influencing human performance, which enables recommendations for system improvement and prioritises attention on critical tasks that may jeopardise system reliability.

4.2 Failure processes

The process that describes how a multi-component system goes from operating state to a failed state or degraded state is known as the failure process. This process is a result of forces and stresses generated during the operation of systems or from external sources. A failure process is characterised by the structure of a system and the failure modes of its components. Failure is the termination of the ability of an entity to perform a required function. As a result, failures have different effects on the operation of a system and the failure effects need to be assessed to determine the impact on system performance. A scale of criticality can be used to classify failures with respect to their effects on the system. An example applicable to railway systems is shown in Table 4-2. Alternatively, failures can be classified according to their causes, which can be due to primary or secondary causes. Primary failures are not caused directly or indirectly by the failure of another component within the system. On the other hand, secondary failures are directly and indirectly caused by the failure of another component within a system.

Table 4-2 : Failure categorisation

Failure Category	Consequence
Significant	Cancellations
Major	Delays
Minor	Reduction in capacity

Failures in a railway network occur in different parts of the network and may only be studied together within comparable parameters. In that case when failures are recorded the criteria on the infrastructure and impact on traffic must be provided [12]. Esveld [73] suggested that failure data should be grouped into comparable sets by presenting guidelines on the process of recording failures. Furthermore, when collecting failure data it is important to highlight each failure mode separately. A failure mode is an effect by which a failure is observed. There is a difference between failure causes and failure modes. Failure causes of a component are failures of that part whereas failure modes are the tangible effects that these failures produce on the functions of the asset. More significantly it must be noted that failure modes have a direct impact on system reliability in terms of the probability of occurrence of the failure modes. Additionally, failure modes depend on the response time to restore a system into safe mode and the maintenance support for effective and safe maintenance procedures.

When analysing system reliability, particularly that of railway infrastructure which has a complex configuration, it is required to critically ascertain the root cause of infrastructure failures and their effects in order to understand the nature and occurrence of system failures. Studying railway infrastructure failure modes assists in assessing the impact of infrastructure defects on the performance of the network. McNaught [14], Jidayi [24] and Brinkman [47] identified and categorised critical railway perway failure modes. The failure modes identified that have secondary effects on the infrastructure system include rail breaks, faulty block joints, and pantograph hook-ups. Hassankiade [74] performed a failure analysis of railway switches and crosses and identified the critical failure modes in railway signalling infrastructure based on historical data and failure frequency. Saba [50] presented a hazard log list showing the different failure interfaces between the electrical, signalling and perway railway infrastructure subsystems. Patra and Kumar[75] also performed an availability analysis on a railway track circuit and highlighted rail breaks and rail joint failure as one of the most critical failure modes.

The study of failure processes of complex systems can be defined either as failure-based reliability approach or as degradation-based approach. The random variable of interest in a reliability-based approach is the failure time of components while degradation-based models are interested in the remaining useful life of components [57]. A failure-based reliability approach will be the focus of

this study. Figure 4-6 shows a typical process to be followed when performing a failure-based reliability study. It can be seen that the first step to a successful reliability evaluation is establishing the system characteristics and related failure modes.

• Failure data collection • Failure modes • Infrastructure characteristics	• Parametric • Non-parametric analysis	Calculation of failure rate expression λ(t)	Calculation of reliability R(t)
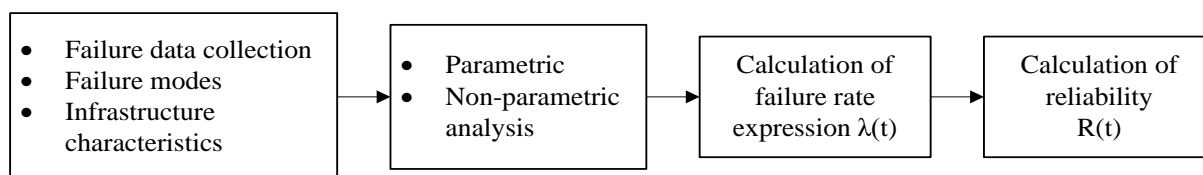
Figure 4-6 : Reliability and failure rate forecasting procedure (adapted from Pereira [12])

4.2.1 Failure Mode Effect Analysis (FMEA)

Failure Modes and Effects Analysis is a reliability assessment technique developed for the USA defence industry but it has been extended in practice to be used in different areas of system failure analysis. The FMEA is a systematic structure method that can be used to identify and assess the effect and/or consequences of failure modes on the infrastructure system. This approach utilises an inductive and experiential technique to provide qualitative information about a system's design and operation. FMEA operations have been used to create hierarchical lists of maintenance items and subsystems for improvement and modification. These hierarchical lists can be implemented to achieve the required infrastructure performance by applying the appropriate maintenance strategy. Figure 4-7 shows the iterative process of identifying the causes, effects, and modes of failure in a system.
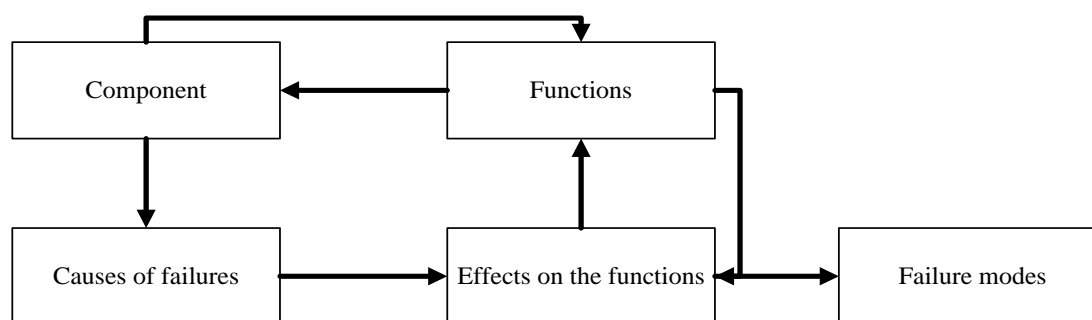
Figure 4-7 : Causes effects and modes of failure

FMEA can be extended to classify potential failure effects according to their severity and criticality to become FMECA (Failure Modes, Effect, and Criticality Analysis). FMECA documents the catastrophic and critical failures in a system. Identifying these critical and catastrophic failures implies that the criticality of the consequence and severity of the failure in a system can be established. The fundamental objective of a criticality assessment is to determine the failure modes on the basis of their consequence and the probability of occurrence. Using the FMECA, the successful assessment of asset criticality is achieved by utilising two common methods which are the Risk Priority Number (RPN) technique and the Military standard technique (MIL-STD-1629). The RPN technique calculates the risk priority number which is based on the probability of the

failure occurrence (O_r), the severity of its effects (S_r) and the detectability (D_r) of the failure [66]. Failures that score high RPN values are areas of greatest risk requiring their causes to be minimised.

$$RPN = O_r \times S_r \times D_r \qquad\qquad [4.3]$$

The military standard technique (MIL-STD-1629) categorises and prioritises failure modes according to severity so that the appropriate interventions can be recommended and it looks at two types of criticality analysis; qualitative and quantitative. Qualitative criticality analysis looks at the severity of the potential effects of failure and the likelihood of occurrence for each potential failure mode. A criticality matrix is developed to identify and compare each failure mode with all other failure modes with respect to severity [76]. Quantitative criticality analysis considers the reliability or unreliability of system components at a given operating time and identifies the portion of the component's reliability that can be attributed to each potential failure mode.

4.2.1.1 Application of FMECA to railway infrastructure

Famurewa [77] utilised FMECA to support an analysis to increase railway infrastructure capacity through improved maintenance management practices. Brinkman [47] utilised FMECA to model failure behaviour and to measure the effects of maintenance concepts using a simulation process that expressed results in terms of the performance indicators for railway infrastructure. McNaught [14] recommended FMECA in the development of a risk-reliability model for the perway subsystem because of the comprehensive results it provides over other methods. The FMEA and FMECA are preliminary analysis methods that can be complemented by other methods to identify the combinations of relevant failures. Jidayi [24], Carretero et al [3] and Network Rail [26] utilised FMECA and Pareto methodologies in evaluating the risk and reliability of railway infrastructure networks. The Pareto analysis is a statistical technique in decision-making used for selecting a limited number of tasks that produce a significant overall effect. The technique uses a Pareto principle also known as the 80/20 rule, which is useful in a case where many possible courses of action are competing for attention. The Pareto principle states that 'in any series of elements to be controlled, a selected small factor in terms of the number of elements almost always accounts for the large factor in terms of effort' [78]. The Pareto analysis is a creative way of identifying the cause of problems, but it is limited by the fact that it excludes possible important problems which may seem small at first but grow with time.

Saba [50] utilised FMECA to develop a RAMS program for railway infrastructure identifying failure modes and potential hazards within the infrastructure system. To identify the potential hazards, two common methods were found in literature which are the preliminary hazard analysis (PHA) and the Hazard and Operability analysis, which place priority on hazards and not on failure [41]. Preliminary hazard analysis (PHA) utilises pre-existing experience or knowledge of a hazard or

41

failure to identify potential hazards and events that might cause harm. On the other hand, the Hazard and Operability Study (HAZOP) is a rigorous analysis method that utilises guide words to identify potential deviations from a system's normal operating conditions. The guide words utilised describe functional losses at system and subsystem level. PHA and HAZOP are more useful when applied to safety analysis than to reliability evaluations, but they can apply in the initial stages of reliability studies to understand failure modes and unwanted events that led to those failures.

Fault Tree Analysis has been extensively used to evaluate the reliability, assess the failure effects, and investigate the impact of maintenance practices on railway electrical systems [19], [20], [79], [80]. Fault Tree Analysis (FTA) is a diagnostic tool used to predict the most likely failure to cause system breakdown. In a systematic way, the combination(s) of conditions required for an event to occur are delineated by identifying how failure-related events at the higher level are caused by events at the lower level, known as 'primary events'. The results from an FMEA analysis can be used as an input for performing FTA methods. However, when Fault Tree Analysis is compared with FMEA/FMECA, it can be seen that an FTA predicts the causes for usually known problems. In contrast, FMEA/FMECA methods systematically predict new problems and their causes. In other words, the FTA identifies part failure as a cause of functional failure whereas FMEA/FMECA identify functional failure as a result of part failure. For all the above-mentioned techniques, it is worth noting that the best performance of the methodologies is achieved when the techniques are used properly for a particular requirement at a specific stage within the framework of modelling and quantifying railway infrastructure reliability.

4.2.2 Modelling failure characteristics

We may analyse the reliability of a system in terms of the component or mode failures, provided they are independent of one another. For each mode, we may define a probability density function for a time to failure and an associated failure rate. The important point in all this is that the definition of the failure modes totally determines the system's reliability and dictates the failure mode data required at the component level [67]. Reliability is best understood in term of rates of failure; time then becomes an important variable in reliability studies. To gain a thorough insight into the nature of failures, one needs to examine the time dependence of failures throughout the design life of infrastructure systems. This will differentiate failures caused by the different system mechanisms from those caused by the different components of a system. The failure rate or hazard rate is thus an important function in reliability analysis because it shows the changes in the probability of failure of a component over its design life.

Generally, a failure rate function exhibits a bathtub shape often referred to as the bathtub curve shown in Figure 4-8. A bathtub curve displays three distinct phases in a component's life cycle as it is a superposition of three different failure distributions. The curve in the early failure region,

also known as 'infant mortality region', exhibits a decreasing failure rate which can be attributed to design defects or the period of adjustment for interacting components in a system. The constant failure rate region referred to as the 'useful life' is a period in the life cycle characterised by random failures of the component likely caused by random events resulting from external factors and other unavoidable loads. The 'wear out' region in contrast to the early lifetime region exhibits an increasing failure rate characterised mainly by complex ageing and degradation processes.
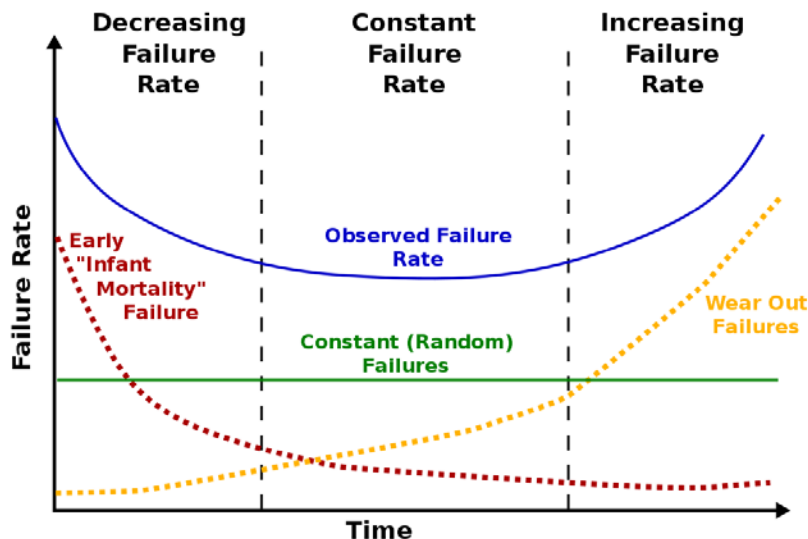
Figure 4-8 : Bathtub curve for failure studies

Not all components exhibit the bathtub-shaped failure rate curve. Mechanical components do not show a constant failure rate region but rather exhibit a gradual transition between the early failure rate and wear out stages [65]. Electrical devices exhibit a relatively constant failure rate distribution. The distributions in the wear out curve are believed to be the dominant failure distributions in most components. Failure rates grow with the load for railway infrastructure components. Jorge et al [12] in the study of the failure of railway infrastructure, recommended the use of a formula with non-constant failure rate. When working with variable failure rates it is of little value to consider the actual failure rate since only reliability and MTBF are meaningful. The non-constant failure rate is often used when working with reliability and MTBF directly because it does not require knowledge of the actual failure rate of the components. Performing an analytical calculation when dealing with non-constant failure rate will result in extremely complicated functions. As a result, several expressions and statistical models can be written and assigned to non-constant failure rate using empirical datasets.

43

4.2.3 Repairable systems theory

Railway infrastructure systems contain electrical and mechanical equipment, such as point machines, track circuits, and trip stops. This means these components usually exhibit varying deterioration and or improvement in the reliability performance over time, therefore, a constant failure rate will not always be sufficient or appropriate when performing a reliability evaluation of multicomponent systems.

Railway infrastructure systems are repairable systems. A repairable system is a collection of items, which after failing to perform at least one of its required functions, can be restored to performing all of its required functions by any method other than replacement of the entire system [81]. Non-repairable systems are discarded the first time they cease to perform a function satisfactorily. Upon failure, they cannot be repaired and are generally replaced. When working with repairable systems it is often preferred to count the events which influence the performance of a system. This approach assumes the event-driven modelling approach presented in section 3.3 where the events are either system failures or system repairs.

The Renewal Process (RP), Homogeneous Poisson Process (HPP), and Non-Homogeneous Poisson Process (NHPP) are the general stochastic processes employed in analysing the reliability of repairable systems. A stochastic point process is a mathematical model for a physical phenomenon characterised by highly localised events distributed randomly in a continuum [81]. RP methods analyse data on the assumption that the times between failures are independent and identically distributed in the time domain. This assumption makes the RP appropriate for non-repairable systems. In scenarios where the RP is applied to repairable systems the assumption that the repair returns the system to 'as good as new' is taken [82]. When the HPP and NHPP are applied to repairable systems the continuum is the time and the highly localised events are failures or repairs which occur at instants within the time continuum. Figure 4-9 represents a portion of a sample path of a stochastic point process representing successive failures of a single system. The failure rate of the process is the instantaneous rate of change of the expected number of failures with respect to time, which means it is a failure rate of the process that measures wear-out of the system.
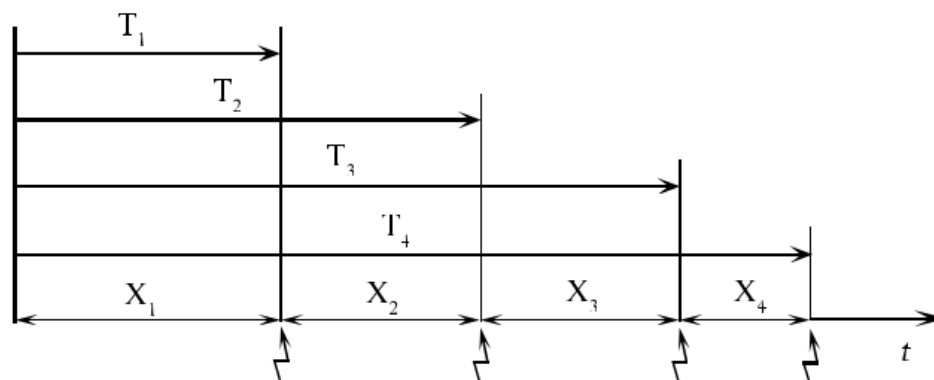
Figure 4-9 : Stochastic process

When dealing with reliability evaluations for repairable systems Basile et al [81] posed two assumptions: 1) the system will be operated wherever possible; and 2) repair times are negligible. Reliability evaluations of repairable systems study the process of failures and repairs of a system. Typically times between failures will be neither independent nor identically distributed. As a result, in reliability analysis, the time is measured in terms of the operating time between failures ignoring repair times [82]. O' Connor [83] supports this when recommending the use of time-based failure distributions stating that replacement or repair times are usually small as compared with standby or operating times hence is it is feasible to assume that the failure of the component is independent of its repair actions.

4.2.3.1 Non-homogeneous Poisson Process

The distinction between HPP and NHPP is that the rate of occurrence of failures (ROCOF) for the NHPP varies with time and is not constant as in the case of HPP [25]. The NHPP process describes a sequence of random variables which are neither independently nor identically distributed. For NHPP models the rate of occurrence of failures varies with time. An NHPP is more applicable and can be easily used for modelling data that exhibits a trend [25]. When failure data is ordered chronologically and a trend is observed, the interpretation is that the time to failure is not independent or identically distributed (IID). If ordered by magnitude, however, which implies IID, misleading results will be produced because once failure data is reordered, the trend information is lost [83]. The NHPP is used to model repairable systems that are subjected to a minimal repair strategy with negligible repair times. The implications of minimal repair mean that when a system fails and the system is restored to the functioning state, the likelihood of system failure is the same before and after a failure repair. This assumption draws more attention to the NHPP because most repairs involve the replacement of only a small fraction of a system's constituent parts. It is, therefore, plausible to assume that the system's reliability is the same as it was just before the failure occurred. When an NHPP model is used to model a repairable system, the system is treated as a black box in that there's no concern about the internal system of the components [39]. There

are two functional parametric NHPP models that have been highlighted in literature [25] [14][81]; the log-linear model and the power law model. When dealing with repairable systems the focus is on predicting the probability of system failure, the expected number of failures, the probability structure of time between failures and the probability structure of the time to failure as a function of system age [82]. The equations related to the NHPP for the power law and log-linear law to determine these parameters are given as follows.

4.2.3.1.1.1 Power law NHPP

The power law model ROCOF NHPP is given by

$$\rho_2(t) = \lambda \beta^{\beta-1} \; where \; \lambda, \beta > 0, t \geq 0 \qquad [4.4]$$

Expected number of failures

$$E_p\left(N(T_2) - N(T_1)\right) = \lambda\left(T_2{}^{\beta} - T{}^{\beta}{}_1\right) \qquad [4.5]$$

Reliability

$$R(T_1, T_2) = e^{-\lambda\left(T_2{}^{\beta} - T_1{}^{\beta}\right)} \qquad [4.6]$$

Mean time between failures

$$MTBF_2(T_1, T_2) = \frac{T_2 - T_1}{\lambda\left(T_2{}^{\beta} - T_1{}^{\beta}\right)} \qquad [4.7]$$

4.2.3.1.1.2 Log-linear law NHPP

The log-linear law model ROCOF NHPP is given by

$$\rho_1(t) = e^{\alpha_0 + \alpha_1 t}, with -\infty < \alpha_0, \alpha_1 < \infty, t \geq 0 \qquad [4.8]$$

Expected number of failures

$$E_{\log}\left(N(T_2) - N(T_1)\right) = \frac{1}{\alpha_1}\left(e^{\alpha_0 + \alpha_1 T_2} - e^{\alpha_0 + \alpha_1 T_1}\right) \qquad [4.9]$$

Reliability

$$R(T_1, T_2) = e^{\frac{-\left(e^{\alpha_0 + \alpha_1 T_2} - e^{\alpha_0 + \alpha_1 T_1}\right)}{\alpha_1}} \qquad [4.10]$$

Mean time between failures

$$MTBF_{\log}\left(T_1, T_2\right) = e\frac{\alpha_1\left(T_2 - T_1\right)}{e^{\alpha_0 + \alpha_1 T_2} - e^{\alpha_0 + \alpha_1 T_1}} \qquad [4.11]$$

4.2.3.2 Homogeneous Poisson Process

An HPP describes the sequence of independently and identically exponentially distributed random variables. For the HPP the rate of occurrence of failures does not vary with time. Despite its simplicity, the HPP model is used widely for repairable systems. Classical statistical distributions such a lognormal, exponential and Weibull can be used for modelling HPP models for repairable systems. The HPP is applicable in scenarios where there is no evidence of a trend or dependence in the failure data. Widely used functions in reliability engineering include failure rate, mean time function and the reliability functions. The functions can be derived from the PDF (probability density function) of the statistical distributions used to model the HPP. Commonly used distributions to represent life data include the exponential, lognormal and Weibull distribution and will be discussed as follows.

