

Development of a maintenance possession scheduler for a railway

Dewald Cillie

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



Study leader: Professor JF Bekker

April 2022

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 sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: April 2022

Abstract

Maintenance of rail infrastructure is an important element in rail operations in order to keep traffic moving. However, maintenance causes infrastructure to be taken out of service, which impacts traffic flow. In this study, the requirements of a maintenance possession scheduler for a South African application was investigated, and a proposed solution was subsequently developed. The main objective of the scheduler was to minimise the deviation of the train service on a subset of rail infrastructure while ensuring that the required maintenance is done.

To achieve this, a literature study was done on a number of themes, which include an overview of the local railway operator with a look at the role of industrial engineering as a function in the railway operator business, railway infrastructure and operations, planning of railway operations, and maintenance in the context of rail operations. The topic of possession scheduling was then studied; the previous themes helped the researcher to learn the bigger picture while understanding possession scheduling is critical for this study. Past and recent works were studied and research areas and trends were synthesised, including time span of possession scheduling in optimisation models, and whether it was done on microscopic, mesoscopic or macroscopic level. The various optimisation objectives formulated by researchers were also noted, among other subthemes.

An application case was identified as the railway infrastructure between Bellville and Wellington in the Western Cape province of South Africa. A novel mixed-integer linear programming model was formulated for this case and implemented in Cplex, after which it was validated. The model can do possession scheduling for 24 hours on a microscopic level. Finally, several experiments were conducted to investigate the performance and results of the model. It was found that the model delivered optimal results in less than

eight minutes, which makes it a feasible maintenance possession scheduler for day-to-day work in the immediate planning horizon.

Opsomming

Instandhouding van spoorinfrastruktuur is 'n belangrike element in spoorweg-operasies ten einde verkeervloei te verseker. Instandhouding veroorsaak egter dat infrastruktuur uit diens geneem word wat verkeer weer belemmer. In hierdie studie was die vereistes van 'n besitskedgeleerder vir instandhouding vir 'n Suid-Afrikaanse toepassing ondersoek, en die voorgestelde oplossing was daarna ontwikkel. Die hoofdoelwit van die skedgeleerder was om die afwykings van die trein diens te minimeer op 'n gedeelte van spoorinfrastruktuur terwyl verseker word dat die nodige instandhouding gedoen word.

Om dit te bereik is 'n literatuurstudie op 'n aantal temas gedoen. Dit sluit in 'n oorsig van die plaaslike spoorwegoperateur en die rol wat bedryfsingenieurswese as funksie daarin vervul, spoorweginfrastruktuur en operasies, en instandhouding in die konteks van spoorwegoperasies. Die onderwerp van besitskedgeleding was daarna bestudeer; die vorige temas het die navorser gehelp om die groter prentjie te verstaan, terwyl die studie van besitskedgeleding kritiek was vir hierdie studie. Navorsingswerk uit die verlede asook onlangse werk was bestudeer en navorsingsareas en tendense is deur sintese bepaal. Dit sluit in die tydsduur van besitskedgeleding in optimeringsmodelle en of dit op mikro-, meso- of makroskopiese vlak gedoen word. Die verskillende optimeringsdoelwitte wat navorsers formuleer het is ook waargeneem, asook met ander subtemas.

'n Gevallestudie vir toepassing van 'n besitskedgeleerder vir instandhouding is identifiseer as die spoorweginfrastruktuur tussen Bellville en Wellington in die Wes-Kaap provinsie van Suid-Afrika. 'n Nuwe gemengde heeltal-lineêre programmeringmodel was geformuleer vir hierdie gevallestudie en in Cplex implementeer, waarna dit gevalideer is. Die model kan besitskedgeleding vir 24 uur doen op mikrovlak. Verskeie eksperimente is uiteindelik uitgevoer

om die prestasie en resultate van die model waar te neem. Dit is bevind dat die model optimale resultate in minder as agt minute kon lewer, wat dit 'n aanvaarbare instandhouding besitskedgeerder maak vir dag-tot-dag werk in die nabye beplanningshorison.