4.2.3.2.1 Exponential distribution

The exponential distribution is commonly used to model constant failure rate models. Meeker [84] states that the exponential distribution is appropriate for some electrical components and can describe failure times for components that exhibit physical wear-out. Furthermore, it is suitable for modelling the time between system failures but is highly inappropriate for modelling the life of mechanical components which are subjected to a combination of fatigue, wear, or corrosion. The two-parameter exponential distribution has a CDF, PDF, hazard function and reliability function given as below. Ө is a scale parameter and must be greater than zero, γ is a location and threshold parameter. If γ= 0 the exponential distribution becomes the well-known one parameter exponential distribution. In special circumstances, the exponential distribution can be useful in determining the time between system failures and other inter-arrival time distributions [84].

$$F\left(t, \theta, \gamma\right) = 1 - e^{-\left(\frac{t-\gamma}{\theta}\right)} \qquad [4.12]$$

$$f\left(t, \theta, \gamma\right) = \frac{1}{\theta}e^{-\left(\frac{t-\gamma}{\theta}\right)} \qquad [4.13]$$

$$h\left(t, \theta, \gamma\right) = \frac{1}{\theta}, t > \gamma \qquad [4.14]$$

$$R\left(t\right) = 1 - F\left(t\right) = e^{-\left(\frac{t-\gamma}{\theta}\right)} \qquad [4.15]$$

4.2.3.2.2 Lognormal distribution

The lognormal distribution represents the distribution of a random variable whose logarithm follows a normal distribution. This distribution model is particularly useful for modelling failure processes that are a result of many multiplicative errors. Meeker [84] highlighted that the model is appropriate to model time to failure that is caused by a degradation process involving combinations of random rate constants that combine multiplicatively. Some specific applications of a lognormal distribution are modelling time to failure of components due to fatigue cracks and failures attributed to maintenance activities [66]. The lognormal distribution has been widely used to describe the time to fracture from fatigue growth in metals and has been used to model electronic components that exhibit a decreasing failure rate. The CDF and PDF of lognormal distributions are given as follows:

$$F(t,\mu,\sigma) = \Phi_{nor}\left[\frac{\log(t)-\mu}{\sigma}\right] \qquad [4.16]$$

$$f(t,\mu,\sigma) = \frac{1}{\sigma t}\phi_{nor}\left[\frac{\log(t)-\mu}{\sigma}\right], \quad t > 0 \quad [4.17]$$

4.2.3.2.3 Weibull distribution

The Weibull distribution has a broad range of applications in reliability analysis mainly because of its flexibility in describing all three regions of the bathtub curve. Todinov [85] describes the Weibull model as a universal model for the times to failure of structural components of systems which fail when the weakest component in the system fails. Modarres [66] showed that it is possible to use a Weibull distribution for a system composed of a number of parts whose failure is governed by the most severe defect of its components, known as the weakest link model. Antoni [86] simulated different ageing scenarios using the Weibull lifetime model to investigate the impact of different maintenance strategies for the railway signalling equipment. Meeker [84] further recommended the use of the Weibull distribution to model failure time with decreasing or increasing hazard functions. In general, the Weibull case requires three parameters. They do not have a physical meaning in the same way that failure rate does. They are parameters which allow us to compute reliability and MTBF. These parameters are the shaping parameter, scaling parameter and the location parameter. The Weibull CDF, PDF, hazard function, and reliability function can be written as:

$$F(t,\mu,\sigma) = \Phi_{sev}\left[\frac{\log(t)-\mu}{\sigma}\right] \qquad [4.18]$$

$$f(t,\beta,\eta) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}e^{\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]} \qquad [4.19]$$

48

$$h\left(t,\mu,\sigma\right)=\frac{1}{\sigma e^{\mu}}\left[\frac{t}{e^{\mu}}\right]^{\frac{1}{\sigma-1}}=\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\ t>0 \qquad\qquad [4.20]$$

$$R\left(t\right)=e^{-\left(\frac{t}{\eta}\right)^{\beta}} \qquad\qquad\qquad\qquad\qquad [4.21]$$

The equations presented for time to failure distributions need to fit the failure data. Ahmad et al [87] presented a new approach to failure distribution fitting and established that the application of incorrect failure distribution in maintenance optimisation studies will yield inaccurate results. Maillart and Pollock [88] in their study of the effect of failure distribution specification errors found that if the failure distribution is incorrectly specified, the cost per unit time will significantly increase in the long run. Preventative maintenance strategies are more effective in cases where the failure rate increases with time. If a preventative maintenance strategy is carried out at decreasing or constant failure rate, the replacement and downtime costs will increase significantly by time. As a result, it is important to employ the correct failure distributions. This is achieved by utilising statistical methods that will be the subject of the next section.

4.3 Statistical methods for reliability evaluations

Taking up the question of statistics, given a set of data, how do we infer the properties of the underlying distribution from which the data has been drawn? At this point distinguishing between the statistical analysis of a component and the analysis of system failure data is important. Components have distributions with a single time to failure whereas time between successive failures of a system are modelled by a sequence of distribution functions. Therefore the failure of a single system is sufficient for the statistical analysis if there is enough observed inter-arrival times for time to failure distribution approximation. The railway infrastructure system contains several components. The statistical failure approximation of an infrastructure system can, therefore, be modelled by multiple failures from different parts of the infrastructure system. The system approach is less data intensive and will thus be the focus of further investigation.

When using observed failure data to select and estimate failure distribution models to perform a reliability evaluation there are non-parametric and parametric methods that can be utilised for this exercise. Empirical methods provide a non-parametric graphical estimate of the failure rate versus the asset age or rate of asset utilisation. Furthermore, empirical methods do not assume the form of the mean function or the process of generating system histories. Parametric methods, on the contrary, use probability distributions like the Weibull or exponential distributions to model the failure behaviour of the system components. Meeker [84] recommended that data analysis should begin with empirical techniques which do not require assumptions in assigning

49

models. Therefore empirical analysis can be interpreted as an intermediate step towards a more complex model. Lewis [67] supported this by stating that empirical analysis can provide insight toward selection of the most appropriate time to failure distribution. The use of parametric methods can complement empirical methods precisely because parametric models provide smooth estimates of failure time distributions and can be described accurately with just a few parameters, unlike empirical methods which have to report an entire curve.

To determine which failure distribution to assign in the reliability evaluations, three stages are usually employed when analysing statistical data. The stages which enable the development of a probabilistic model of a system are trend testing, parameter estimation and selection of the best fit for the appropriate point process model [14]. The data analysis for the reliability modelling of repairable systems can follow a basic methodology as presented in Figure 4-10 . The flow chart presents criteria for model identification and can be used as a basis for the analysing of failure data.
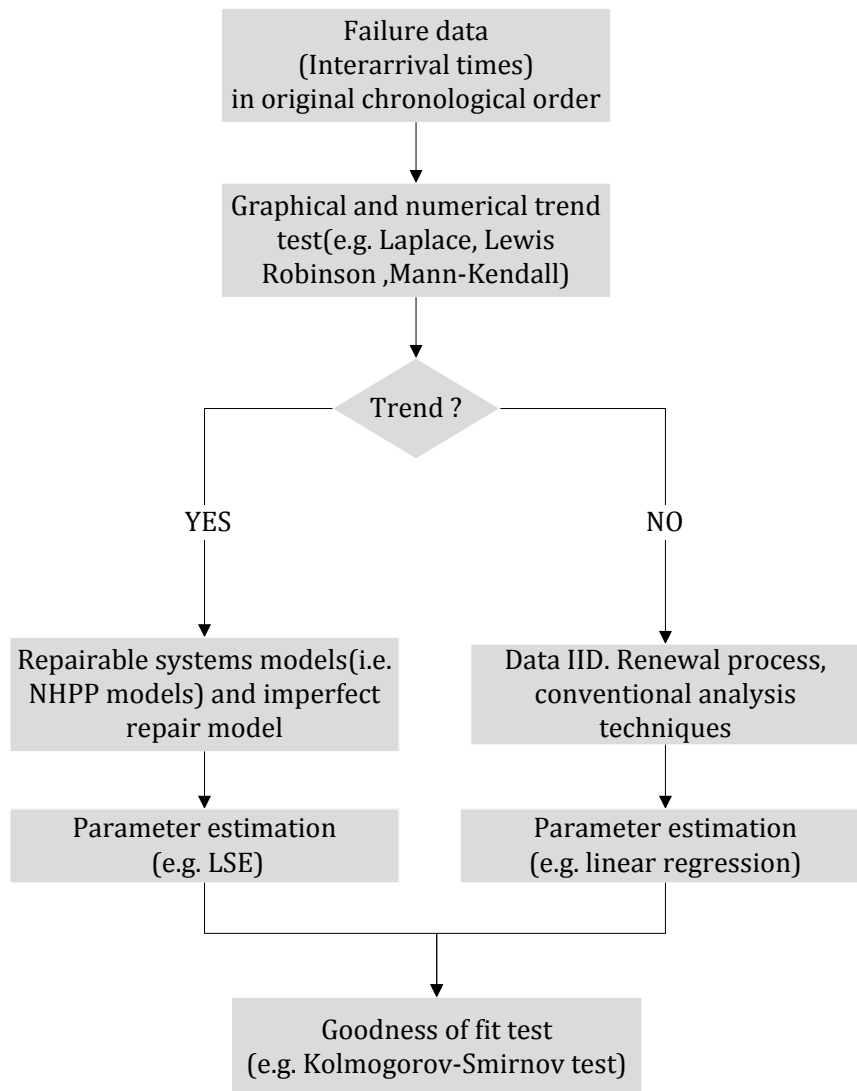
Figure 4-10 : Framework for analysis of failure data for reliability evaluations

4.3.1.1 Trend testing

To identify which point process model to apply to available failure data, trend testing is employed. A graphical assessment of observed failure data is not sufficient, hence a numerical validation is required to confirm the graphical assessment results and to establish if the data observed is statistically significant or just accidental. The main objective of a trend test is to identify if failure patterns are significantly changing with time. In a pattern of failures, the trend can be either monotonic or non-monotonic. A monotonic trend has a concave or convex shape whereas non-monotonic trends occur when trends change with time or when trends repeat themselves in cycles [89]. An example of a non-monotonic trend as discussed is the bathtub curve. A trend test is conducted by testing a null hypothesis that a system failure pattern is a point process. If interoccurrence times are independent and identically distributed (IID) it implies an HPP, otherwise, the alternative hypothesis is adopted implying an NHPP. There are several methods to perform a trend test. This study will describe the frequently used tests which are the Laplace test, The Military Handbook Test (MLK-HDBK-189) and the Lewis Robinson Trend Test.

4.3.1.1.1 The Laplace test

This is the most used trend test for data sets. The test statistic where the system is observed until n failures have occurred where S_1, S_2 denote the failure times.

$$U = \frac{\frac{1}{n-1}\sum_{j=1}^{n-1}S_j - \left(S_n/2\right)}{S_n/\sqrt{12(n-1)}} \qquad [4.22]$$

Where the system is observed until a time t_0, the test statistic is given

$$U = \frac{\frac{1}{n}\sum_{j=1}^{n}S_j - \left(t_0/2\right)}{t_0/\sqrt{12n}} \qquad [4.23]$$

In the both cases, the test statistic U is approximately standard normally distributed when the null hypothesis H_0 is true. The numerical value of U will indicate the direction of the trend with U <0 for a happy system and U >0 for a sad system. Table 4-3 shows the different interpretations of the Laplace Trend Test values U. The rejection criteria is based on the assumption that U follows a standard normal distribution. Conradie [25] and Lindqvist [90] advised that the use of the Laplace Trend Test (LTT) should not be done without questioning the data and the results. For Laplace Trend Test values within the grey area as highlighted in Table 4-3, further tests are required such as the Lewis- Robinson test, Mann- Kendall Test and the Weibull test.

Table 4-3: Interpretation of the LTT value U [25]

Value of u	Description	Type of theory
$-2 < U_L < -1,$ $1 < U_L < 2$	Grey area, more tests required	Either renewal theory or repairable systems theory
$U_L < -2$	Reliability improvement, data non-homogeneous	Repairable system theory, use NHPP
$U_L > 2$	Reliability degradation, data non-homogeneous	Repairable system theory, use NHPP
$-1 < U_L < 1$	Non-committal, data homogeneous	Renewable theory, use HPP

4.3.1.1.2 The Military Handbook Test

The test statistic from the military handbook test for the case where the system is observed until *n* failure where to occur is given by:

$$Z = 2\sum_{i=1}^{n-1} In\frac{S_n}{S_i}$$

[4.24]

Where the system is observed until time t_0, the test statistic is given by:

$$Z = 2\sum_{i=1}^{n} In\frac{t_0}{S_i}$$

[4.25]

For the Military handbook test, the null hypothesis is 'HPP' which is rejected when the z values are small or large. Low values correspond to deteriorating systems, while the large values of Z correspond to improving systems. In strict terms, the rejection of the null hypothesis implies that the process is not HPP but in principle, it could still be a renewal process and thus still have no trend. These false rejections can be avoided by utilising the Mann test or the Lewis-Robinson test.

4.3.1.1.3 Lewis-Robinson Trend Test

When the null hypothesis is rejected with the Laplace Trend Test and Military Handbook trend tests it is important to avoid drawing the wrong conclusions. To counter this, the Lewis-Robinson Trend test is introduced to provide a modification to the Laplace trend test. In this instance, the null hypothesis is the distribution of the arrival times that correspond to a renewal process. The test statistic for the Lewis-Robinson Test is defined in terms of the Laplace test statistic and the coefficient of variation for the inter-arrival times.

$$U_{LR} = \frac{U_L}{CV}$$

[4.26]

4.3.1.2 Parameter estimation

Parameter estimation is a process that provides tools to use data for aiding in reliability modelling and estimation of constants appearing in the time to failure models [91]. When the suitable time to failure model is selected for a random variable of interest the variables that govern the characteristics of the particular distribution need to be determined. When estimating time to failure parameters it is important to consider confidence intervals in the process. In many cases, failure data is not always complete and thus the estimation process has a degree of uncertainty. A number of techniques are available to perform this process.

4.3.1.2.1 The Least Square Estimation Method

The Least Squares method produces estimated parameters with the highest probability of being correct if critical assumptions are observed. The estimation follows the statistical curve fitting approach of plotting a line that produces the smallest difference between the expected and observed values [14]. The basis of this method lies in minimising the sum of the squared errors ($e_1{}^2 + e_2{}^2 + e_3{}^2 + e_4{}^2$) as shown in Figure 4-11. Linear model parameters estimations can be determined analytically, but for non-linear models, an analytical solution becomes complex and very time-consuming. This can be avoided by transforming a non-linear model to a linear model but care should be taken when performing the transformation [92].
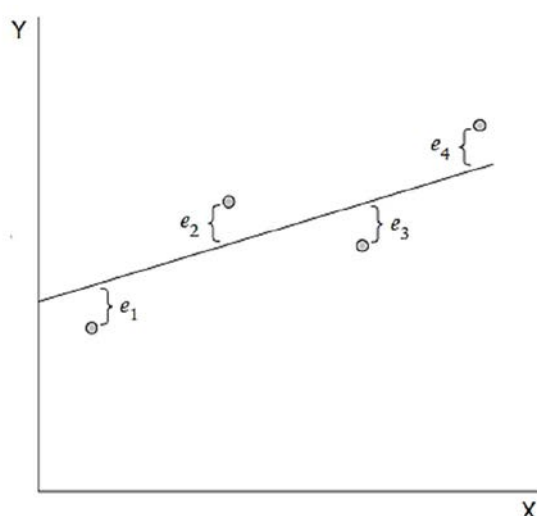
Figure 4-11 : Errors for the Least Square method

4.3.1.2.2 Maximum Likelihood Estimation

Likelihood – a basic measure of the quality of a set of predictions with respect to observed data [78]. Maximum Likelihood Estimator (MLE) is consistent in most cases, provides intuitive results, and is widely accepted as one of the most powerful methods for parameter estimation. MLE for multinomial distribution is unbiased but its variance is problematic when estimating parameters that calculate probabilities of events with low expected counts [93]. Suppose data consists of random observations $x_1,\ldots\ldots,x_n$ of a random variable coming from the same population with

53

probabilities governed by an unknown parameter Θ, the PDF for each of the *n* observations is given as:

$$P\left(X_i = x_i\right) = f\left(x_i \mid \theta\right), \; i = 1, \ldots . n \qquad\qquad [4.27]$$

These random observations are independent and as such the joint probability is the product of the PDFs for all the *n* observations and is called the likelihood function given as:

$$L = f\left(x_i \mid \theta\right) \ldots \ldots \ldots f\left(x_n \mid \theta\right) \qquad\qquad [4.28]$$

The concept behind the maximum likelihood function is maximising the natural logarithm L and solving for Θ from which the maximum likelihood estimate Θ is obtained [94]. This is achieved by taking the derivative of the natural logarithm of L (In L) with respect to Θ and equating it to zero as shown.

$$\frac{\partial In \; L\left(\theta; x\right)}{\partial \theta_i} = 0 \;\; for \, i = 1, 2 \ldots \ldots . m \qquad\qquad [4.29]$$

The maximum likelihood method is applicable for both part components and systems and as such the variable *x* can be replaced with time *t* [14].

4.3.1.3 Selection of best fit

To determine whether a sample of data belongs to the hypothesised theoretical distribution, a test to determine the adequacy of fit needs to be determined. These tests establish the level of confidence to which a specific distribution with known parameters fits a given set of data [67]. This test is done by establishing the difference between the frequency of occurrence of a random variable as seen from the observed sample and the expected frequencies obtained for the hypothesised distribution. These are known as the goodness-of-fit tests. There are several goodness-of-fit tests. Two commonly used methods will be discussed; the Chi-square and the Kolmogorov goodness-of-fit tests.

4.3.1.3.1 Chi-square tests

This test is based on a statistic that approximates the chi-square distribution. An observed sample taken from the population representing a random variable X must be split into *k* non-overlapping intervals. The hypothesised distribution model is then used to determine the probabilities p_i that the random variable X would fall into each interval *i (i=1,2,…,k)*. Multiplying the probability p_i by the size of the sample *n*, we get derive the expected frequency as e_i. The observed frequency for each of the intervals *i* is denoted by o_i, the difference between e_i and o_i characterises the adequacy of fit. The test statistic for the chi-square test is χ^2 which is defined as:

$$W = \chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \qquad [4.30]$$

From the equation of the statistic χ², if o_i differs significantly from e_i the value of W will be large implying that the fit is poor [95]. The chi-squared test performs poorly for small data samples.

4.3.1.3.2 Kolmogorov – Smirnov Goodness-of- Fit Test

The Kolmogorov-Smirnov (K-S) test is a commonly used goodness-of-fit test based on cumulatively ranked data mainly because it is simpler to use when compared with the chi-squared test [83]. A hypothesised cumulative distribution function (CDF) is compared with the empirical or sample cumulative distribution function. If the maximum discrepancy between the experimental and theoretical frequencies is larger than that normally expected for a given sample size, then the theoretical distribution is not acceptable for modelling the underlying population. On the other hand, if the discrepancy is less than the critical value then the theoretical distribution is acceptable at the prescribed significance level. The K-S test statistic can be defined as:

$$d = Max\,|F(x) - E(x)| \qquad [4.31]$$

where F(x) and E(x) are the theoretical and empirical distribution functions respectively. The function F(x) is a continuous function and the distribution of d does not depend on the underlying hypothesised distribution which makes the K-S test method computationally attractive.

Ahmad et al [87] developed a new approach to identify the best-fit time to failure distribution methods which provide a different perspective to reliability modelling. In the traditional approach, Least Square Estimator (LSE) and the Maximum Likelihood Estimator (MLE) are used. The LSE is utilised to specify the best failure fit failure distribution by examining all the possible time to failure distributions (lognormal, Weibull etc.). The MLE is then applied to calculate the parameters of the selected time to failure distribution. With the new approach, the LSE method is used to determine the β parameter of the Weibull distribution. The value of the β parameter can then be used to determine the best-fit failure distribution using the MLE technique. A comparison of the old method and the new approach is presented in Figure 4-12.
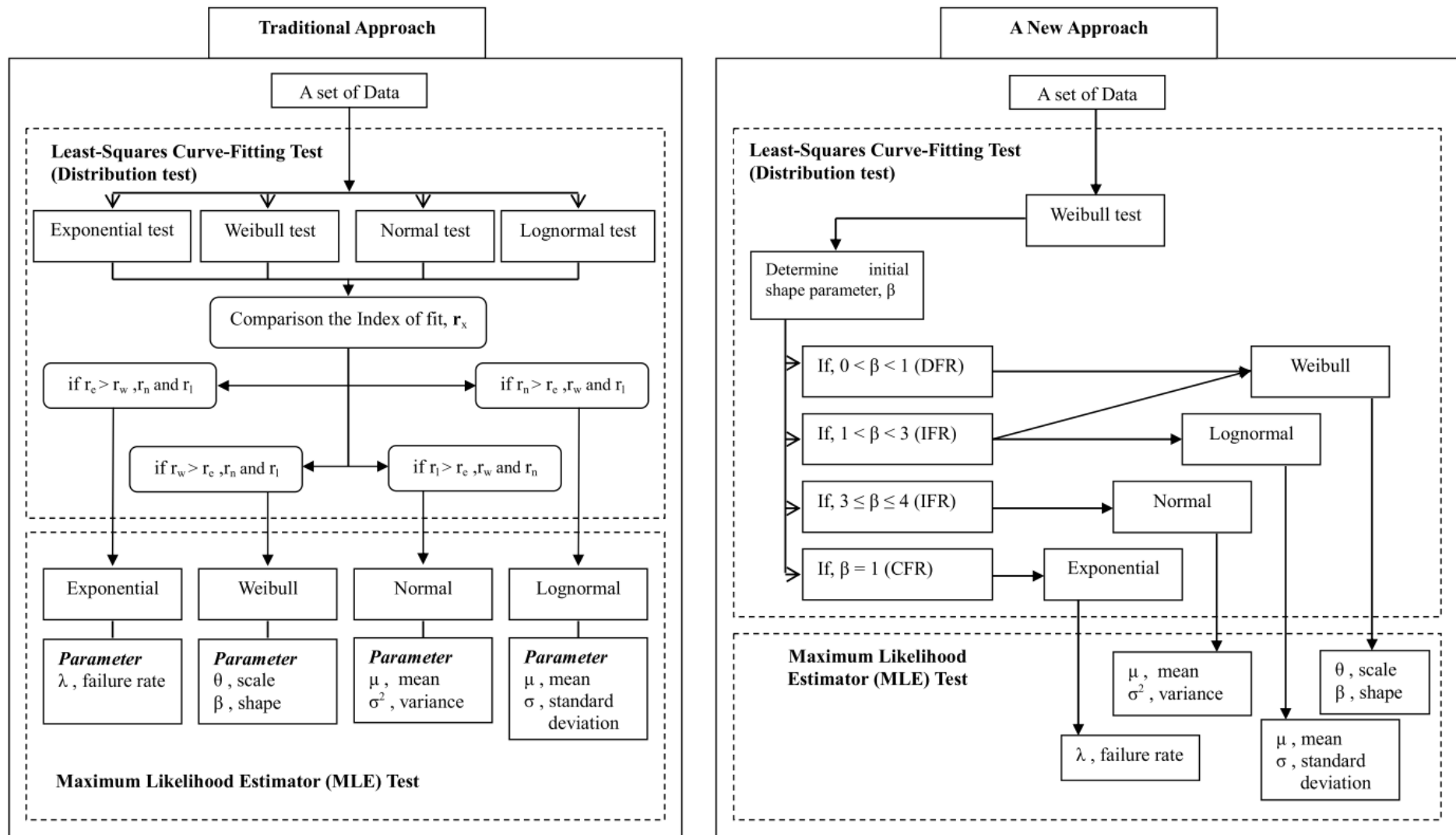
Figure 4-12 : Comparison of the traditional and new approach adopted from Ahmad et al [87]

4.4 Section summary

The author studied reliability concepts that contribute towards developing a reliability model to quantify the reliability of railway infrastructure. This section presented a basic understanding of a reliability theory. The failure processes and characteristics governing railway infrastructure were explored and methodologies of analysing failure data were provided. A methodology for quantifying the reliability of railway infrastructure will follow a general process summarised in Figure 4-13. The model is built on a high-level approach to quantify the reliability of railway infrastructure. The initial step is to characterise the system based on the failure data collected for railway infrastructure failures. A system function will be developed to model the configuration and behaviour of the system using reliability theory. The failure modes are established using the methodologies presented in this section and will be utilised to construct a functional model of the railway infrastructure system. Each railway infrastructure subsystem contributes to the overall performance of the infrastructure system, therefore each subsystem will have a function which represents its behaviour that will ultimately be modelled into a single system function using reliability theory. A subsystem function is one which assumes the correct failure distribution to match the specific subsystem. A methodology for selecting the appropriate time to failure distribution has also been presented. Once modelling is complete the reliability of the infrastructure system can be computed.

| Failure data collection | Functional and qualitative analysis | Statistical data anlaysis | Reliability modelling | Quantitative analysis | Quantifying reliability | Reliability forecasting |

Figure 4-13 : Reliability modelling procedure

5 Development of reliability model

The preceding sections have presented the theory that is required to develop a holistic reliability model for railway infrastructure systems. This section applies the theory to a practical case study by developing a reliability model to quantify the reliability of PRASA's Western Cape railway infrastructure system. A background of PRASA's maintenance management will be provided along with a data analysis approach to study the collected data for parameter estimation in developing the reliability models. Additionally, a comprehensive failure mode analysis is presented to assist in characterising the functional relationships and interdependencies in the infrastructure.

5.1 PRASA maintenance management

The passenger Rail Agency of South African is a wholly owned state company which operates the Metro commuter long-distance, intercity, and cross-border services known as Metrorail. Metrorail operates in the major metropolitan areas in South Africa transporting over 1.7 million passengers per week across 3 180 km of rail line. Of the 468 passenger rail service stations, 374 are owned by PRASA [96]. The network lines are developed and maintained by the regional Metrorail offices. Figure 5-1 shows the Metrorail Network for the Western Cape Province that will formulate the basis of this case study.
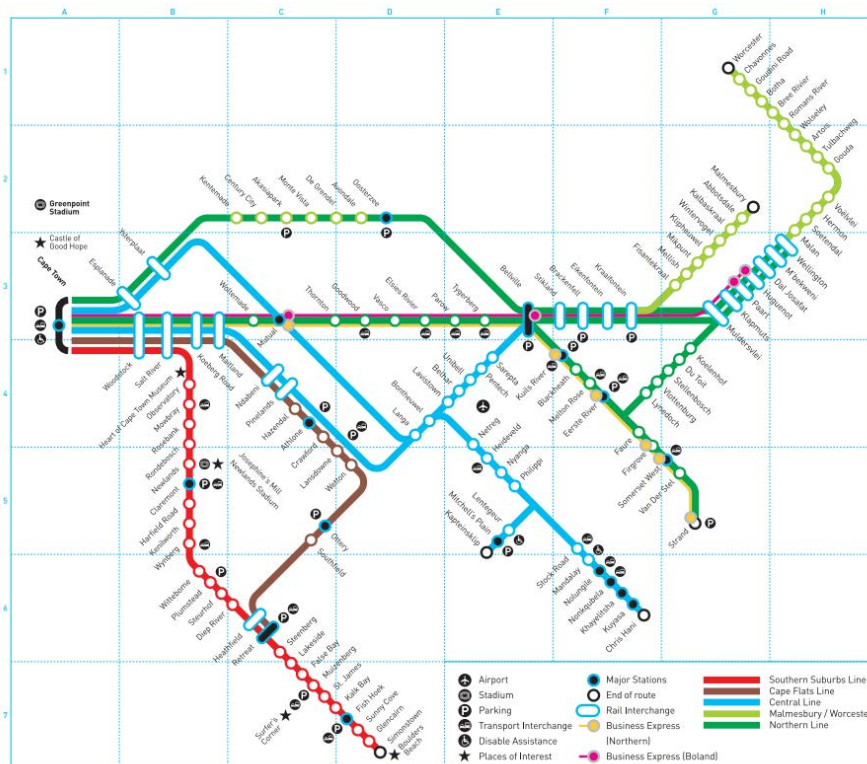
Figure 5-1 : Map of the Cape Town Metrorail network

58

The organisation of the PRASA maintenance department is split into engineering services and maintenance operations. The engineering services department is responsible for planning, policies, and procedures in facilitating the execution of maintenance-related tasks. The engineering services department is divided according to the infrastructure subsystem. Each department within engineering services has its own specific RAMS and RCM framework that are followed in executing the infrastructure asset management strategy. The maintenance operations department is responsible for executing the plans and procedures and provides maintenance support to the engineering services department. The two divisions, therefore, mirror each other and coordinate all infrastructure-related interventions on the railway network. It is, however, part of a bigger framework which has parallel strategic and delivery components relating to the operation of the network such as supply chain and human resource management as shown in Figure 5-2.
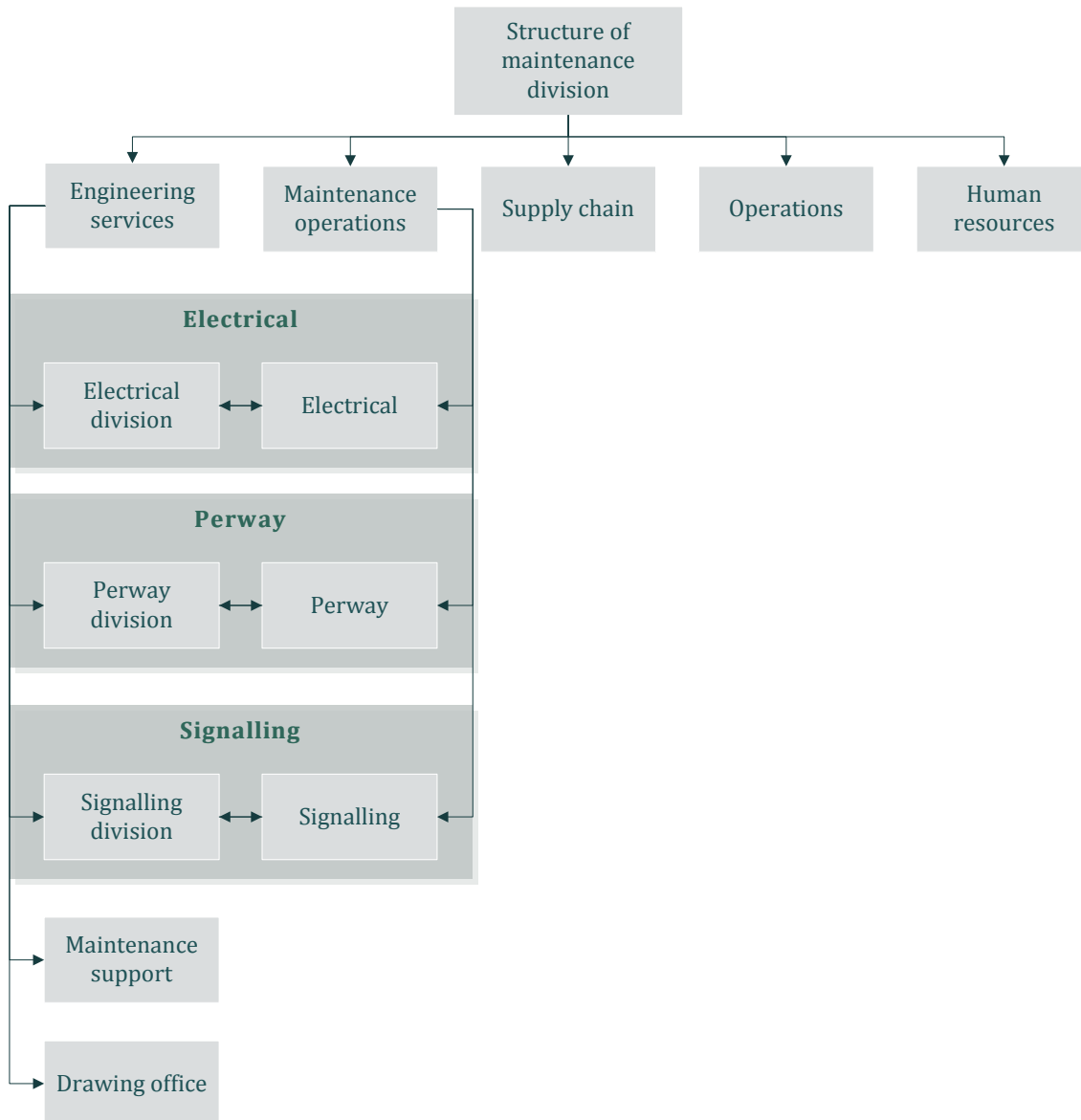
Figure 5-2 : Organisational structure of Metrorail maintenance division

The Enterprise Maintenance Planning and Control (EMPAC) is the Integrated Management System (IMS) used in PRASA's maintenance management operations. This system documents the performance, planning and budgeting of all maintenance management-related activities by generating statistics, reports, and summaries on the performance of the railway network. The general indicators of infrastructure performance obtained from the system include the number of delays and cancellations caused by each infrastructure system. The infrastructure access and planning process of PRASA articulates a maintenance strategy which comprises of all activities that require secure access to the railway passenger service by improving the availability and reliability of rolling stock and infrastructure systems.

Maintenance dimensioning in PRASA addresses the issue of resource allocation across the infrastructure network by considering traffic volumes, safety, reliability and the economic needs that impact the decision-making process. The performance of the maintenance intervention strategies developed by the engineering services is measured using the number of productive hours spent on an asset during maintenance operations. The travel time to restore system failure is categorised as unproductive hours; unavailable hours refer to the time where maintenance resources are unavailable. This performance measure captures the scope of PRASA`s infrastructure maintenance management which focuses on a preventative maintenance plan strategy. A preventative maintenance plan is the first line of defence for ensuring minimal infrastructure failures and consists of routine tasks, planned tasks, and feedback systems on the tasks performed. Figure 5-3 summarises PRASA's asset management decisions and activities arranged in a plan-do-review framework. The framework provides a simple representation of the major building blocks of asset management and the key interfaces between them. In addition, it provides a detailed process mapping the different responsibilities assigned in the asset management systems strategy. PRASA recognises that maintenance is a technical process and as such a maintenance programme needs to be managed in a manner that yields greater service reliability, ultimately enhancing the commuter experience. To achieve this means spending more productive time on the infrastructure assets to keep the condition of the assets at acceptable operating levels.

Figure 5-3 : Scope of activities for PRASA`s asset management framework.

5.2 Data analysis

Data from PRASA`S Information management system (IMS) was analysed to demonstrate the theory of failure statistics used to develop reliability models presented in section 4.3. Data analysis is the process of cleaning and analysing raw data for input into a developed model to produce the desired outcome. The fundamental aspect of the data analysis and modelling approach is based on the relationship that is established between the railway system reliability and the transportation service level offered by the system itself. This information is important to the railway company from a practical point of view because it enables a systematic evaluation of maintenance policies and plans while identifying and verifying the reliability targets for the infrastructure subsystems.

The framework for reliability evaluations is a continuous systematic analysis which must be applied at the relevant levels of the railway network. To achieve this the researcher established the format and structure in which the data is recorded within the IMS. From a maintenance perspective, there is a difference between a point and linear assets depending on the criticality and the length of the asset. For point assets maintenance is not assigned to a particular length of the asset but rather to the entire asset or to some of its indenture levels. A linear asset, on the other hand, is an asset whose length plays a central role in its maintenance, an example being the track or catenary system. The inventory from the IMS accounts for these characteristics and defines the location of a point or a section of the network to describe an infrastructure asset. When a failure event occurs on the railway network the location of the failure is defined by a point or section along the asset between the geographically closest stations. Using the network topology

61

map the location of failures on a network section can be identified and traced in accordance with the asset tags specified in the asset registry. A hierarchical representation of the infrastructure indenture levels formulates the modelling methodology. This approach ensures that infrastructure failures with critical consequences on the operation of the railway service are given attention. The indenture levels followed to analyse the data in developing the reliability model for the analysis are shown in Figure 5-4.
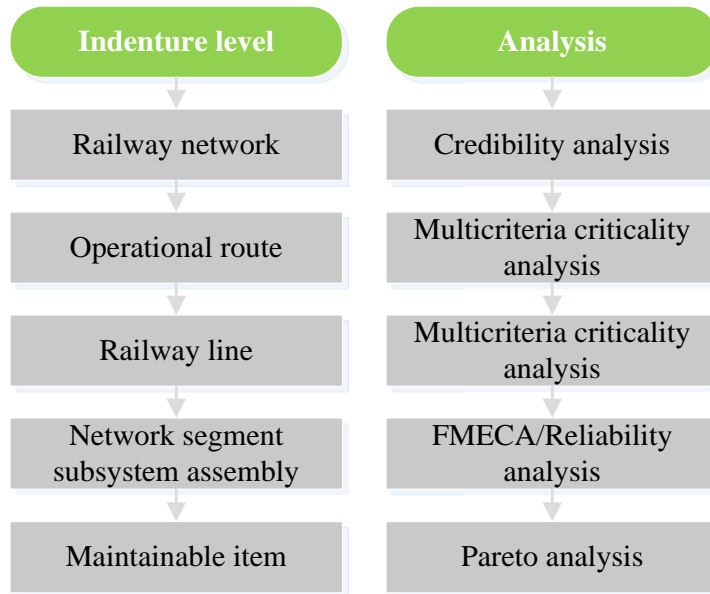
Figure 5-4 : Breakdown structure for reliability evaluation to support the modelling of the infrastructure network [53]

5.2.1 Failure data analysis

Service reliability is measured by the number of trains cancelled and delayed. Therefore, from the point of view of a transportation process, the most significant failure consequences are delays and cancellations. The researcher identified the infrastructure-related incidences that impact the quality of service as interpreted by train delays and cancellations. This approach only takes into account recorded failures that cause system downtime (unavailability). Referring to Figure 5-5 a typical failure episode of an individual linear asset is shown. An asset can be in either of two possible states (available or unavailable). The state of a system at a time t can be described by a state variable $s\,(t)$:

$$s(t) = \begin{cases} 1, & \text{if the subsystem is functioning at time } t \\ 0, & \text{if the subsystem is in failure state at time } t \end{cases} \qquad [5.1]$$

Once an asset experiences a failure at a given time the asset starts to malfunction. After a reaction time, the failure is registered and a work order is opened with the aim of restoring the normal activity of the asset. The distinction between the time to the first failure and the time between failures will be applied using repairable systems theory. The time to failure is understood as the

time elapsing from when the item was put into operation until it fails for the first time and is interpreted as a random variable T. To apply statistical analysis the variable is not always interpreted in calendar time but can be a discreet or continuous variable which determines the random distribution used to model the reliability. To simplify the process of data analysis, suspension of the railway infrastructure system was assumed to occur between the start and end of the data sets. This means that the period under study using the collected data assumes uninterrupted operation. Thus any system downtime is assumed to be as a result of infrastructure failure events as recorded in the database.

TTF : Time to first failure
TBF : Time between failures
TTM /DT : Time to maintain/Down time
UT : Up time, available state
WT : Waiting time
TTR : Time to restore
WO : Work order
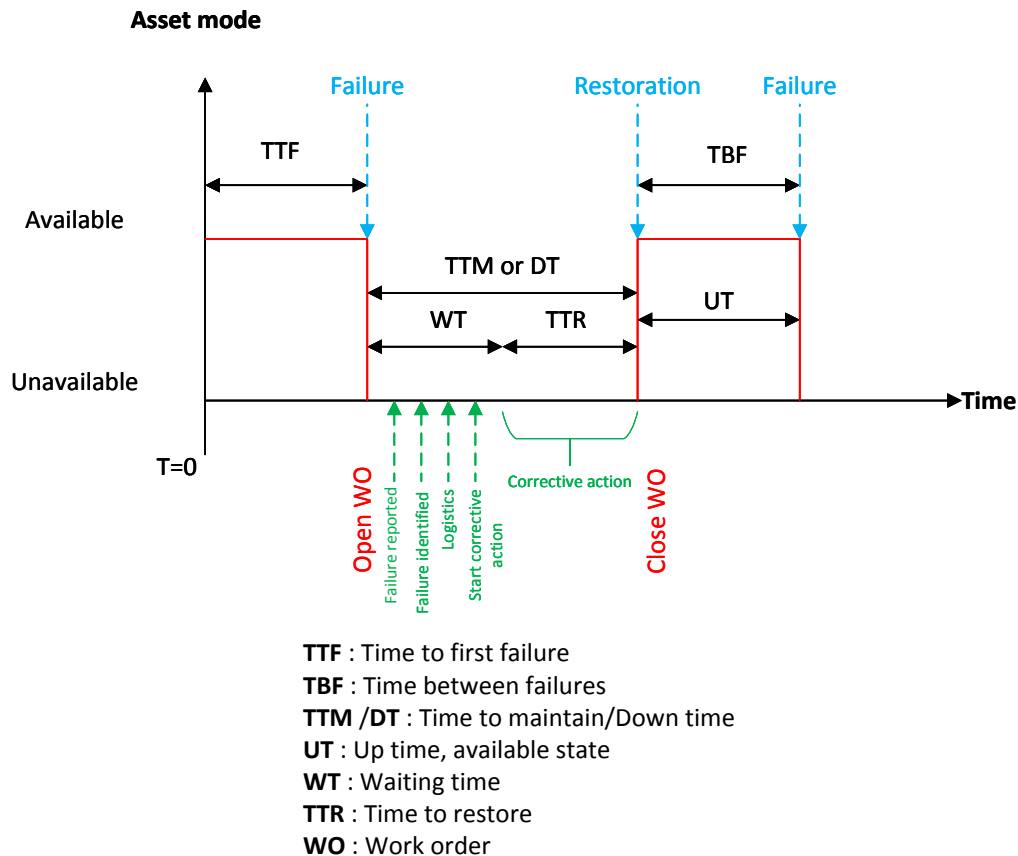
Figure 5-5 : Failure episode and definition of terms

The researcher studied the weekly failure data recorded to identify the failures reported on the network by looking at the different corridors and operational routes on the network. The weekly report logs all the daily failure information according to infrastructure type and provides information on the location, asset ID, failure date, and cause of failure for the different infrastructure subsystems. Table 5-1 shows an example of a daily failure log for the signalling system and the impact that such an incident has on train service reliability. To trace a failure to a line corridor the train stations that fall on that corridor must be determined to establish the failures collected at these stations on the network. The reality is that not all items under study registered in the asset registry will contain a failure event. The modelling approach taken by the researcher can only be certain that a number of items have not failed in a particular period, not

63

knowing whether they would have failed after a longer period. Sections, where smaller errors are present, were preferred for the analysis. Additionally, a study of the different corridors in the network performance revealed duplicate data entries that had to be removed to avoid multiple entries.

Table 5-1 : Daily failure logging for signal failures

	Direct			Consequential			Total			Manageable Delays & Cancellations
	Train	Min	CX	Train	Min	CX	Train	Min	CX	
Signals	54.1	952	1				54.1	952	1.0	Points 5433 defective at Bellville, Adjust Blade Tension. Points 6013 defective at Salt river affecting the number 2 Flats\Kapteinsklip line. Defective signal WTD 6240 at Ottery, cleared after passage of train. TCO Philippi panel reported that train 9357 reported that the signal fell back at danger. Ongoing closure of Firgrove station after the intervention of the RSR since 05\11\2016. Track Circuit A1052 faulty at Bonteheuwel, Cable rusted. Defective Track Circuit 5842 at Salt River, Repaired Staggering.

From the analysis of the different corridors, the cleaned failure data was utilised to model the set of arrival times of each infrastructure subsystem for the application of repairable systems reliability theory presented in section 4.2.3. The prediction intervals chosen account for the statistical uncertainty in reliability predictions that occurs because of limited data samples and variability in system failures. In cases where the failure times of the infrastructure subsystems took values in a particular range, the data was truncated to remove the uncertainty and bias that may occur in statistical approximation because of inconsistencies in the recording of failures. To validate the model, resampling or cross-validation techniques will be used. With these techniques, a complete data set is divided into two subsets. The first set becomes the training set that is used for model selections and parameter estimations; the second set, which is known as the validation set, is used for model validation and error estimation. Application of future forecasting will be tested in this manner.

5.3 Failure mode and effect analysis

Systems fail because of different failure modes. For infrastructure systems, some failure modes behave differently and it is generally easier to establish the time to failure distributions of the individual failure modes. For some infrastructure subsystems, only one failure mode is of critical importance. The failure mode information was traced and classified from the failure event data. Failure modes were assumed to be statistically independent, meaning that one failure mode increases the probability of failure of another failure mode. The analysis of failure mode data with this assumption simplifies the analysis of multiple failure modes similar to that used by single failure mode. Additionally, failure modes with negligible severity in terms of service interruption were omitted.

The aim of this exercise is to assist in reliability modelling of the infrastructure system, in that the analysis gives attention to the failures that disrupt the performance of the system more in comparison to others. Infrastructure failure modes are classified according to the consequence that they have on the system. The most significant failure consequence is a delay and in extreme instances a service cancellation with other consequences related to a reduction in track capacity and speed restrictions. The classification of failures used is shown in Table 5-2 according to the consequences. The combination of the frequency of occurrence and severity of impact guides the classification of the infrastructure failure modes. The probability of occurrence used by the researcher to classify the failure modes is shown in Table 5-3. The correlation between the type of infrastructure elements and the number of occurring failures and the methods provided in section 4.2 were used to establish the criticality of failure modes. A matrix shown in

Table 5-4 is created using the Military Handbook technique to determine the criticality of the infrastructure failure modes. The criticality index is shown in Table 5-5.

Table 5-2 : Classification of infrastructure failure modes

	Consequences for system
Catastrophic	Cancellations
Critical	Delays
Marginal	Capacity lowered
Insignificant	No service disruption

Table 5-3 : Probability of occurrence of the infrastructure failure modes

Occurrence	Frequency(week)	Description
Very high	>30	Persistent infrastructure failures
High	>20	Failures will occur frequently
Moderate	>15	Likely to occur occasionally
Low	>10	Relatively low failures. Probability of occurrence low
Remote	>5	Unlikely to occur but possible.

Table 5-4 : Matrix to evaluate criticality

		Severity			
		Insignificant	**Marginal**	**Critical**	**Catastrophic**
Frequency	**Very high**	R 3	R 4	R 4	R 4
	High	R 2	R 3	R 4	R 4
	Moderate	R 2	R 3	R 3	R 4
	Low	R 1	R 3	R 3	R 4
	Remote	R 1	R 2	R 3	R 4

Table 5-5 : Relationship between level of risk and mitigation measures

Criticality index	Evaluation	Definition
R 1	**Negligible**	Acceptable
R 2	**Tolerable**	Acceptable with adequate controls and agreement with different infrastructure departments
R 3	**Undesirable**	Acceptable only when impact in impracticable
R 4	**Intolerable**	Should be eliminated

5.3.1 Railway infrastructure failure modes

Based on a failure mode and effect analysis, several objects in the railway infrastructure system are prone to failure. This section discusses the outcomes of the failure modes and effect analysis performed on the railway infrastructure system.

5.3.1.1 Signalling subsystem

Failures attributed to faulty track circuits accounted for the highest rate of occurrence within the signalling subsystem. The functionality of track circuits is affected by the failure of its components and changes in the characteristics of the track. Track maintenance activities were also identified to affect the functionality of track circuits. Other failure modes that had a significant occurrence were related to the interlocking and point-to-point machines. The occurrence of cable and wire discontinuities is attributed to high levels of vandalism and were identified as the biggest contributors to power-related signalling failures. An example of a failure analysis of an occupied track failure event is shown in Figure 5-6. Failures related to false occupation occur randomly and are unpredictable but they constitute a significant subset of track circuit failures and can be triggered by bad workmanship during preventative maintenance actions or by the vibrations due to rolling stock during uptime.
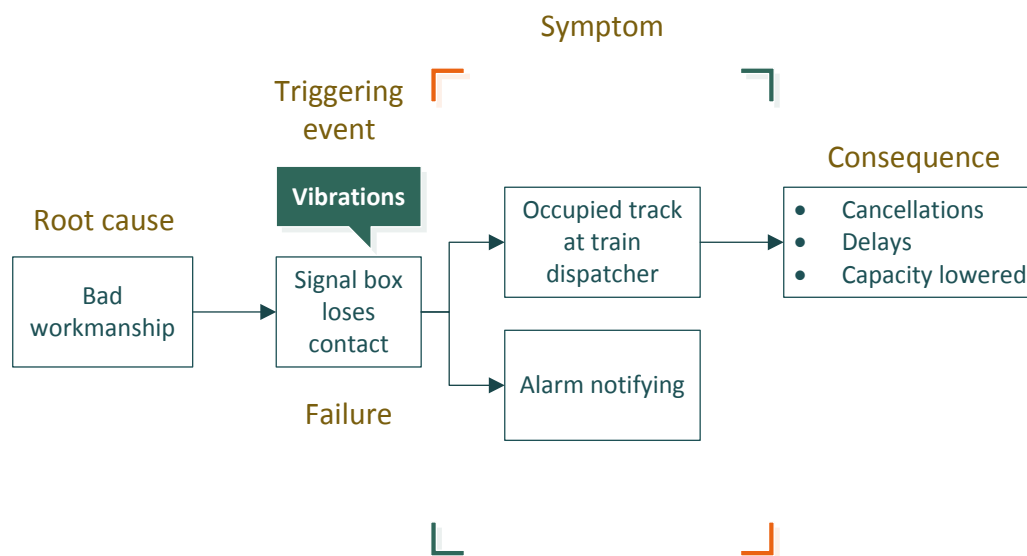
Figure 5-6 : Failure analysis of 'Occupied track events'

5.3.1.2 Perway subsystem

A study of the failure data related to perway failures reveals that insulated rail joints have a shorter life cycle than other components of the track. The failure frequency of rail joints is evidenced by the high occurrence of failures related to faulty block joints. This is caused by continuous tonnage due to traffic use. The ballast is a significant component within the track substructure as it influences the failure pattern of the perway infrastructure system. The

identified causes of ballast failure are related to voiding and settlement. In addition, the condition of the ballast influences the track circuit in the signalling subsystem. This is because the ballast offers electrical resistance and the track circuit is only functional at specific ballast resistance levels. If this resistance drops to values lower than that specified, the flow of current drops and makes the track circuit non-functional. The occurrence of such failures is intermittent in nature and more likely to occur during the wet winter season than in summer.

The most severe infrastructure failure related to the perway subsystem is a derailment. Falling levels of infrastructure renewal and worsening track quality results in high dynamic forces during operation which may lead to broken and or defective rails. Broken and defective rails are the highest causes of derailments which can have fatal consequences. Furthermore, faulty rails trigger track circuit failures which affect the performance of the signalling subsystem. The study observed that there is a distinction between a broken rail and a defective rail as such – a defective rail is not considered a broken rail. A broken rail is a rail with a complete break or a missing piece. Exceptions are rails that break in possessions and in sidings. A defective rail, however, is a rail identified as containing defects that are related to geometry and the characteristics of the track such as alignment defects. Other failures related to the perway subsystem can be attributed to rail clip and sleeper failures which are a result of high rates of vandalism on the network's infrastructure.

5.3.1.3 Electrical subsystem

The electrical subsystem is the core of any electric railway transport system and its criticality is emphasised by the impact that it has on service cancellations as compared with other infrastructure subsystems. The network under investigation has 3 kV and 11 kV transmission lines which supply power to the overhead track equipment and the signalling system. Failures related to the overhead track equipment are attributed to pantograph hook-ups, fallen trees, and electrical power failures at the substations. Activities related to the maintenance interventions are prone to trigger failures related to the overhead track equipment. Routine maintenance actions on the perway subsystem may lead to a track settlement which increases the gap between the pantograph on the train and the overhead contact wire. Substations are characterised by failures related to faulty switches and circuit breakers. There is, however, redundancy at the substations of the 3 kV and 11 kV substations which ensures that the power is always available for the OHTE (Overhead traction equipment) and signalling subsystems equipment. An investigation of infrastructure-related failures reveals that electrical subsystem failures have a relatively low frequency but cause a significant number of delays.

5.4   Characterising infrastructure dependencies

As highlighted in section 2, railway infrastructure is a complex system that has interdependencies and dependencies between the different subsystems. To develop a reliability model for railway infrastructure the researcher mapped out all possible interdependencies and flow relationships between system components. The method of empirical approaches to characterise interdependencies has been presented in the literature. Mapping the key interdependencies between the infrastructure assets was based on the operational requirements of the assets, and the results of the failure mode and effect analysis. Furthermore, after consultation with the engineering services department personnel, the functional relationships between the infrastructure subsystems to develop a reliability model for the railway infrastructure systems was established.

The researcher utilised a dependency matrix to map out the different infrastructure dependencies that exist in the railway infrastructure system. This can be seen in Appendix A. Some infrastructure assets exhibit both unidirectional and bidirectional interdependencies. Modelling these characteristics is important to estimate the physical and functional propagation effects of failures. Failure propagation decreases the quality of service due to the loss of physical interactions and functional relationships between connected assets in the infrastructure system. Figure 5-7 shows the interdependencies and functional flow diagram of the railway infrastructure system. Single arrows indicate a unidirectional relationship while double arrows indicate a bidirectional interdependence of the infrastructure assets. The track circuit, OHTE, and signalling power depend on the uninterrupted availability of electric power from the substations and transmission lines exhibiting a unidirectional dependence. On the other hand, the OHTE and perway superstructure exhibit a bidirectional interdependence between the infrastructure components.
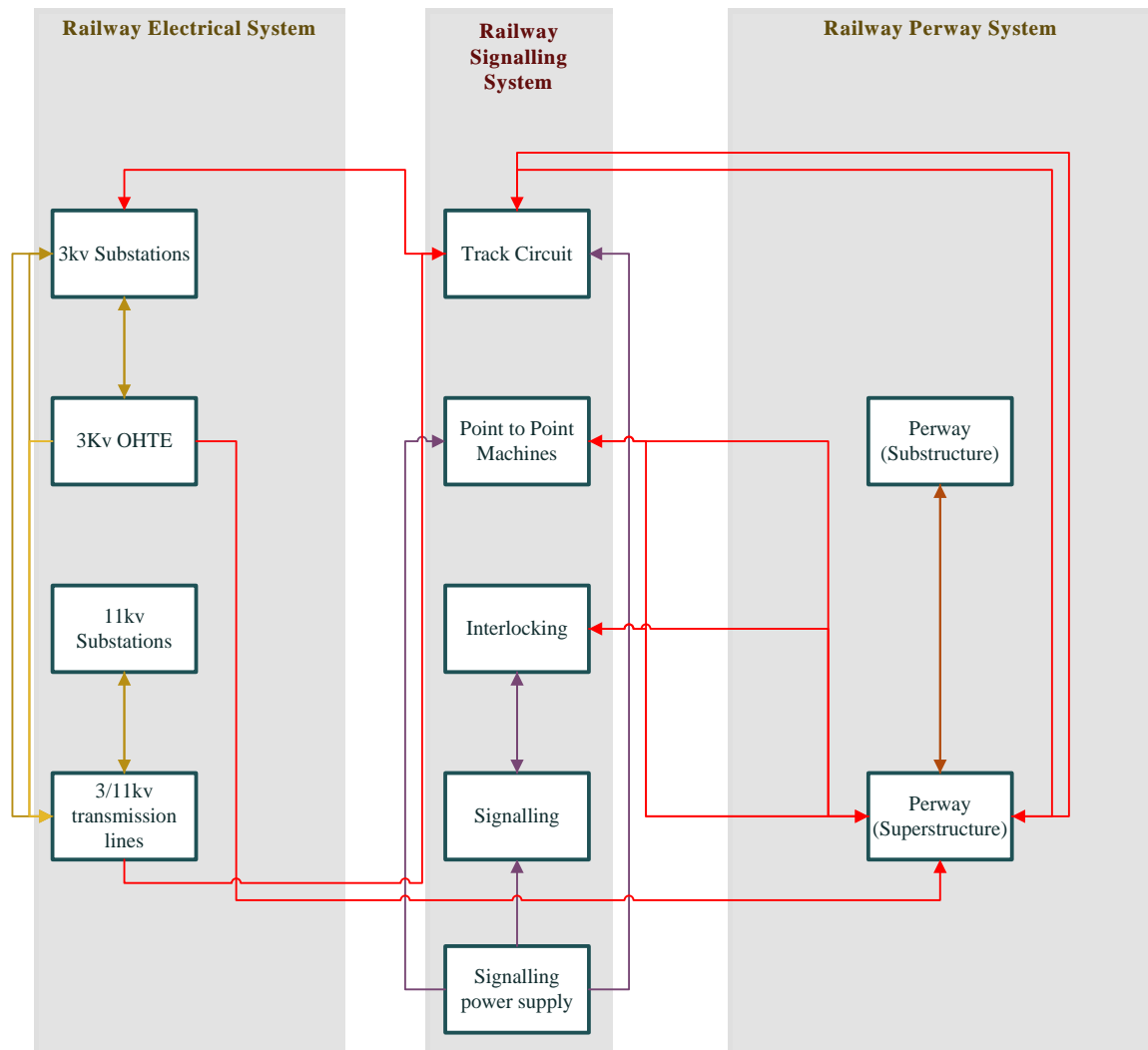
Figure 5-7 : Interdependencies and Flow Relationships

5.5 Railway infrastructure reliability model

The concept that formulates the modelling methodology is based on a hierarchical representation of the railway infrastructure network. This allows the analysis to be performed at different levels of granularity ranging from an individual maintainable item to a large multi-asset network.

The railway network topology for the infrastructure system can be assumed to consist of indenture levels as shown in Figure 5-8. Utilising a top-down approach, the whole rail infrastructure network can be broken down into operational routes representing the different parts of a railway network. The operational routes constitute a specified number of lines made up of multiple segments representing a corridor between two locations (stations) or a section between two signals called a signal block. Multiple segments characterise individual maintainable items according to technical and functional properties to represent the distinct infrastructure subsystems. Individual maintainable assets for which degradation mechanisms and intervention processes can be determined are lowest on the indenture level.

70
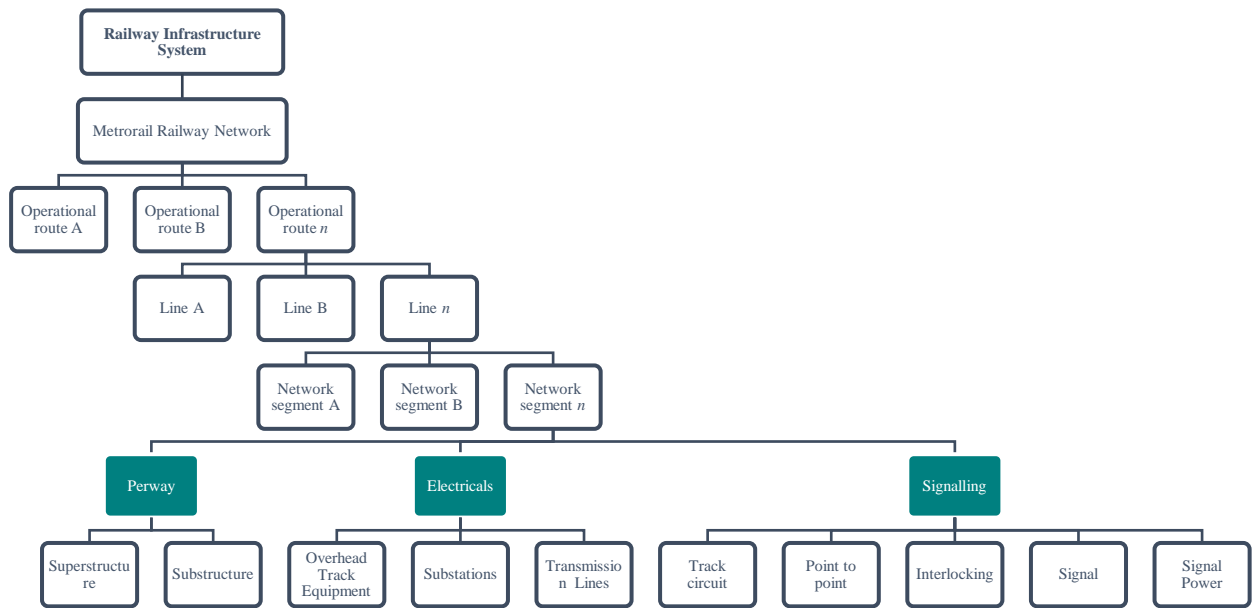
Figure 5-8 : Infrastructure indenture levels for reliability modelling approach

Figure 5-9 shows an example of an operational route between two stations that constitute part of a larger network with point and linear assets. This configuration examines the relationship between the point and linear assets and formulates the basis of the holistic infrastructure reliability model. The redundancies that exist in railway infrastructure systems, particularly the electrical system, were accounted for in the functional mapping of the reliability model developed for the network segment. This approach takes into account the most essential functional properties of the system to be modelled in order to provide a comprehensible reliability model.
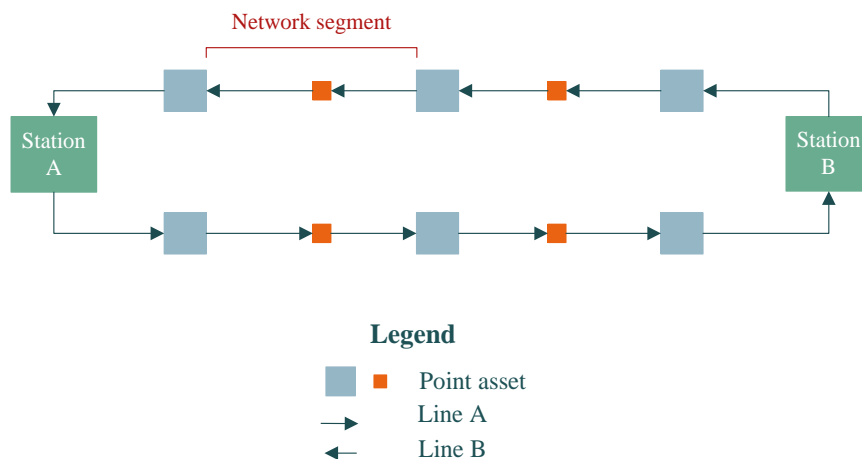
Figure 5-9 : An example of an operational route

To identify critical components that constitute a network segment for railway infrastructure system, systematic and exhaustive consequence investigations for the different component failures were performed. The practical issue, however, was analysing combinations of failures by assuming that they increase as the number of simultaneous failures increase. This assumption

71

allows a combination of failed components to not be restricted to only one particular infrastructure but rather to a combination of simultaneous component failures in the different infrastructure subsystems. It was further observed that some components are highly critical in themselves, therefore combinations of failure including these components will also be highly critical. However, highlighting these components as critical when looking at simultaneous failure adds minimal input to the modelling information, since their criticality would have already been taken into account when considering single failures. A functional reliability model representing a network segment is shown in Figure 5-10 for the railway infrastructure system. The model constitutes the core maintainable components required for a complete transportation process between two stations. It is assumed that there is no loss of service for as long as a path exists for train passage between two stations. Loss of service as a result of a malfunctioning infrastructure system is interpreted through delays and cancellations.
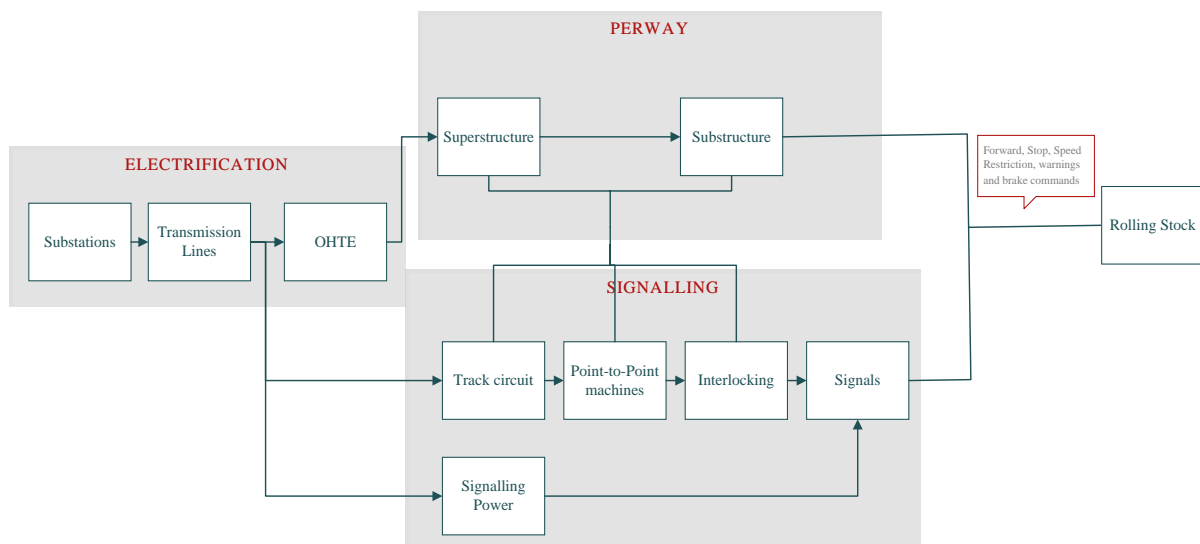
Figure 5-10: Functional reliability model of a network segment

From the functional reliability model the asset state models for the different infrastructure subsystems can be developed. The individual asset state model is built for a specific infrastructure's subsystems, taking into consideration the integration of the degradation-failure and intervention processes to simulate its state changes over time. The reliability block diagram for each infrastructure subsystem has a series configuration that represents the infrastructure state models as shown in Figure 5-11.
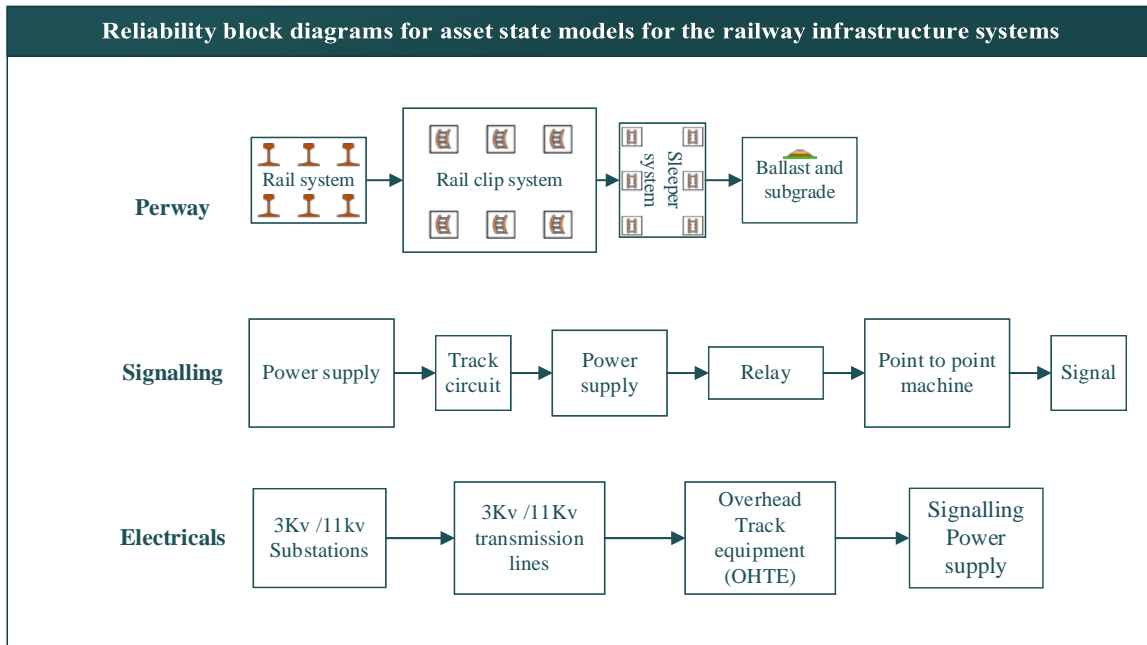
Figure 5-11 : Reliability block diagrams for the infrastructure asset state models

The infrastructure asset state models are used as building blocks for the infrastructure system state model. From the functional reliability model presented the system state model is a series configuration of the infrastructure subsystems as shown in Figure 5-12. A collection of infrastructure system state models assembled together construct network segment models that can be used to model higher network hierarchical and/or infrastructure indenture levels. The abstraction level and network system details govern the configuration of the network segment models. If the network segment models are combined at the relevant abstraction levels, railway lines and operational routes can be modelled holistically for performing reliability evaluations of railway infrastructure systems. The modelling approach shown in Figure 5-13 uses the system and subsystem utilisation information and the possible strategic interventions that influence the degradation process of the different infrastructure subsystems.
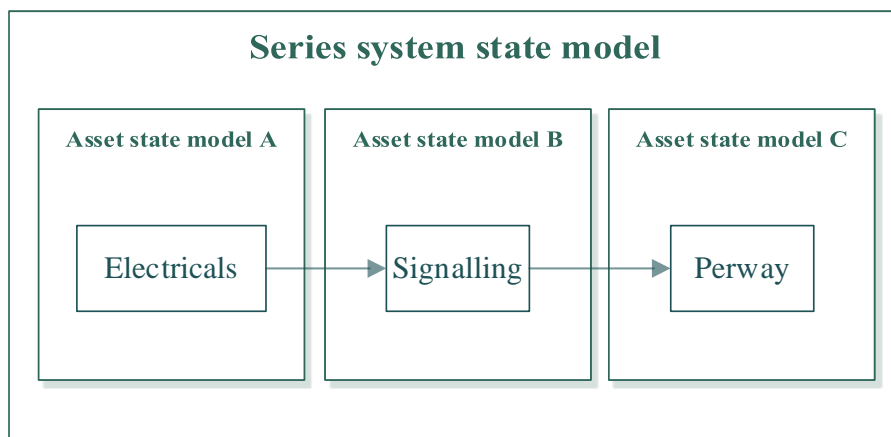
Figure 5-12 : Reliability block diagram for network segment railway infrastructure systems

73
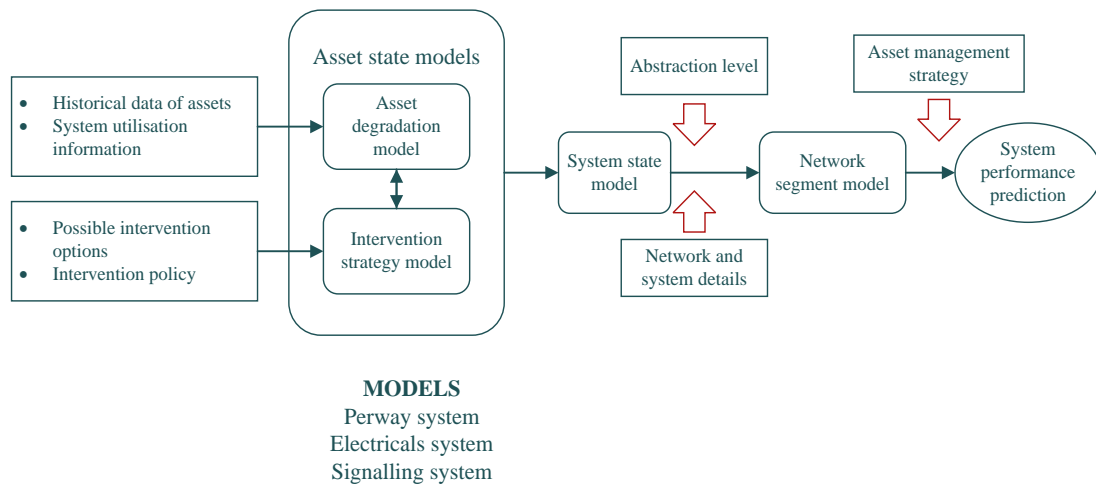
Figure 5-13 : Modelling approach showing asset state and system reliability model

5.6 Section summary

This section presented the reliability model of the railway infrastructure system to quantify the reliability performance of the railway infrastructure system. Attention was given to the complex functional and operational relationships between the different infrastructure subsystems. The methodology utilises infrastructure asset state models as the core building blocks of the reliability model of the system. Linear assets were segmented to identify the hierarchical taxonomy and the relations among their various component assets. This procedure helps to identify and analyse the system at the appropriate level to accurately quantify the reliability performance. The model is of a stochastic nature as such the data quality and quantity will be of fundamental importance as more quality in data results in less biased predictions of infrastructure performance.

6 Application of reliability model

This section demonstrates the application of the reliability model to quantify the reliability of railway infrastructure systems. The modelling methodology that has been presented in the previous section will be illustrated on PRASA's Metrorail network.

6.1   Reliability analysis of a single corridor

6.1.1 Data collection

The Western Cape Metrorail network has five lines which are the Northern, Southern, Central, Cape Flats, and Malmesbury-Worcester line. Indenture levels illustrated in section 5.5 highlighted that operational routes are constituted of multiple line sections. The simplest unit on which to apply the reliability model is a single multi-directional line. The Simons Town-Steenberg line is a single multi-directional traffic line that runs on the Southern Line which makes it suitable for the application of the reliability model. Applying repairable system theory to the corridor the data between January 2015 and December 2015 formed the scope of the analysis. The daily and weekly failure information for the corridor was scrutinised for all the failure data collected for the infrastructure assets in the scope of the study. The arrival times of failures were extracted from the failure data on the line corridor using the reliability modelling approach given in the appendix between January 2015 and June 2015. Cross-validation of the reliability predictions will be conducted using the second subset of data between July 2015 and December 2015. The extracted inter-arrival failure times for each infrastructure subsystem are given in Figure 6-1. The signalling subsystem registered 36 failures, the perway 9, and the electrical subsystem 8 failures.

SIGNALLING	
Interarrival time	N(t)
4	1
13	2
19	3
29	4
39	5
45	6
46	7
47	8
59	9
62	10
67	11
79	12
84	13
89	14
90	15
96	16
98	17
99	18
101	19
102	20
109	21
119	22
124	23
128	24
137	25
152	26
154	27
155	28
156	29
160	30
160	31
163	32
165	33
168	34
169	35
177	36

PERWAY	
Interarrival time	N(t)
20	1
34	2
66	3
67	4
84	5
88	6
132	7
168	8
178	9

ELECTRICALS	
Interarrival time	N(t)
70	1
77	2
96	3
125	4
128	5
129	6
145	7
165	8

Figure 6-1 : Inter-arrival times for the infrastructure failures

76

6.1.2 Trend tests

Following the modelling approach as given in Appendix A and utilising the appropriate statistical methods highlighted in section 4.3 the researcher obtained the test statistics for the infrastructure subsystems obtained from the inter-arrival times. The results of the trend tests are summarised in Table 6-1. The Laplace test statistic for the electrical subsystem $U = 2.041$, concludes that an NHPP model is applicable for modelling the electrical subsystem. Furthermore, a Laplace test statistic of $U > 2$ shows a system in a degrading state. The Laplace test statistic for the signalling subsystem is in range $1 < U < 2$ which is a grey area that cannot classify a trend. Further Lewis Robinson Tests yielded a test static which concluded an NHPP model to be more appropriate for modelling the signalling subsystem. The NHPP log-linear and power law models were applied for both the signalling and electrical subsystems and were subject to further tests to determine the appropriate model that best fits the data. Laplace tests for the perway subsystem were non-committal; however, using the Lewis Robinson tests suggested an HPP model that follows a two-parameter Weibull distribution.

Table 6-1: Summary of the test statistic and the recommended modelling distributions.

Subsystem	Data points	Laplace Trend Test	LTT interpretation	Lewis Robinson	Model
Perway	9	0.234	Non-committal	0.313	Weibull
Signalling	40	1.782	Grey area	2.028	NHPP
Electricals	8	2.041	Reliability degradation		NHPP

6.1.3 Parameter estimation

A best of fit test was performed on the NHPP log-linear and power law models before determining the parameters of the distributions, to establish whether the model is representative of the data. The cumulative number of failures against time provides a good indicator as to whether a system is deteriorating or improving and is a standard tool for fitting failure models to failure data. A graphical comparison shown in Figure 6-2 reveals that both the power law and log-linear law are suitable for modelling the failure processes of the signalling subsystems. Similarly, a graph given in the appendix for the electrical subsystem shows the same trend. The CDF (cumulative distribution function) for the Weibull function is shown in Figure 6-3. The graphical fit shows that the Weibull function approximates the data sufficiently. A selection of best fit was performed on the data sets for all subsystems using the Kolmogorov–Smirnov Test. The results summarised in Table 6-2 concluded that the Weibull distribution is representative of the data for the perway

subsystem whereas the power law is more representative of the data for the electrical and signalling subsystems.
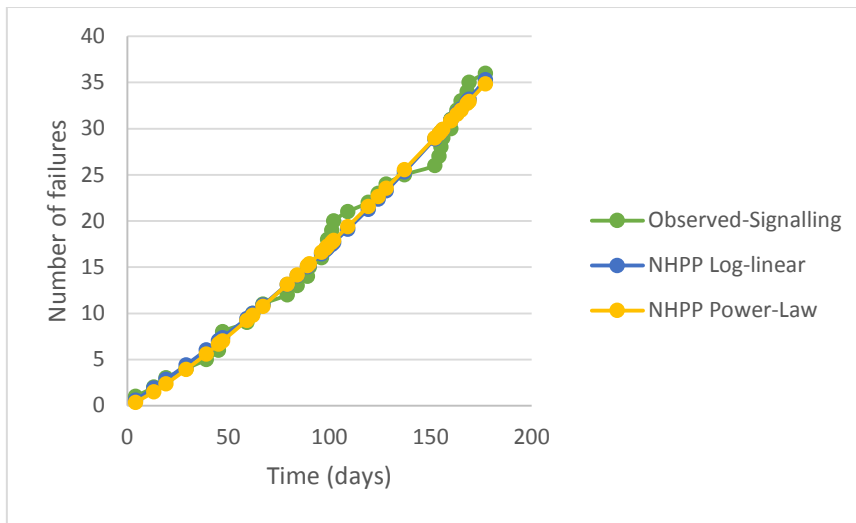
Figure 6-2 : Graph of the power law and log-linear law for the signalling system
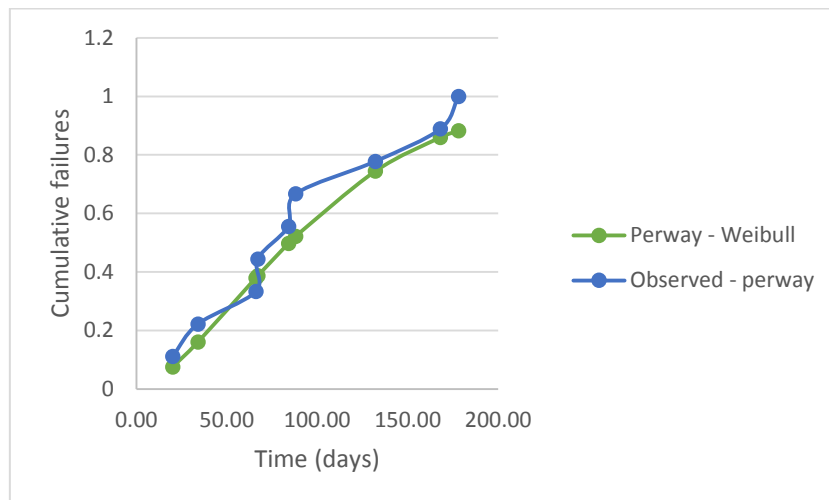
Figure 6-3 : Cumulative distribution function for the Weibull distribution and observed values

Table 6-2 : Summary of parameter estimation and K-S test

Subsystem	Models	K-S Test Result		Parameters	
Perway	Weibull HPP	d_{max} < $d_{critical}$ 0.1816 < 0.6082	Good fit	η = 107.54	β = 1.5127
Signalling	NHPP	d_{max} < $d_{critical}$ 0.0103 < 0.0475	Power law	λ = 0.0663	β = 1.2104
Electricals	NHPP	d_{max} < $d_{critical}$ 0.0303 < 0.2267	Power law	λ = 0.000345	β = 1.9770

6.1.4 Reliability predictions

Using the parameter values obtained from the statistical analysis the reliability of the corridor can be determined using the equations described in section 4.2.3. The reliability function for perway with parameter values of η = 107.54 and β = 1.5127 is used to calculate the reliability of predictions of the perway subsystem. The shape function, β lies in the range 1 < β < 3 which indicates an increasing failure rate. The reliability of the perway system at a time T_n of the railway infrastructure system can be calculated using equation 6.1.

$$R(t) = e^{-\left(\frac{T_n}{\eta}\right)^{\beta}} \qquad [6.1]$$

$$R(t) = e^{-\left(\frac{T_n}{1.5127}\right)^{107.54}} \qquad [6.2]$$

Similarly, the reliability equation for the power law shown in equation 6.3 applied to the signalling and electrical subsystem using the estimated parameters yields equations 6.4 and 6.5 respectively. For each of the infrastructure subsystems, the reliability predictions are determined from the time of the last failure.

Power law

$$R(t) = e^{-\lambda\left(T_2^{\beta} - T_1^{\beta}\right)} \qquad [6.3]$$

Signalling subsystem

$$R(t) = e^{-0.0663\left(T_2^{1.2104} - T_1^{1.2104}\right)} \qquad [6.4]$$

Electrical subsystem

$$R(t) = e^{-0.000345\left(T_2^{1.9770} - T_1^{1.9770}\right)} \qquad [6.5]$$

Using the reliabilities of the individual asset state models, the reliability of the railway infrastructure system state model can be determined using the appropriate reliability modelling equations. The reliability block diagram for the railway infrastructure system state model developed in section 5 concluded that the railway infrastructure system state model assumes a series configuration which follows the equation below.

$$R(t)_{system} = \prod_{i=1}^{n} R(t)_i \qquad [6.6]$$

$$R_{system}(t) = R_{perway}(t) \times R_{signal}(t) \times R_{electricals}(t) \qquad [6.7]$$

$$R(t)_{System} = e^{-\left(\frac{T_n}{1.5127}\right)^{107.54}} \times e^{-0.0663\left(T_2^{1.2104} - T_1^{1.2104}\right)} \times e^{-0.000345\left(T_2^{1.9770} - T_1^{1.9770}\right)} \qquad [6.8]$$

Equation 6.8 is used to calculate the reliability performance of the Southern line for the first 150 days of operation, Figure 6-4 shows a graphical representation of the reliability performance of the Southern line with time. Table 6-3 shows the predicted reliability performance of 48.2 % for the railway infrastructure system after 7 days. Reliability predictions were conducted from the last recorded failure for all the infrastructure subsystems.
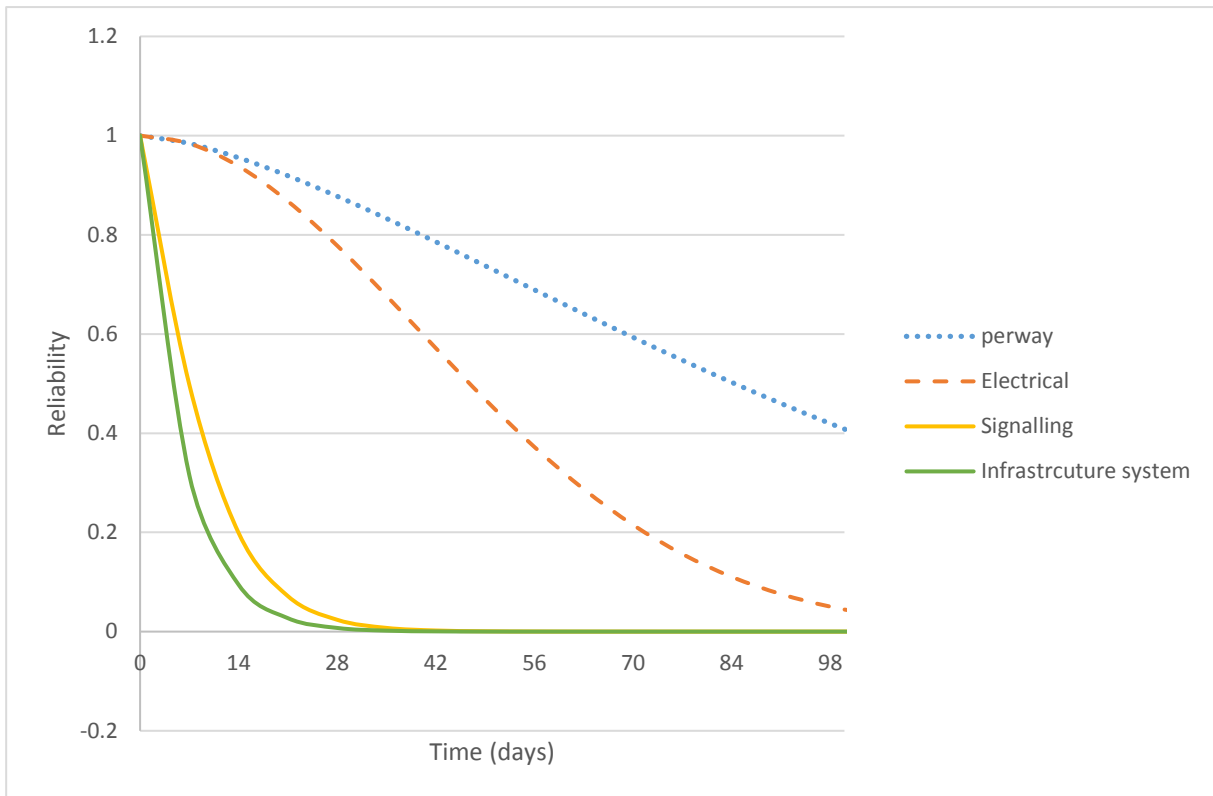
Figure 6-4 : System reliability for the railway infrastructure system

Table 6-3 : Reliability of the railway infrastructure system in the first 14 days of operation

R(t)	Corridor	Perway	Electricals	Signalling	System
14 days	Southern line	98.4 %	98.4 %	49.7 %	48.2 %

6.1.5 Validation of reliability predictions

The failure prediction for the different subsystems was conducted to check the extent of variations in the predicted and observed values of the time to failure (MTBF) and expected number of failures E (N). The cross-validation technique will estimate the time to the first failure using the equations given in section 4.2.3 for the NHPP log-linear, power law and Weibull functions. The observed values to be compared with those obtained from the model are extracted from the data in the second subset between July and December 2015. To present an accurate validation process, the prediction period begins from the last observed failure recorded in the first subset of data. Figure 6-5 shows the dates for the last observed failure for each of the subsystems. Time T=0 will be set at the date of the last observed failure for each of the subsystems.

<div style="border:1px solid #999; padding:8px; text-align:center;">
Perway subsytem
1 January, 2015 - 28 June, 2015
</div>

<div style="border:1px solid #999; padding:8px; text-align:center;">
Signalling subsystem
1 January, 2015 - 27 June, 2015
</div>

<div style="border:1px solid #999; padding:8px; text-align:center;">
Electrical subsystem
1 January, 2015 - 15 June, 2015
</div>

01 January 2015   **31 December 2015**
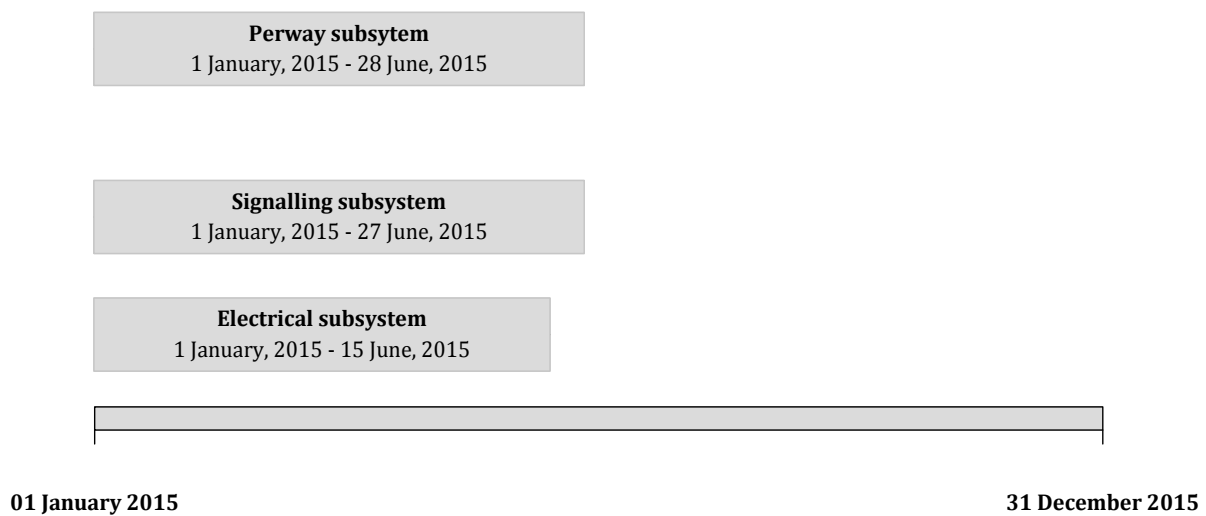
Figure 6-5 : Timeline showing the location of the last failure for the infrastructure subsystems

Using the equations presented in section 4.2.3 for determining the time to first failure (MTBF) and expected number of failures E (N) for the infrastructure subsystems, the validation of the results from the reliability predictions for each of the infrastructure state models follows.

6.1.5.1 Perway

The parameter values for the two-parameter Weibull function modelling the perway subsystem are η = 107.54 and β = 1.5127. To predict the time to first failure (MTBF) of the perway infrastructure subsystem. Setting T_2 = 186 days for the perway state model. The predicted time to first failure and expected number of failures for the perway subsystem is given as follows:

$E\left(T_2, T_1\right)$ MTBF (days):

$$E(T_2,T_1) = \eta\Gamma\left(1+\frac{1}{\beta}\right) \text{ where } \Gamma(n) \text{ is the gamma function}$$

$$E(T_2,T_1) = 107.54 \times \Gamma\left(1+\frac{1}{1.5127}\right) \tag{6.9}$$

$$E(T_2,T_1) = 107.54 \times 0901828$$

$$= 97\,days$$

Expected number of failures:

$$E(N(T_2 \rightarrow T_1)) = \left(\frac{\beta}{\eta}\right)\left(\frac{T_2}{\eta}\right)^{\beta-1} T_2 - \left(\frac{\beta}{\eta}\right)\left(\frac{T_1}{\eta}\right)^{\beta-1} T_1$$

$$E(N(0 \rightarrow 186)) = \left(\frac{1.5127}{107.54}\right)\left(\frac{186}{107.54}\right)^{1.5127-1}(186) \tag{6.10}$$

$$= 3.46\,failures$$

The actual inter-arrival time from the failure data is 17 days which means that the perway subsystem lasted 79 days shorter than predicted. The deviation in the results can be attributed to various factors. Weibull models with values of β > 1 have a failure rate that increases with time. This highlights that the reliability model assumes high failure rates with time. The reliability at the observed MTBF is 94.1%. The number of failures from the observed data N (t) = 4 failures, while the predicted number of failures in the same period is 3 failures.

6.1.5.2 Signalling

The power law parameters for the signalling subsystem are given as λ = 0.0599 and β = 1.2503. Setting T_2 = 187 days for the signalling subsystem. The predicted time to first failure and expected number of failures for the signalling subsystem is given as follows:

Time to first failure (TFF):

$$MTBF_2(T_1,T_2) = \frac{T_2 - T_1}{\lambda\left(T_2^\beta - T_1^\beta\right)}$$

$$MTBF_2(0,187) = \frac{187-0}{0.0599\left(187^{1.2503}\right)} \tag{6.11}$$

$$= 4.5\,days$$

Expected number of failures E (N):

$$E_p(N(T_2) - N(T_1)) = \lambda\left(T_2^\beta - T_1^\beta\right)$$

$$E_p(N(187) - N(0)) = 0.0599\left(187^{1.2503}\right) \tag{6.12}$$

$$= 41.48\,failures$$

The observed inter-arrival time after the last failure is 5 days, which means the signalling subsystem lasted 0.75 days longer than the prediction. The observed number of failures E (N) =

37 failures versus the predicted E (N) = 41 failures. The reliability of the signalling system when the first failure is observed yields 63.9 %.

6.1.5.3 Electrical

The power law parameters for the electrical subsystem are given as λ = 0.0599 and β = 1.2503. Setting T_2 = 199 days for the electrical subsystem. The predicted time to first failure and expected number of failures for the electrical subsystem is given as.

Time to first failure (TTF):

$$MTBF_2\left(T_1, T_2\right) = \frac{T_2 - T_1}{\lambda\left(T_2^{\beta} - T_1^{\beta}\right)}$$

$$MTBF_2\left(0, 199\right) = \frac{199 - 0}{0.000345\left(199^{1.9770}\right)}$$

$$= 16.45\, days$$

[6.13]

Expected number of failures E (N):

$$E_p\left(N\left(T_2\right) - N\left(T_1\right)\right) = \lambda\left(T_2^{\beta} - T^{\beta}_1\right)$$

$$E_p\left(N\left(199\right) - N\left(0\right)\right) = 0.000345\left(199^{1.9770}\right)$$

$$= 12.1\, failures$$

[6.14]

The observed inter-arrival time for the electrical subsystem was 48 days from the day of the last recorded failure, which means the electrical subsystem lasted 30.56 days longer than the prediction. The observed number of failures E (N) = 11 versus the predicted E (N) =12.13. The reliability of the subsystem at the observed time to failure is 48.2%. The researcher conducted the predictions on shorter intervals for each of the subsystems for the expected number of failures. The predictions were compared with the observed values in the same time frame. The results are presented Table 6-4 below.

Table 6-4 : A comparison of the subsystems for the expected and observed number of failures

X_i (days)	Perway		Signal		Electricals	
	N(t)	E(N)	N(t)	E(N)	N(t)	E(N)
7	0	0.024	1	0.682	0	0.0160
14	0	0.069	3	1.623	0	0.0638
28	2	0.1976	5	3.861	0	0.1975
56	2	0.5637	7	9.1858	1	0.9888

6.2 Section summary

This section demonstrated the application of the model to quantify the reliability of railway infrastructure systems. The model was further validated to test for variations and deviation in the predicted values. It can be concluded that it is possible to quantify and predict infrastructure failures in railway systems using a reliability centred approach.

7 Multi-criteria analysis

The aim of the study seeks to quantify the reliability of railway infrastructure systems to assist in the maintenance and management of railway infrastructure assets. Reliability as a performance measure can assist maintenance managers in prioritising infrastructure assets during maintenance interventions on the railway network. In this section, the model will be applied to multiple corridors and the application of the reliability model in maintenance management prioritisation will be demonstrated.

7.1 Application of multi-criteria analysis

Following the data analysis approach given in section 4, two corridors on the central line of the Metrorail network had sufficient data for reliability analysis. The two corridors are the Nyanga-Phillipi corridor and the Langa-Belhar corridor. The reliability predictions were conducted using failure data between the same periods (Jan-Dec 2015). The predictions were conducted from the day of the last observed failure for each subsystem. A summary of the results from the statistical analysis is given in the Appendix. The reliability performance of the infrastructure system for the selected corridors is shown in Figure 7-1. From the figure the Langa-Belhar corridor shows better reliability performance over time, implying that the Nyanga-Phillipi corridor requires prioritisation in order to improve its reliability performance. These results do not show which prioritisation of the subsystems should occur to holistically improve the reliability of the infrastructure system on the network. Studying the reliability performance of the infrastructure subsystems across the board will provide more insight on the prioritisation required to improve the reliability.

85
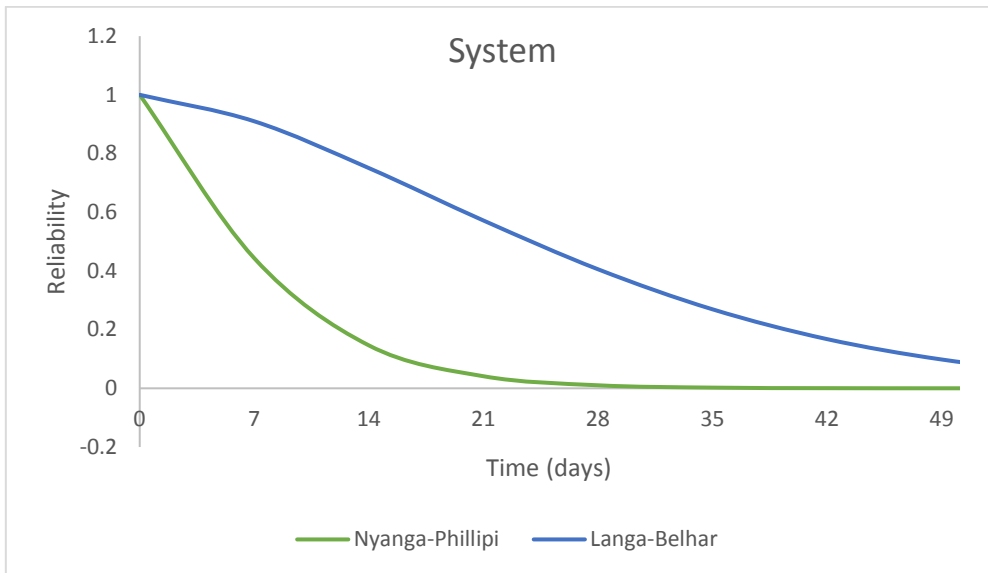
Figure 7-1 : Reliability performance for the Nyanga-Phillipi and Langa-Belhar corridors

Figure 7-2 and Figure 7-3 show the reliability performance of the subsystems for the Langa-Belhar and Nyanga-Phillipi corridors respectively. For the Langa-Belhar corridor, it can be seen from Figure 7-2 that the poor reliability performance of the perway subsystem has the governing criticality that influences the performance of the infrastructure system on that corridor. For the Nyanga-Phillipi corridor shown in Figure 7-3, the signalling subsystem has the governing criticality on that corridor. These results show the subsystems that require prioritisation for each individual corridor.
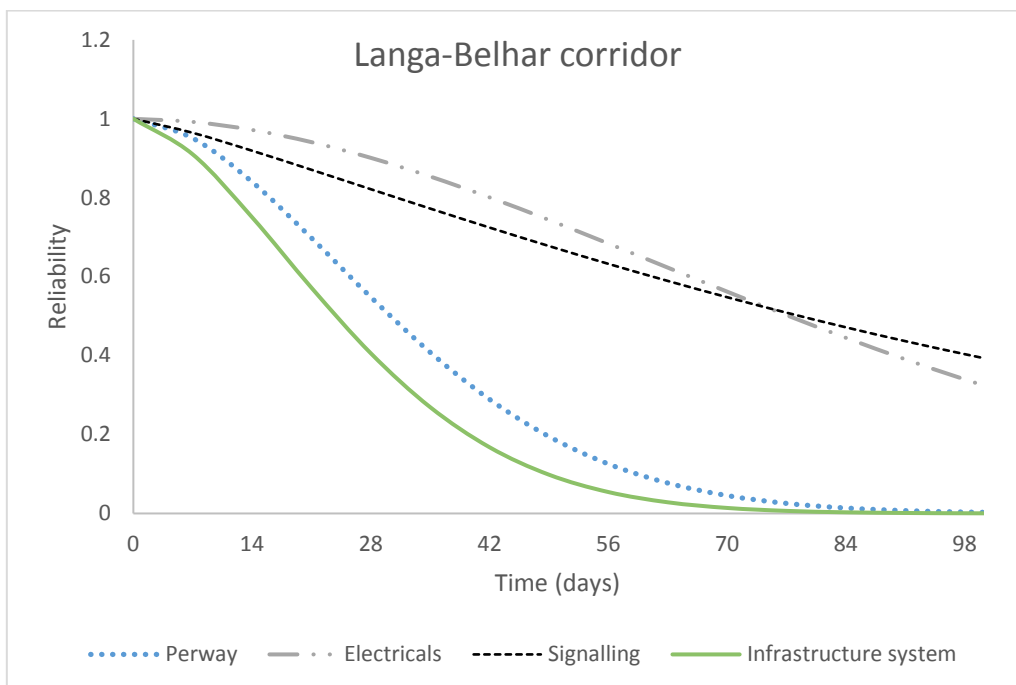
Figure 7-2 : Reliability performance of the Langa-Belhar corridor

86
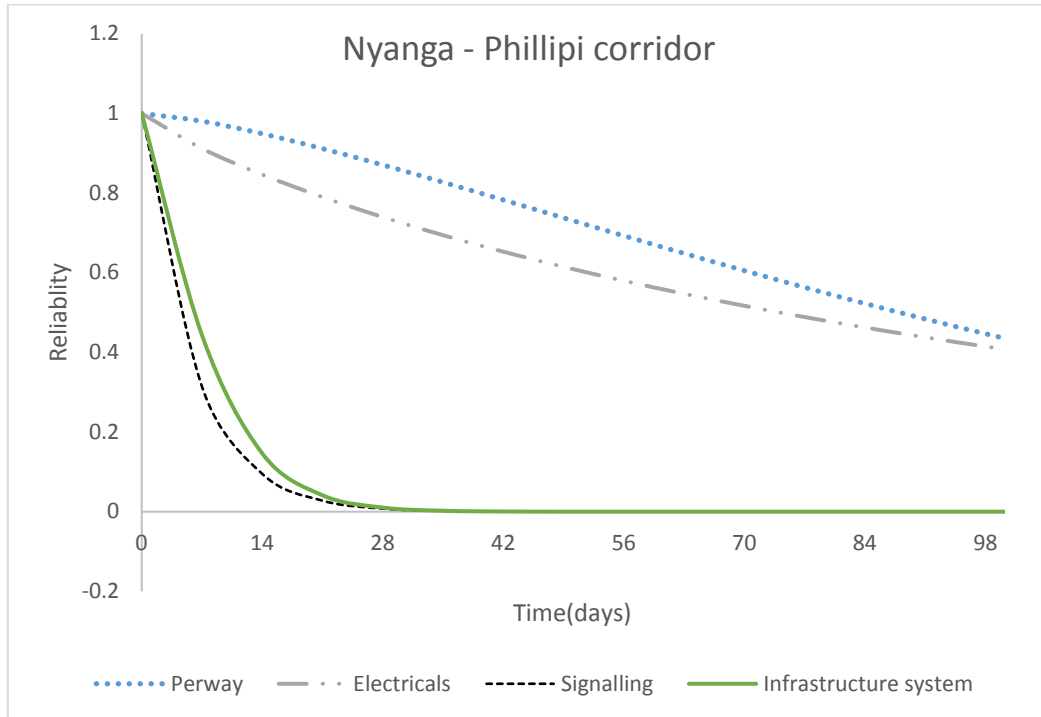
Figure 7-3 : Reliability performance of the Nyanga-Phillipi corridor

For maintenance planning on a large network, the reliability model can be applied to assist in decision-making for prioritising and selecting the best intervention methods on the two corridors that will improve the reliability performance of the railway network. Figure 7-4 and Figure 7-5 shows that the performance of the Langa-Belhar corridor presents better reliability performance for the signalling and electrical subsystems. The outcome of this prediction means that the Nyanga-Phillipi corridor must be prioritised for maintenance for both the signalling and electrical subsystems to improve system performance. The observed time to first failure for the Nyanga-Phillipi signalling subsystem was 4 days, this reflected the predicted value over the same period of 4 days. The predicted time to first failure for the Nyanga-Phillipi electrical subsystem was 102.4 days against an observed value of 208 days, which means the first failure was observed 94.2 days later than the predicted value. For longer maintenance windows however the Langa-Belhar corridor must be prioritised for maintenance because after 84 days the rate of reliability degradation for the electrical subsystem on the Langa-Belhar corridor increases in comparison to that of the Nyanga-Phillipi corridor.
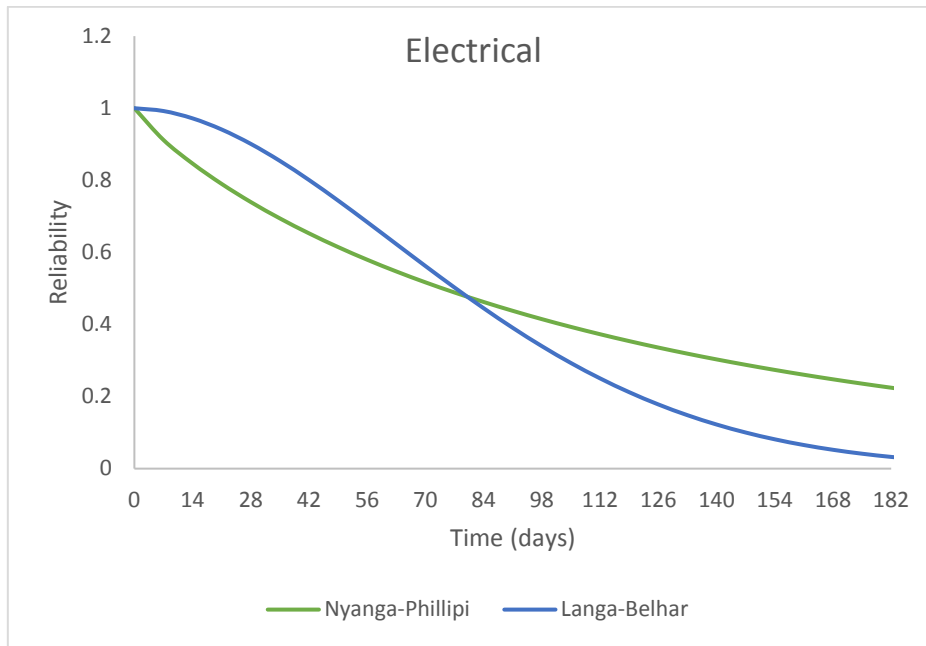
Figure 7-4 : Comparison of the reliability performance of the electrical subsystem
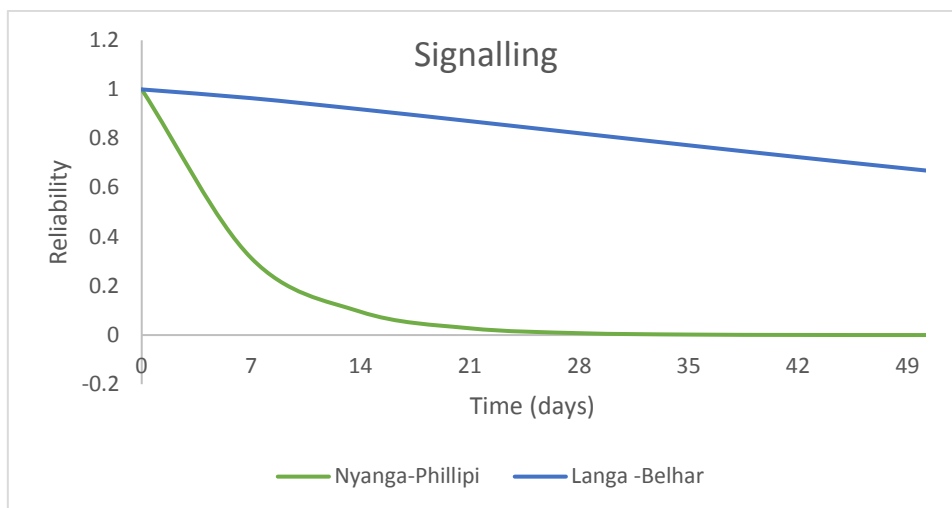
Figure 7-5 : Comparison of the reliability performance of the signalling subsystem

Figure 7-6 for the perway subsystem performance shows reliability performances that contrast to that of the electrical and signalling. Instead, the reliability performance of the perway system for the Nyanga-Phillipi corridor registers high-reliability performance over the same period. This result shows that for the perway subsystem, priority should be given to the Langa-Belhar corridor to maintain acceptable levels of reliability performance. The predicted time to the first failure for the perway subsystem on the Langa-Belhar corridor was 10.61 days against an observed value of 4 days. This means for the Langa-Belhar corridor the time to first failure was recorded 6.61 days earlier than the predicted value.
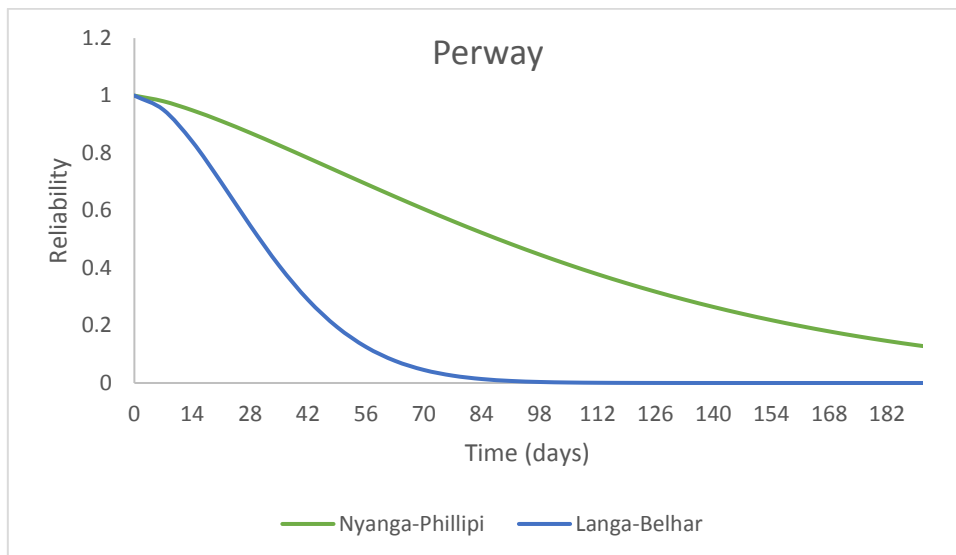
Figure 7-6 : Comparison of the reliability performance of the perway subsystem

7.2 Section summary

This section looked at the application of the reliability model in a multi-criteria analysis to establish the appropriate maintenance prioritisation strategies on a railway network. The model was applied to two corridors and the performance of the corridors was evaluated to establish appropriate maintenance interventions.

8 Discussion of results

From the preceding section, it is evident that the reliability modelling approach given for railway infrastructure systems can assist in maintenance prioritisation by highlighting sections/lines and routes that require attention based on the reliability performance of the infrastructure assets. The study identified that in railway infrastructure environments, two factors influence infrastructure quality. The ability to continuously measure infrastructure quality over time and the ability to employ the necessary measures to restore infrastructure quality suppose it falls below acceptable levels. This section discusses the results of the reliability model which quantify infrastructure quality along with their implication on the asset management strategy to restore infrastructure quality to acceptable levels.

8.1 Reliability as an infrastructure quality measure

The asset failure data collected on the infrastructure network was utilised to generate useful information for decision-making. This information identified the critical subsystems which impact service performance highlighting the asset groups with the highest unreliability. The reliability model predicted the reliability performance of the infrastructure assets over time based on historical asset failure data. These predictions measure how the infrastructure quality of the subsystems evolves on the operational routes from a reliability perspective. The predictions assume that if all managerial and operational decisions remain constant then the system is likely to perform according to the behaviour modelled using the historical asset failure data.

To support primary decisions in the maintenance and renewal of infrastructure systems spread over wide geographic areas, the asset information and performance data must be synthesised into information that can be useful to make informed decisions. Figure 8-1 shows a summary of the results from the multi-criteria analysis for the two routes. From these results at operational route level, the Nyanga-Phillipi line exhibits low reliability performance as compared with the Langa-Belhar line. In addition, the results from the analysis show that the critical subsystems governing the reliability performance of each line is the signalling and perway subsystems for the Nyanga-Phillipi and Langa-Belhar lines respectively. Using the information produced by the proposed modelling framework, all the potential asset management decisions are incorporated, allowing policies and regulations to be formulated that deliver the required performance level of the infrastructure assets on the railway network. From the summary of results in Figure 8-1 the electrical and signalling subsystem of the Nyanga-Phillipi line should have maintenance resources prioritised whereas for the Langa- Belhar line the priority asset group for maintenance is the perway subsystem.

Metrorail Western Cape	Nyanga - Phillipi		Langa - Belhar	
Operational Route	❌			
Criticality	❌	Electrical		Electrical
		Signalling		Signalling
		Perway	❌	Perway
Priority	❌	Electrical		Electrical
	❌	Signalling		Signalling
		Perway	❌	Perway

Figure 8-1 : Summary of multi-criteria analysis

The reliability-based approach quantified the variations that arise at the subsystem interfaces and identified the effects of various intervention strategies related to improving the reliability performance of the railway infrastructure assets. Results from the Pareto analysis seen in Figure 8-2 show the type of infrastructure component and its contribution to infrastructure system downtime as recorded by the number of failures. From the figure, points-and-crossings failure mode demonstrate a high frequency of occurrence highlighting the impact of the signalling subsystem on the reliability performance of the infrastructure system. In addition, the block joint and defective rail failure modes register high frequency of occurrence highlighting the impact of the perway subsystem on the reliability performance of the infrastructure system. Overhead Track Equipment (OHTE) and cable related failure modes of the electrical subsystem although registering a relatively low frequency of occurrence significantly impact the reliability performance of the infrastructure system.

The criticality ranking of the failure modes is summarised in the Appendix from the results of the FMECA study. From the FMECA study, points and crossings and interlocking failure modes ranked intolerable on the criticality scale. The effect of these failures is severe causing on-track machine failures and loss in detection between interlocking components and point to point machines of the signalling subsystem. The effect of failures in the perway subsystem is observed by faulty track circuits, derailments and burnt out catenary. Failure modes related to the perway subsystem like faulty block joints and defective rails caused most track circuit related failures in the signalling subsystem. Studying the failure cause variation in the railway infrastructure system reveals that low-frequency events that have high impact are inherently difficult to predict. This was observed with the electrical and perway subsystem which registered low failure incidences as compared with the signalling subsystem. On the other hand, high-frequency low-impact events are constantly active in the system and can be predicted easily. This was observed on the signalling

subsystem which showed relatively high rates of failure occurrence when compared with the other subsystems.
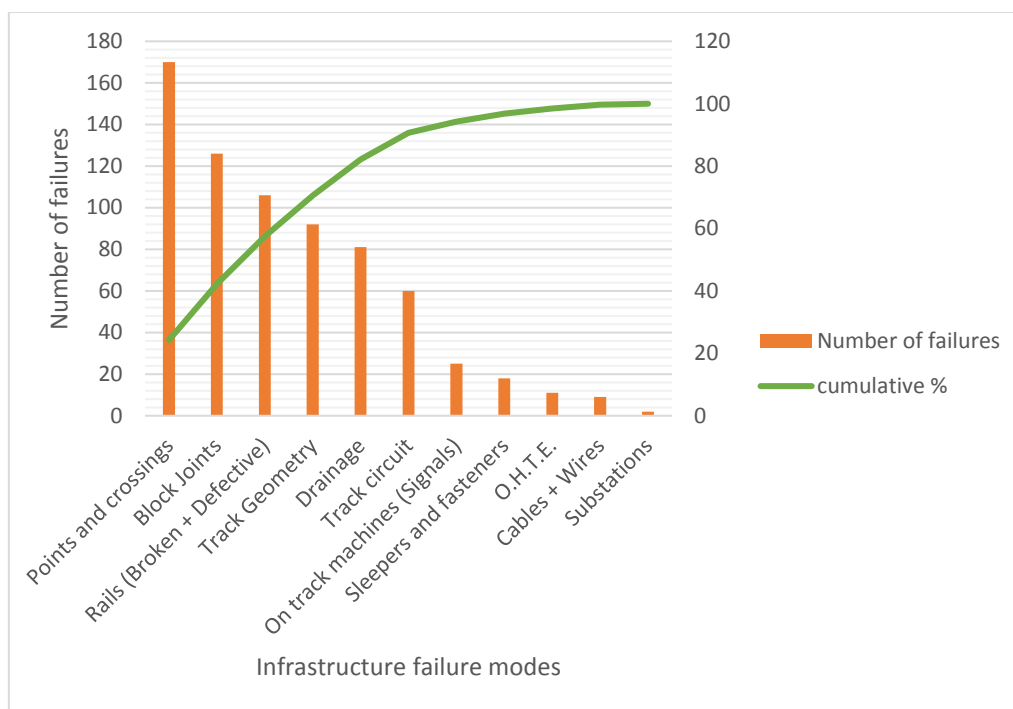
Figure 8-2 : Pareto analysis for failure modes and frequency of failure.

8.2 Reliability-based infrastructure asset management

The researcher studied annual failure records from the IMS together with the results from the reliability analysis. A graphical presentation of the annual contribution of each infrastructure subsystem to train cancellations and delays on the Metrorail network is given in Figure 8-3 and Figure 8-4 below. The trend shows that the signalling-related incidents contribute significantly to train delays as compared with the other subsystems. However, the electrical subsystem contributes more to train cancellations in comparison with other infrastructure subsystems. In addition the results from the FMECA analysis exhibit varying relationships between the failure modes of the different infrastructure subsystems. Despite the relatively low failure incidences reported for the perway subsystem on the railway network, a significant number of signalling failure modes were caused by perway related incidences. This can be attributed to the fact that the system utilisation information of the perway subsystem recorded a corrective and time based maintenance strategy. Allocation of maintenance resources using this strategy does not necessarily follow or respond to the condition of the asset but instead follows consistent interventions guided by manuals or knowledge of local maintenance experts. These "blind" periodic interventions have devastating effects on the performance of other infrastructure subsystems as seen by the severe impact of the perway subsystem on the performance of the signalling subsystem. Failing to respond to this reality by adapting policies based on the operating condition of the asset means railway infrastructure managers are likely to expend resources

92

inefficiently. To improve the infrastructure system therefore based on these outcomes means supporting maintenance policies that emphasise spending more productive hours on infrastructure assets i.e. condition and reliability-based maintenance, than policies based on the operating time of the components i.e. corrective and time-based maintenance. A holistic reliability-based integrated maintenance planning approach based on system status compliments preventative and condition-based maintenance to support overall system improvement. From a reliability-based perspective the results recommend that focusing on high-frequency and low consequence events (incidences) can yield as much benefit to infrastructure reliability performance as focusing on low frequency and high-consequence events.
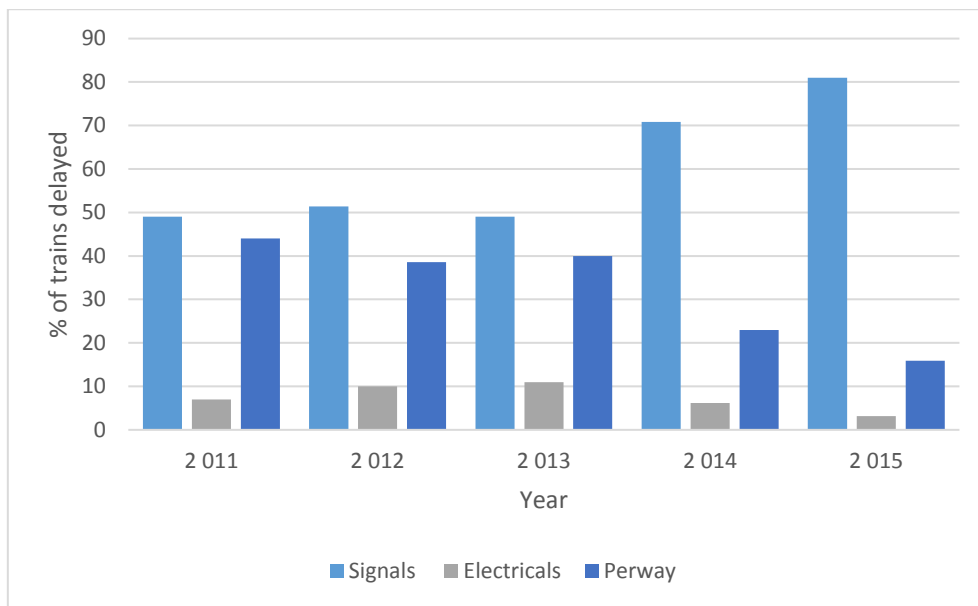

Figure 8-3 : The impact of the different infrastructure subsystems failures to train delays
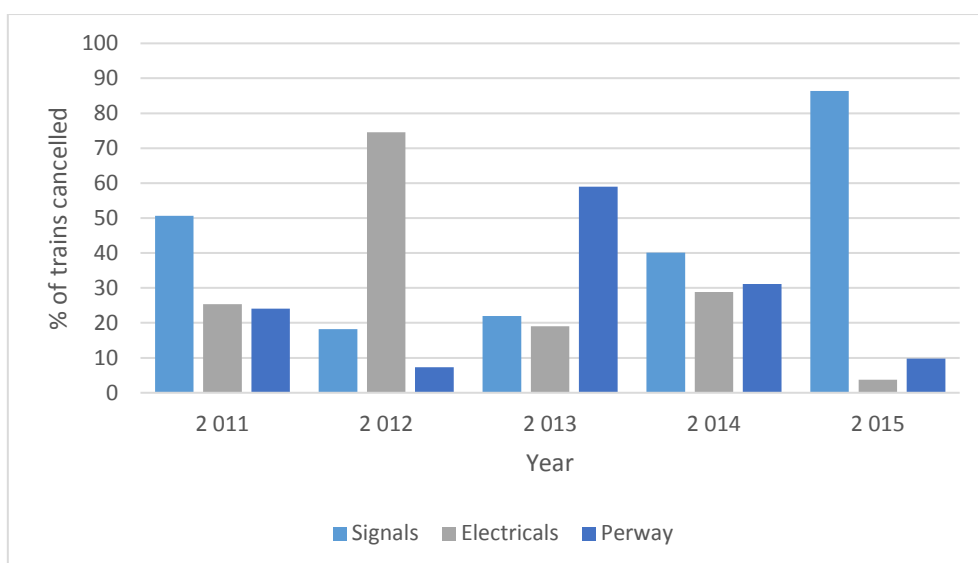

Figure 8-4 : The impact of the different infrastructure subsystems to train cancellations

93

8.3 Research findings

To successfully benefit from a holistic approach to infrastructure asset management presented in the study, the core building blocks that ensure the sustainable application of reliability analysis to improve the maintenance and management of railway infrastructure assets must be identified. In addition, various limitations need to be overcome to effectively develop infrastructure management systems that utilise a reliability-based integrated approach to railway infrastructure maintenance and management. The successful application of the reliability modelling framework presented in this study relies on the availability of a common data structure which is coherent and accessible across the different functions. Additionally, the development of reliable asset degradation models relies on good quality data based on the operational history and condition of assets to achieve sustainable maintenance improvements. Good quality data enables a 'dynamic' identification of priority areas, which allows early detection and prevention of unexpected failures, thus increasing the availability, reliability and the safety of the railway infrastructure system.

The performance of railway organisations is governed by the ability to form a consistent, integrated, and evidenced based approach in the maintenance and management of assets in the medium to long term. To achieve this in railway organisations like PRASA is a challenge because of the separate siloed processes for long-term demand forecasting, asset enhancement planning, and maintenance planning activities. In addition, maintenance intervals for infrastructure systems are determined 'statistically', based on operating time or on the amount of productive hours spent on the infrastructure asset. These intervals are derived from previous experiences or from specifications made by the infrastructure managers based on the life of components involved. To transform this requires re-engineering the strategic asset planning processes to enable the analysis and forecasting of asset conditions and degradation patterns which can be used to develop integrated short and long term asset replacement and management strategies. An extension of this re-engineering process leverages on new technologies to improve monitoring, modelling, and forecasting tools that consolidate the current infrastructure asset management processes in the railway industry.

Although whole life and whole system thinking is difficult to initiate in the short term due to various resource constraints, railway organisations need to actively promote the right values and behaviours to support a holistic approach to asset management. Part of this requires organising around a common asset management strategy and having the right organisational and governance structure that cuts across functions. To deliver a reliable railway infrastructure system a multi-disciplinary and function based thinking approach is required which promotes partnerships to develop solutions that meet the internal needs by building new internal capabilities and competencies.

8.4 Limitations

Access to accurate information supports new processes and ways of thinking and is a requirement for the successful application of a holistic reliability-based approach to infrastructure asset management. In addition, infrastructure performance can be considerably improved if the Information Management Systems are populated with accurate failure data that correctly references failure causes for the different assets in the registry. During the failure analysis, the root cause triggering certain events in some datasets could not be determined. Some failure records studied by the researcher indicated causes that are likely not to be accurate. The root cause in some cases was hard to tell from a single instance, which suggests that further checks were required. The data, however, was detailed with regards to components and functionality but did not concisely define and describe all the events that led to failure. During the failure analysis for the reliability model the researcher concluded that causes given just to complete the data may be misleading, hence the necessity of filling in all fields was not overemphasised. It becomes, therefore, essential for railway organisations to have a technological infrastructure which supports the collecting, organising and managing of the correct data.

8.5 Section summary

The reliability model presented in the study quantifies the reliability performance of the infrastructure system by linking failures, asset data, and the utilisation rate of the railway infrastructure assets. The linking of all infrastructure asset failures assists in identifying complex relationships between the infrastructure subsystems. In this section, it was demonstrated that knowledge of these relationships can improve the operational reliability of the passenger railway infrastructure system by facilitating informed decision making in maintenance and management activities.

9 Conclusions and recommendations

The aim of the research study was to develop a model to measure the reliability performance of railway infrastructure systems to facilitate integrated maintenance planning in railway infrastructure environments. A systematic analysis to develop a holistic reliability model for railway infrastructure systems to improve railway infrastructure asset management processes has been presented. The model presented in this research is an evidence-based decision-making tool which uses asset failure information to account for the joint dependability attributes that characterise railway infrastructure systems. The model developed in the study was applied to a case study on PRASA`s railway network to support the development of appropriate maintenance strategies to improve infrastructure reliability. The model identified critical infrastructure subsystems that impact the reliability performance of the railway infrastructure systems which enables the strategic alignment of asset management plans for the different subsystems to maintain the railway network at acceptable operating levels. Aligning asset management plans using a reliability-based maintenance and management approach moves away from the silo approach which currently characterises railway infrastructure asset management in the South African passenger railway industry. This enables railway organisations to exploit opportunities that can increase capacity and improve the resilience and reliability of railway infrastructure systems in the short to long term period. The reliability modelling approach presented in the study has the capacity to improve asset performance to meet the increasing demands of service quality and infrastructure reliability in railway environments. It can be concluded that reliability analysis can be utilised to develop an integrated reliability-based approach in the maintenance and management of railway infrastructure assets.

9.1 Summary of findings

Asset information supports the primary decisions and activities related to components covered in an asset management framework. These decisions include the development of informed asset policies and the implementation of asset management plans. To fully realise the benefits of information-based asset management strategies such as reliability analysis requires a significant commitment in aligning planning processes, functional and technical specifications, approvals, installations and commissioning processes. Asset management is multidisciplinary and cross-functional and as such it requires personnel who are open to evidence and have the ability to work

in multidisciplinary teams to integrate and interpret the different factors that influence decision-making in such environments. Furthermore, it was observed that it is important to have tools that capture high-quality asset data to support decision-making which enables efficient asset management strategies in collaborative environments. This requires a diverse mix of practical and thinking skills sustained by knowledge and understanding relevant to the planned intervention processes. This must further be complemented by collaborative behaviour and enhanced mechanisms for automated data capture, collation, and visualisation.

9.2 Recommendations

Asset management is no longer a matter of trading off one asset against the other, but rather a matter of trading off how each asset impacts the performance of the whole system in achieving the highest functional performance in terms of safety, availability, and reliability with least possible costs. Railway infrastructure maintenance interventions need to minimise train disruptions, this requires efficient and effective coordination of maintenance planning activities of the railway infrastructure assets. The current structure around asset management in PRASA has two divisions which are the engineering services and maintenance operations. Each department has its own planning process. To facilitate the practical application of the reliability model presented in this research it is recommended that PRASA Metrorail division adopts an integrated planning process in maintaining and managing railway infrastructure assets. An integrated approach will facilitate collaborative sharing of knowledge for decision-making by considering all aspects of required outcomes, including skills required to evaluate cost and reliability performance trade-offs. In addition, increasing the productive time on infrastructure assets can significantly improve the reliability performance of the railway infrastructure system. This means that an integrated approach to maintenance must have the capacity to consistently evaluate and monitor the implementation of the asset management strategies for continuous reliability improvements. However, support for developing integrated maintenance planning in the South African passenger railway requires an increase in awareness within the leadership structure and willingness across the different functional departments to seek, share, and adopt others' learning.

9.3 Theoretical contributions and future research

The researcher developed a reliability model which supports a holistic approach to evaluate the reliability performance of railway infrastructure assets. The reliability modelling approach presented in this study identified critical failure modes for railway infrastructure systems using a FMECA methodology. In addition, the functional and operational interdependencies in railway infrastructure systems were modelled to accurately quantify the joint dependability attributes that characterise railway infrastructure systems. This sets the basis for the development of rail

infrastructure network models that enable the railway system to be viewed both topographically as a map and topologically as schematic logical views showing how individual assets are connected. Network models provide a geospatial view of the railway network showing the location of assets on the network and the underlying asset information for each infrastructure asset. Rail infrastructure network models can bring together infrastructure data sets describing system-level utilisation and performance, connecting asset management, operations, and maintenance allowing infrastructure managers to understand relationships between assets.

10 References

[1]   DoT, "National Household Travel Survey in South Africa (NHTS)," Pretoria, 2013.

[2]   D. Rama and J. D. Andrews, "A Holistic Approach to Railway Infrastructure Asset Management," vol. 11, pp. 3–16, 2014.

[3]   J. Carretero, J. M. Pérez, F. García-Carballeira, A. Calderón, J. Fernández, J. D. García, A. Lozano, L. Cardona, N. Cotaina, and P. Prete, "Applying RCM in large scale systems: a case study with railway networks," *Reliab. Eng. Syst. Saf.*, vol. 82, no. 3, pp. 257–273, Dec. 2003.

[4]   D. Rama and J. D. Andrews, "Railway infrastructure asset management: the whole-system life cost analysis," *IET Intell. Transp. Syst.*, vol. 10, no. 1, pp. 58–64, 2016.

[5]   M. Macchi, M. Garetti, D. Centrone, L. Fumagalli, and G. Piero Pavirani, "Maintenance management of railway infrastructures based on reliability analysis," *Reliab. Eng. Syst. Saf.*, vol. 104, pp. 71–83, Aug. 2012.

[6]   H. Fukuoka, "Reliability evaluation method for the railway system: A model for complicated dependency," *Q. Rep. RTRI (railw. Tech. Res. Institute)*, vol. 43, no. 4, pp. 192–196, 2002.

[7]   S. M. Famurewa, C. Stenström, M. Asplund, D. Galar, and U. Kumar, "Composite indicator for railway infrastructure management," *J. Mod. Transp.*, vol. 22, no. 4, pp. 214–224, 2014.

[8]   D. Prescott and J. Andrews, "Modelling maintenance in railway infrastructure management," *Proc. - Annu. Reliab. Maintainab. Symp.*, pp. 3–8, 2013.

[9]   J. González, R. Romera, J. Carretero, and J. M. Pérez, "Optimal Railway Infrastructure Maintenance and Repair Policies to Manage Risk Under Uncertainty with Adaptive Control," Madrid, 06–16, 2006.

[10]   P. Mokhtarian, M.-R. Namzi-Rad, T. K. Ho, and T. Suesse, "Bayesian nonparametric reliability analysis for a railway system at component level," in *2013 IEEE International Conference on Intelligent Rail Transportation Proceedings*, 2013, pp. 197–202.

[11]   C. Stenstrom, "Operation and maintenance performance of rail infrastructure: Model and Methods," Luleå University of Technology, 2014.

[12]   P. Prof, J. Jorge, P. Pereira, O. Prof, and P. Manuel, "RAMS analysis of railway track infrastructure (Reliability , Availability , Maintainability , Safety)," Instituto Superior Tecnico, 2008.

[13]   N. Rhayma, P. Bressolette, P. Breul, M. Fogli, and G. Saussine, "Reliability analysis of maintenance operations for railway tracks," *Reliab. Eng. Syst. Saf.*, vol. 114, pp. 12–25, Jun.

2013.

[14] M. D. McNaught, "A Risk-Reliability Comparison Of Track Sections In The Passenger Railway Industry," University of Stellenbosch, 2015.

[15] J. Zhao, a H. C. Chan, and M. P. N. Burrow, "Reliability analysis and maintenance decision for railway sleepers using track condition information," *J. Oper. Res. Soc.*, vol. 58, no. 8, pp. 1047–1055, 2007.

[16] H. Chen and Q. Sun, "Reliability Analysis of Railway Signaling System Based Petri Net," no. Mmat, pp. 71–75, 2012.

[17] S. C. Panja and P. K. Ray, "Reliability Analysis of a 'Point-and-Point Machine' of the Indian Railway Signaling System," *Qual. Reliab. Eng. Int.*, vol. 23, no. November 2006, pp. 517–543, 2007.

[18] D. Rama and J. D. Andrews, "A reliability analysis of railway switches," *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit*, vol. 227, no. 4, pp. 344–363, 2013.

[19] B.-H. Ku and J.-M. Cha, "Reliability assessment of Catenary of Electric railway by using FTA and ETA analysis," in *2011 10th International Conference on Environment and Electrical Engineering*, 2011, pp. 1–4.

[20] B. Ku, J. M. Cha, and H. Kim, "Reliability analysis of catenary of electric railway by using FTA," *Trans. Korean Inst. Electr. Eng.*, vol. 57, no. 11, pp. 1905–1909, 2008.

[21] G. Cosulich, P. Firpo, and S. Savio, "Power electronics reliability impact on service dependability for railway systems: a real case study," *Proc. IEEE Int. Symp. Ind. Electron. ISIE '96*, vol. 2, pp. 996–1001, 1996.

[22] D. J. Pedregal, F. P. García, and F. Schmid, "RCM2 predictive maintenance of railway systems based on unobserved components models," *Reliab. Eng. Syst. Saf.*, vol. 83, no. 1, pp. 103–110, Jan. 2004.

[23] The RAIL Consortium, "Reliability centered maintenance Approach for Infrastructure and Logistics of railway operations," 2000.

[24] Y. M. Jidayi, "Reliability Improvement of Railway Infrastructure," Stellenbosch University, 2015.

[25] P. Daniel, Conradie, C. Fourie, P. Vlok, and N. Treurnicht, "Quantifying System Reliability in Rail Transportation in an Aging Fleet Environment," *South African J. Ind. Eng.*, vol. 26, no. March, p. 128, Aug. 2015.

[26] Network Rail, "Asset Management Strategy." Network Rail, London, p. 48, 2014.

[27] L. Saurabh Kumar, "Reliability analysis and cost modelling of degrading systems," Luleå University of Technology, 2008.

[28] T. Nowakowski, "Problems of Transportation Process Reliability Modelling." Wrocław University of Technology, p. 20, 2004.

[29] A. P. Patra, "Maintenance Decision Support Models for Railway Infrastructure using RAMS & LCC Analyses," Luleå University of Technology, 2009.

[30] J. Andrews and P. Universités, "Maintenance modelling , simulation and performance assessment for railway asset management," Universite De Technologie De Troyes, 2015.

[31] S. C. Panja and P. K. Ray, "Failure mode and effect analysis of Indian railway signalling system," *Int. J. Performability Eng.*, vol. 5, no. 2, pp. 131–142, 2009.

[32] A. Baxter, "Network Rail A Guide to Overhead Electrification," no. February. Network Rail, London, 2015.

[33] V. V Valencia, J. M. Colombi, A. E. Thal, and W. E. Sitzabee, "Asset Management : A Systems Perspective," in *2011 Industrial Engineering Research Conference*, 2011, p. 9.

[34] N. S. Grigg, *Water, Wastewater, and Stormwater Infrastructure Management*. Lewis Publishers, 2003.

[35] G. W. Flintsch and J. W. Bryant, *Asset Management Data Collection Asset Management Data Collection for Supporting Decision Processes*. US Department of Transport, 2006.

[36] International Union of Railways, "Guidelines for the Application of Asset Management in Railway Infrastructure Organisations," Paris, 2010.

[37] J. M. van Noortwijk and D. M. Frangopol, "Two Probabilistic Life-Cycle Maintenance Models for Deteriorating Civil Infrastructures," *Probabilistic Eng. Mech.*, vol. 19, no. 4, pp. 345–359, 2004.

[38] Y. K. Al-Douri, P. Tretten, and R. Karim, "Improvement of railway performance: a study of Swedish railway infrastructure," *J. Mod. Transp.*, vol. 24, no. 1, pp. 22–37, 2016.

[39] M. Rausand and A. Høyland, *System reliability theory : models, statistical methods, and applications*. John Wiley & Sons, Inc, 2004.

[40] G. Muyengwa and Y. N. Marowa, "ANALYZING ADOPTION OF MAINTENANCE STRATEGIES IN MANUFACTURING COMPANIES," in *International Association for Management of Technology*, 2015, p. 25.

[41] T. Åhrén, "Maintenance performance indicators (MPIs) for railway infrastructure : identification and analysis for improvement," *Environ. Eng.*, pp. 1–124, 2008.

[42] A. P. Patra, "RAMS and LCC in rail track maintenance," Luleå University of Technology, 2007.

[43] I. H. Afefy, "Reliability-Centered Maintenance Methodology and Application: A Case Study," vol. 2, pp. 863–873, 2010.

[44]   N. R. Council, *Measuring and Improving Infrastructure Performance*. National Academy of Sciences, 1996.

[45]   EUROPEAN COMMITTEE FOR STANDARDIZATION, *Maintenance - Maintenance Key Indicators*. 2007, p. 8.

[46]   C. Stenström, "Performance Measurement of Railway Infrastructure with Focus on the Swedish Network," 2012.

[47]   P. Brinkman, "Valuing rail infrastructure performance in a multi actor context MSc-Thesis," TU Delft, 2009.

[48]   S. M. Famurewa, M. Asplund, M. Rantatalo, A. Parida, and U. Kumar, "Maintenance analysis for continuous improvement of railway infrastructure performance," *Struct. Infrastruct. Eng.*, vol. 11, no. 7, pp. 957–969, Jun. 2014.

[49]   EUROPEAN COMMITTEE FOR STANDARDIZATION, *EN 50126 - Railway applications - The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS)*. 1999.

[50]   INECO, "GENERAL RAMS PLAN FOR THE RAILWAY LINES AKKO – CARMIEL , HAIFA - BET SHEAN AND HERZELYA-KEFAR SABA." p. 50, 2013.

[51]   FRAZER-NASH Consultancy, "Influence of RAMS in Infrastructure Reliability in Mobility," in *SYSTEMS AND ENGINEERING TECHNOLOGY*, 2015.

[52]   G. VASIĆ, S. INGLETON, A. SCHÖBEL, B. PAULSSON, and M. ROBINSON, "Development of the Future Rail Freight System to Reduce the Occurrences and Impact of Derailment," in *Scientific expert conference on railways RAILCON12*, 2014.

[53]   S. M. Famurewa, "Maintenance Analysis and Modelling for Enhanced Railway Infrastructure Capacity," Lulea University of Technology, 2014.

[54]   F. . Restel, "IMPACT OF INFRASTRUCTURE TYPE ON RELIABILITY OF RAILWAY TRANSPORTATION SYSTEM," *J. KONBiN*, vol. 1, no. 25, pp. 59–74, 2013.

[55]   K. N. Fleming, A. Mosleh, and R. Kenneth Deremer, "A systematic procedure for the incorporation of common cause events into risk and reliability models," *Nucl. Eng. Des.*, vol. 93, no. 2–3, pp. 245–273, May 1986.

[56]   A. Villemeur, *Reliability, availability, maintainability, and safety assessment*. J. Wiley, 1992.

[57]   R. Valenzuela, "Compact Reliability and Maintenance Modeling of Complex Repairable Systems," Georgia Institute of Technology, 2014.

[58]   P. Pederson, D. Dudenhoeffer, S. Hartley, and M. Permann, "Critical Infrastructure Interdependency Modeling: A Survey of U.S. and International Research," 2006.

[59]   A. Schöbel and T. Maly, "Operational fault states in railways," *Eur. Transp. Res. Rev.*, vol. 4, no. 2, pp. 107–113, Jun. 2012.

[60]   T. R. Browning, "Applying the Design Structure Matrix to System Decomposition and Integration Problems: A Review and New Directions," *IEEE Trans. Eng. Manag.*, vol. 48, no. 3, 2001.

[61]   A. N. Singh, M. P. Gupta, and A. Ojha, "Identifying critical infrastructure sectors and their dependencies: An Indian scenario," *Int. J. Crit. Infrastruct. Prot.*, vol. 7, no. 2, pp. 71–85, 2014.

[62]   W.-X. Wang, R.-J. Guo, and H.-L. Zheng, "Application of ISM in service quality analysis of railway passenger train," in *2010 2nd IEEE International Conference on Information Management and Engineering*, 2010, pp. 469–472.

[63]   R. Attri, N. Dev, and V. Sharma, "Interpretive Structural Modelling (ISM) approach: An Overview," *Res. J. Manag. Sci. Res. J. Manag. Sci*, vol. 2, no. 2, pp. 2319–1171, 2013.

[64]   E. Human, "What is Reliability Engineering," *Asset Manag. Reliab. Eng.*, no. June, pp. 3–5, 2012.

[65]   E. A. Elsayed, *Reliability engineering*. John Wiley & Sons, 2012.

[66]   M. Modarres, M. P. Kaminskiy, and V. Krivtsov, *Reliability Engineering and Risk Analysis a Practical Guide, Second Edition.* CRC Press, 2009.

[67]   E. . Lewis, *Introduction to reliability engineering. E. E. Lewis, Wiley, New York, 1987.*, vol. Second Edi. 1987.

[68]   J. Johansson, H. Hassel, and E. Zio, "Reliability and vulnerability analyses of critical infrastructures: Comparing two approaches in the context of power systems," *Reliab. Eng. Syst. Saf.*, vol. 120, pp. 27–38, Dec. 2013.

[69]   A. C. Marquez, A. S. Heguedas, and B. Iung, "Monte Carlo-based assessment of system availability. A case study for cogeneration plants," *Reliab. Eng. Syst. Saf.*, vol. 88, no. 3, pp. 273–289, 2005.

[70]   H. S. Bø and B. F. Nielsen, "Estimation of Reliability by Monte Carlo Simulations Combined with Optimized Parametric Models," Norwegian University of Science and Technology, 2014.

[71]   L. Ljung, "Black-box models from input-output measurements," in *IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No.01CH 37188)*, vol. 1, pp. 138–146.

[72]   F. De Felice and A. Petrillo, "Methodological approach for performing human reliability and error analysis in railway transportation system," *Int. J. Eng. Technol.*, vol. 3, no. 5, pp. 341–353, 2011.

[73]   C. Esveld, *Modern railway track*. MRT-Productions, 2001.

[74]   S. J. Hassankiadeh, "Failure Analysis of Railway Switches and Crossings for the purpose of Preventive Maintenance," Royal Institute of Technology, 2011.

[75]   A. P. Patra and U. Kumar, "Availability analysis of railway track circuit," in *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 2010, pp. 169–177.

[76]   *Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design*. London: Springer London, 2009.

[77]   S. M. Famurewa, "Increased Railway Infrastructure Capacity Through Improved Maintenance Practices," Lulea University of Technology, 2012.

[78]   A. Singh,Vijay P. Jain,Sharad K.Tyagi, *Risk and Reliability Analysis - A Handbook for Civil and Environmental Engineers*. ASCE, 2007.

[79]   T. Ho, S. Chen, and B. Mao, "Application on Fault Tree Analysis in Railway Power Supply Systems," in *5th International Conference on Traffic and Transportation Studies, 2‑4 August*, 2006.

[80]   J. M. Cha and B. H. Ku, "Reliability assessment of railway power system by using tree architecture," *Trans. Korean Inst. Electr. Eng.*, vol. 59, no. 1, pp. 9–15, 2010.

[81]   O. Basile, P. Dehombreux, and F. Riane, "Identification of reliability models for non repairable and repairable systems with small samples," *Proc. IMS2004 Conf. Adv. Maint. Model. Simul. Intell. Monit. Degrad. Arles*, pp. 1–8, 2004.

[82]   J. W.A. Thompson, "On the Foundations of Reliability," vol. 17, no. 3, pp. 333–339, 2016.

[83]   P. D. T. O'Connor and A. Kleyner, *Practical reliability engineering*. Wiley, 2012.

[84]   W. Q. Meeker and L. A. Escobar, *Statistical methods for reliability data*. Wiley, 1998.

[85]   M. T. Todinov, *Reliability and risk models setting reliability requirements*. 2015.

[86]   Marc Antoni, "THE AGEING OF SIGNALLING EQUIPMENT AND THE IMPACT ON MAINTENANCE STRATEGIES," *R&RATA*, vol. 2, no. 4, pp. 28–37, 2009.

[87]   R. Ahmad, S. Kamaruddin, M. Mokthar, and I. Putra Almanar, "Identifying the Best Fit Failure Distribution and the Parameters of Machine's Component: A New Approach," in *International Conference on Man-Machine Systems*, 2006.

[88]   L. M. Maillart and S. M. Pollock, "The effect of failure-distribution specification-errors on maintenance costs," in *Annual Reliability and Maintainability. Symposium. 1999 Proceedings (Cat. No.99CH36283)*, 1999, pp. 69–77.

[89]   A. H. S. Garmabaki, A. Ahmadi, J. Block, H. Pham, and U. Kumar, "A reliability decision framework for multiple repairable units," *Reliab. Eng. Syst. Saf.*, vol. 150, pp. 78–88, Jun. 2016.

[90] B. H. Lindqvist, "Statistical Modeling and Analysis of Repairable Systems 2 \ Major events " in the history of repairable systems relia-," pp. 1–18, 1997.

[91] Z. Zhang, "Parameter estimation techniques: a tutorial with application to conic fitting," *Image Vis. Comput.*, vol. 15, no. 1, pp. 59–76, 1997.

[92] E. H. Sagvolden, "Statistical analysis of failures and failure propagation in railway track," Norwegian University of Science and Technology, 2013.

[93] R. Levy, "Chapter 4 Parameter estimation," *Introd. to Meta Anal.*, vol. 5, pp. 257–294, 2009.

[94] M. O. Locks, *Reliability, maintainability, and availability assessment*. ASQC Quality Press, 1995.

[95] M. (Mohammad) Modarres, *Risk analysis in engineering : techniques, tools, and trends*. Taylor & Francis, 2006.

[96] Department of Transport South Africa, "National Transport Master Plan (NATMAP) 2050," 2016.

11 Appendices

11.1 Railway infrastructure failure modes

Subsystem	Failure mode	Failure cause	Failure effect	Criticality		
				Frequency	Severity	Criticality
Perway	Faulty block joints	• Wear and tear	• Faulty Track circuit	High	Critical	Intolerable
Electrical	Cable + wires	• Vandalism • Maintenance works • Cable faults	• Overhead power failure • Signal power failures •	Moderate	Critical	Undesirable
Signalling	Interlocking (Crossings)	• Wear and tear • Broken blades	• Faulty signalling	Very High	catastrophic	Intolerable
Signalling	Point to point machines	• Wear and tear • Vandalism • Blown fuses • Faulty micro switch	• On-track machine failures • Loss in detection	Very High	Catastrophic	Intolerable
Signalling	Track circuit	• Faulty block joints • Faulty transmitter • Defective rail bond	• Track circuit failures	Very High	Marginal	Intolerable

| Signalling	On track machines (Signals)	• Track circuit failures • Signal power failures • Faulty fuse holder	• False occupation alarm • Loss of signal • Faulty block signal	Moderate	Critical	Undesirable
Electrical	Substation Power	• Feeder cable failures • Blown fuses	• Feeder cable failures • Low overhead supply	Low	Catastrophic	Undesirable
Perway	Broken rail and defective rails	• Wear and tear • Tonnage • Geometric misalignments • Rail to rail bond off	• Loss in signal • Derailments • Short circuit on track circuit. • Burnt out catenary due to short circuit	Moderate	Catastrophic	Intolerable
Perway	Drainage (Track substructure)	• Settlements • Voiding	• Faulty track circuit	Moderate	Critical	Undesirable

107

11.2 Infrastructure dependency matrix

		PERWAY		ELECTRICALS				SIGNALLING				
		S1	S2	OHTE	SUB11kv	SUB3kv	TRANSL	TC	PPM	INTLOCK	SIG	SIGPOW
PERWAY	S1			X			X	X	X	X	X	
	S2			X			X					
ELECTRICALS	OHTE	X	X									
	SUB11kv							X				
	SUB3kv											
	TRANSL											
SIGNALLING	TC	X	X			X						
	PPM											
	INTLOCK	X	X									
	SIG						X					
	SIGPOW						X					

KEY

PERWAY	
Superstructure	S1
Substructure	S2
ELECTRICALS	
OHTE	OHTE
11 kv Substation	SUB11kv
3kv Substation	SUB3kv
3lv/11kv Transmission lines	TRANSL
SIGNALLING	
Track Circuit	TC
Point to Point Machines	PPM
Interlocking	INTLOCK
Signalling	SIG
Signalling power	SIGPOW

11.3 Reliability modelling approach
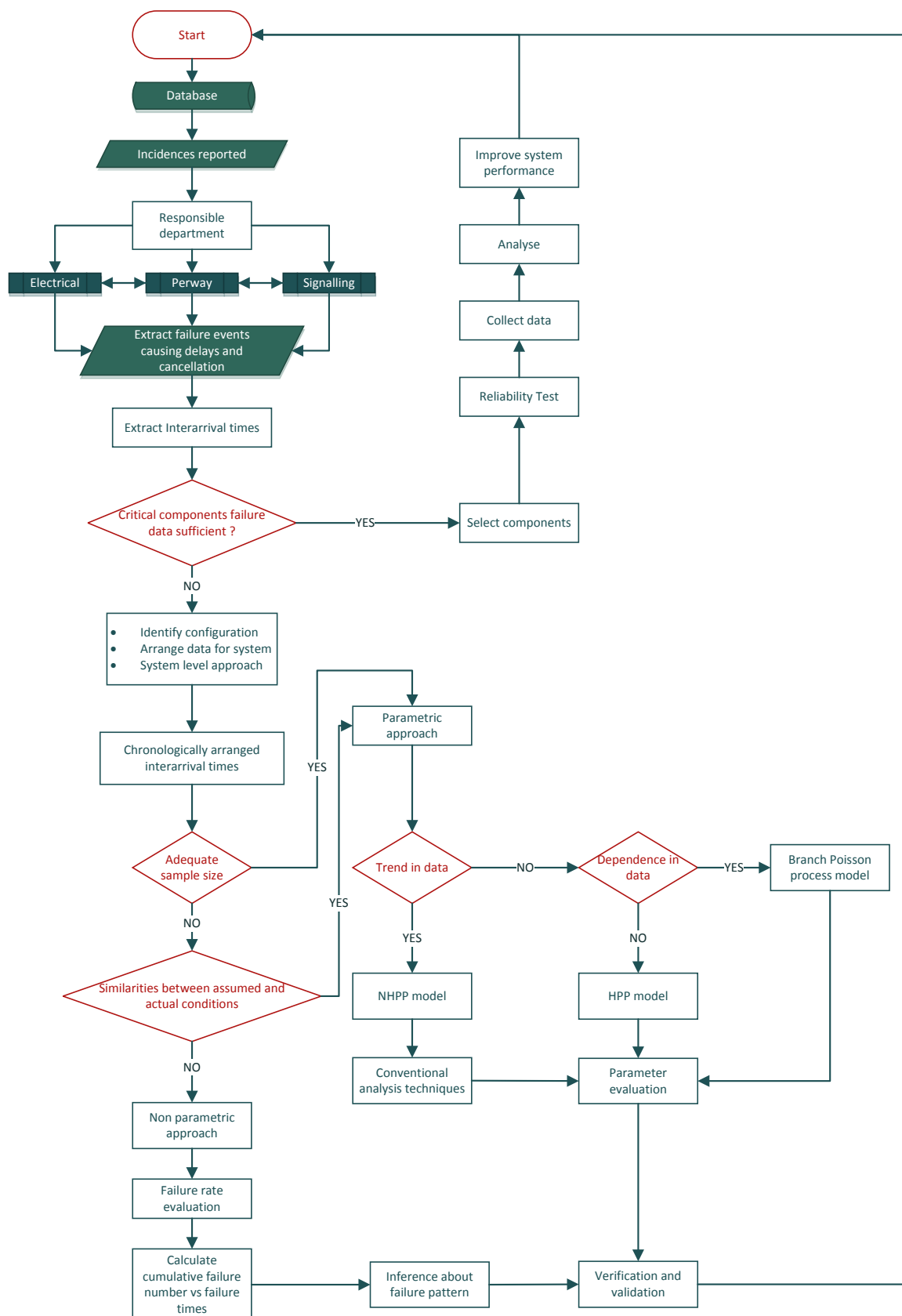
11.4 Langa-Belhar corridor

SIGNALLING	
Interarrival times	N(t)
4	1
5	2
10	3
11	4
12	5
14	6
18	7
19	8
21	9
28	10
32	11
35	12
40	13
41	14
42	15
50	16
52	17
59	18
61	19
62	20
68	21
71	22
74	23
76	24
77	25
84	26
90	27
98	28
104	29
105	30
113	31
114	32
119	33
124	34
126	35
137	36
139	37
144	38
145	39
149	40
150	41
153	42
155	43
158	44
161	45
168	46
170	47
171	48
172	49
174	50
177	51
179	52
180	53

PERWAY	
Interarrival times	N(t)
53	1
67	2
75	3
78	4
79	5
84	6
91	7
130	8
140	9
146	10
148	11
149	12
153	13
161	14
167	15
168	16
179	17

ELECTRICALS	
Interarrival times	N(t)
96	1
168	2
169	3
170	4

Figure 11-1: Arrival times for the Langa-Belhar corridor

110

Table 11-1 : Results from trend test for the Langa-Belhar corridor

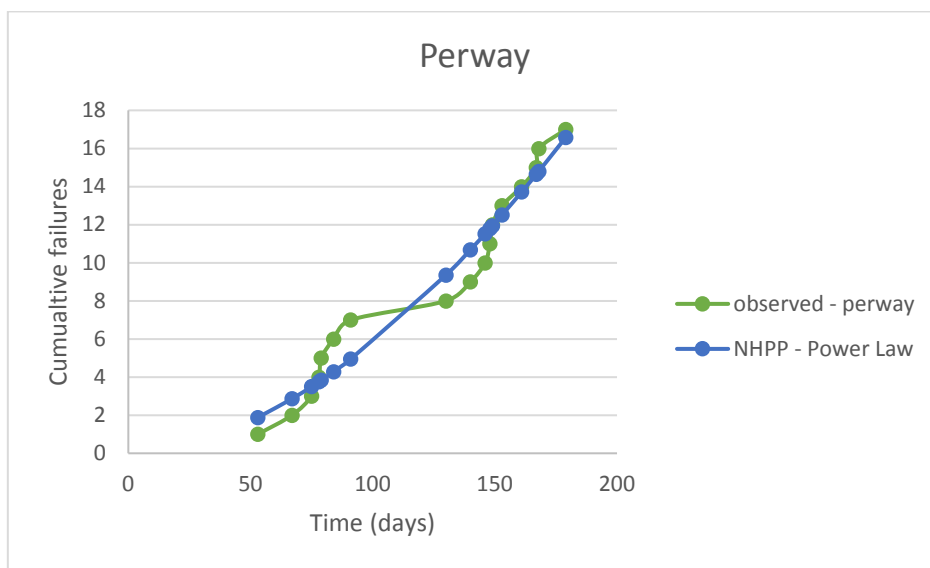| Subsystem	Data points	Laplace Trend Test	LTT interpretation	Lewis Robinson	Model
Perway	17	2.5651	Reliability degradation		NHPP
Signalling	53	0.4520	Non-committal	0.6221	HPP
Electricals	4	2.6796	Reliability degradation		NHPP

Figure 11-2 : Graphical representation of the NHPP power law vs observed values
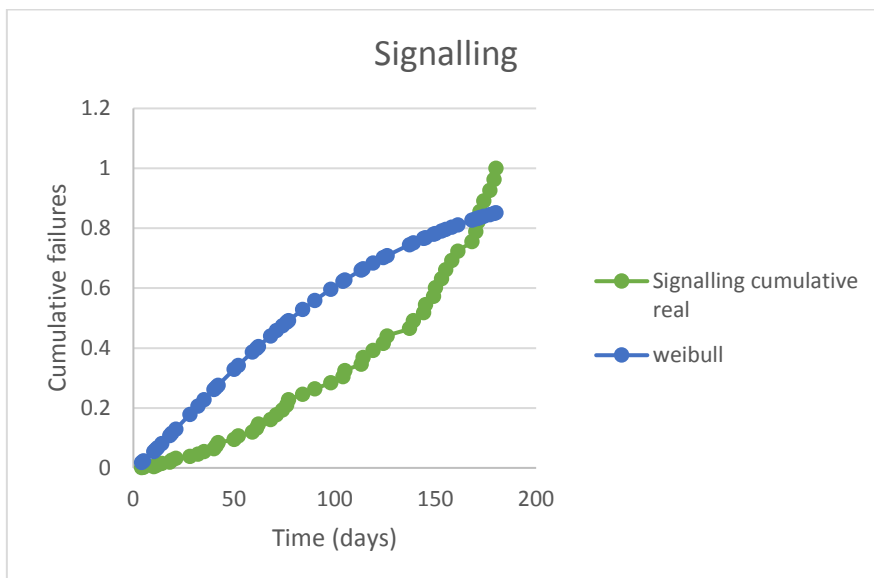

Figure 11-3 : Cumulative distribution function for the Weibull distribution and observed values
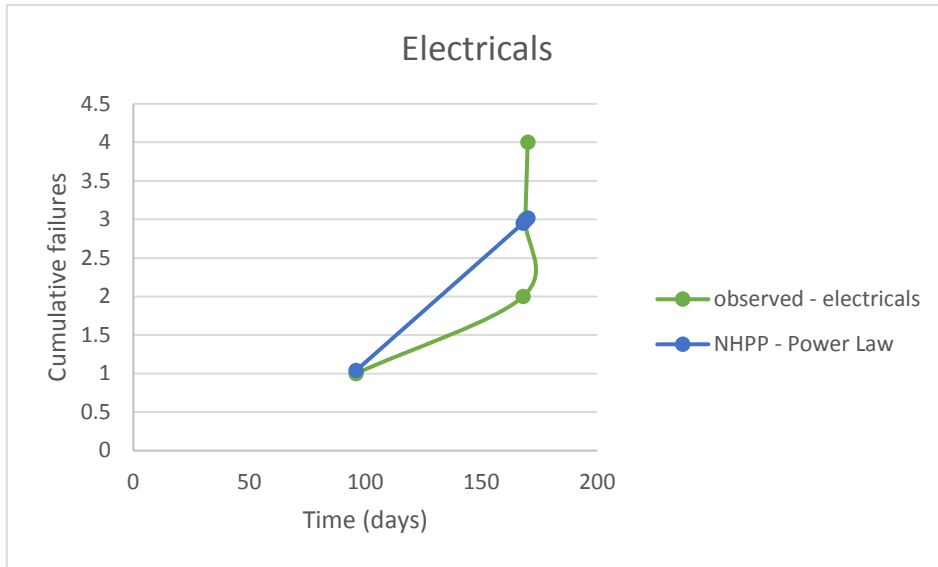
Figure 11-4 : Cumulative distribution function for the Weibull distribution and observed values

Table 11-2 : Parameter estimation results for the Langa-Belhar corridor

Subsystem	Models	K-S Test Result		Parameters		
Perway	NHPP	d_{max} < $d_{critical}$ 0.0187 < 0.1099	Power law	λ = 0.0016	β = 1.7876	
Signalling	Weibull HPP	d_{max} < $d_{critical}$ 0.0103 < 0.0475	Good fit	η = 106.14	β = 1.2207	
Electricals	NHPP	d_{max} < $d_{critical}$ 0.0994 < 0.430	Power law	λ = 0.0002	β = 1.8660	

11.5 Nyanga-Phillipi corridor

SIGNALLING	
Interarrival times	N(t)
4	1
8	2
11	3
13	4
25	5
25	6
32	7
35	8
35	9
47	10
56	11
64	12
71	13
78	14
78	15
88	16
88	17
91	18
93	19
99	20
99	21
99	22
103	23
108	24
127	25
131	26
134	27
135	28
137	29
138	30
138	31
140	32
140	33
141	34
144	35
145	36
149	37
150	38
163	39
165	40
172	41
173	42
174	43
177	44
177	45
178	46
178	47

PERWAY	
Interarrival times	N(t)
27	1
36	2
39	3
112	4
113	5
121	6
155	7
173	8

ELECTRICALS	
Interarrival times	N(t)
13	1
84	2
119	3
138	4

Figure 11-5 : Arrival times for the Nyanga-Phillipi corridor

113

Table 11-3 : Results from the trend test for the Nyanga-Phillipi corridor

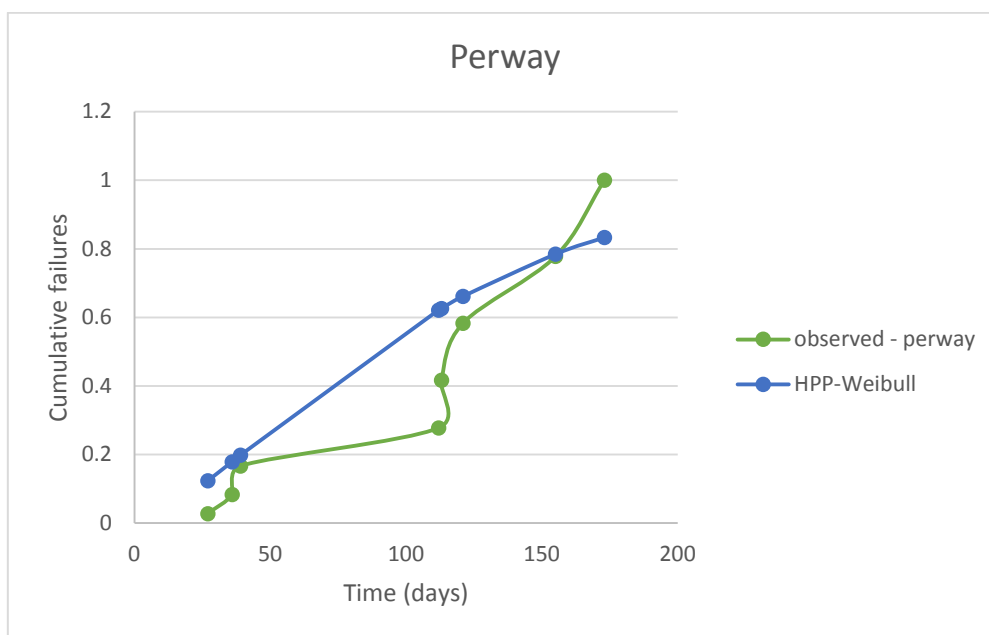| Subsystem	Data points	Laplace Trend Test	LTT interpretation	Lewis Robinson	Model
Perway	8	0.5947	Non-committal	0.5412	HPP
Signalling	47	2.1943	Reliability degradation	2.028	NHPP
Electricals	4	0.9790	Reliability degradation		HPP

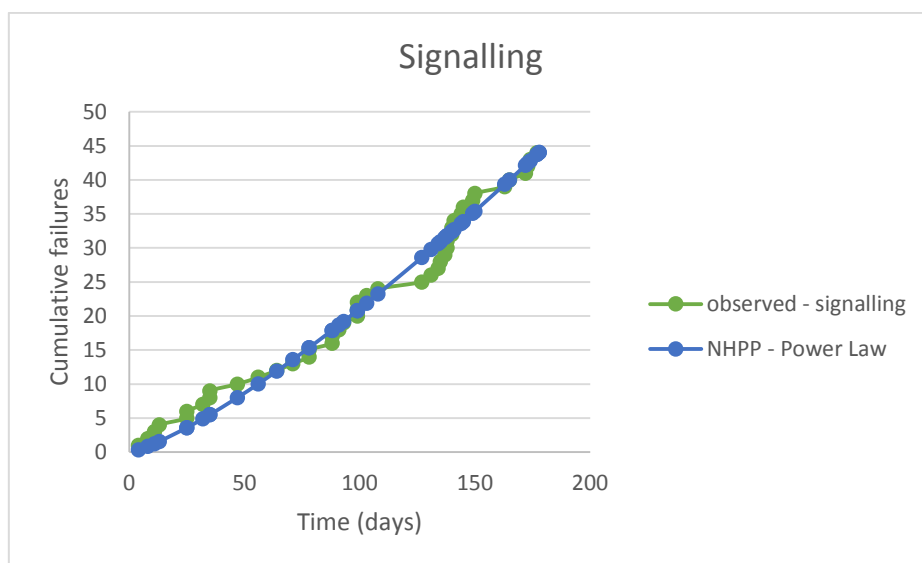Figure 11-6 : Cumulative failures for the observed and Weibull approximations

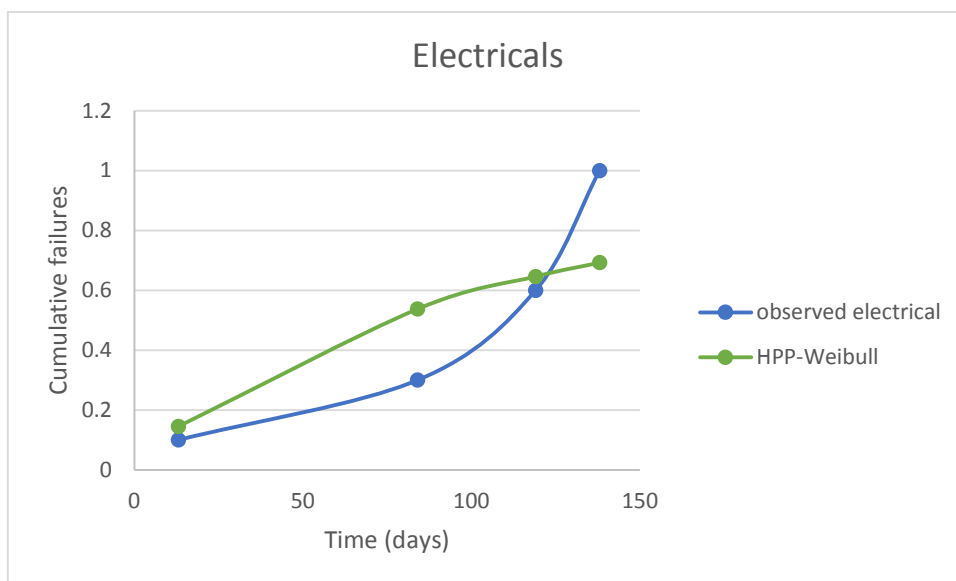Figure 11-7 : Observed vs NHPP power law parameter estimation

Figure 11-8 : Cumulative graph of observed vs Weibull for electrical subsystem

Table 11-4 : Parameter estimation results for the Nyanga-Phillipi corridor

Subsystem	Models	K-S Test Result		Parameters	
Perway	Weibull HPP	d_{max} < $d_{critical}$ 0.200 < 0.6082	Good fit	η = 114.28	β = 1.4047
Signalling	NHPP	d_{max} < $d_{critical}$ 0.0250 < 0.0475	Power law	λ = 0.0582	β = 1.2793
Electricals	Weibull HPP	d_{max} < $d_{critical}$ 0.0502 < 0.430	Good fit	η = 113.80	β = 0.8548

11.6 Map of Metrorail network for the Western Cape region

117