Contents

Declaration	i
Abstract	ii
Opsomming	iv
List of Figures	xiii
List of Tables	xiv
Nomenclature	xviii
1 Introduction	1
1.1 Background	1
1.2 Problem Description	3
1.3 Research Aim	5
1.4 Scope	5
1.5 Research Objectives	6
1.6 Research Methodology	6
1.7 Deliverables	7
1.8 Thesis Organisation	7
2 Literature Study	8
2.1 Overview of Transnet	8
2.1.1 The Transition from Trailblazing Railways to an Unified Govern- ment Railway	8
2.1.2 The Transition from a State Department to a Commercial Company	10
2.1.3 Transnet Limited a State Owned Company	13

CONTENTS

2.1.4	The Role of Industrial Engineering in Transnet Freight Rail . . .	17
2.2	Railway Infrastructure and Operations	26
2.2.1	Fixed Railway Infrastructure	26
2.2.1.1	Track Structure	26
2.2.1.2	Components of Forces Acting on a Track Structure . .	26
2.2.1.3	Main Track Components and Functions	30
2.2.1.4	Track Geometry	33
2.2.1.5	Electrical System	34
2.2.1.6	Overhead Track Equipment Geometry	39
2.2.1.7	Signalling System	39
2.2.1.8	Condition Assessment Systems	54
2.2.1.9	Railway Telecommunication	55
2.2.2	Railway Operations	55
2.2.2.1	Macro Service Planning	56
2.2.2.2	Characteristics of Rail Transportation Demand	58
2.2.2.3	Transportation Requirements	60
2.2.2.4	Rail Traffic Categories	60
2.2.2.5	Railway Operating Strategies	63
2.2.2.6	Train Diagrams	65
2.2.2.7	Rail Capacity	66
2.2.2.8	Train Service Scheduling	68
2.3	Maintenance	69
2.3.1	Physical Asset Management	69
2.3.1.1	Asset Lifetime	69
2.3.1.2	The Objective of Physical Asset Management	71
2.3.1.3	Principles of Physical Asset Management	73
2.3.1.4	Reliability	75
2.3.2	Maintenance Introduction	77
2.3.3	Maintenance Tactics	78
2.3.3.1	Corrective Maintenance	80
2.3.3.2	Time-Based Maintenance	82
2.3.3.3	Condition-Based Maintenance	82
2.3.3.4	Proactive Maintenance	84

CONTENTS

2.3.4	The Maintenance Process	84
2.3.5	Reliability Centred Maintenance	84
2.3.6	Total Productive Maintenance	89
2.3.7	Failure Analysis	92
2.3.7.1	Failure Modes, Effects and Criticality Analysis	94
2.3.7.2	Root Cause Analysis	95
2.4	Maintenance of Railway Infrastructure	97
2.4.1	Track Deterioration	99
2.4.2	Asset Lifecycle Process of Railway Infrastructure	100
2.4.3	Maintenance Process of Railway Infrastructure	102
2.4.4	Maintenance Tactics of Railway Infrastructure	111
2.4.4.1	Routine Preventive Maintenance	113
2.4.4.2	Corrective Preventive Maintenance	113
2.4.4.3	Minor Breakdown Maintenance	113
2.4.4.4	Major Breakdown Maintenance	117
2.4.5	Technical Challenges of Railway Infrastructure Maintenance	117
2.5	Scheduling Techniques	118
2.6	Possession Scheduling	119
2.6.1	Analysis	119
2.6.2	Synthesis	129
3	Case of Application	132
3.1	Railway from Bellville to Wellington	132
3.2	The Monthly Maintenance Planning Process of Bellville Depot	135
3.2.1	Maintenance Possession Scheduling at Transnet Freight Rail	135
3.2.2	The Maintenance Possession Processes of Transnet Freight Rail	139
3.2.3	Maintenance Possession Performance Indicators	147
4	Model	148
4.1	Components of the Microscopic Infrastructure Representation	148
4.2	Assumptions	150
4.3	Definition of Sets, Parameters and Decision Variables	150
4.4	Objective Function	153
4.5	Constraints	153

CONTENTS

4.5.1	Train Movement Constraints	153
4.5.2	Train Routing Constraints	155
4.5.3	Block Section Occupancy Constraints	156
4.5.4	Maintenance Task Scheduling Constraints	157
4.5.5	The Domain of Variables	158
4.6	Validation	159
5	Experiments and Results	160
5.1	Experiments	160
5.2	Results	163
5.3	Evaluation	164
5.4	Transnet Feedback	175
6	Conclusions	178
6.1	Project Summary	178
6.2	Project Appraisal and Future Works	179
	References	189

List of Figures

1.1	TFR core rail network	2
2.1	TFR rail network	16
2.2	Overview of fixed railway infrastructure components	27
2.3	Terminology of track components	28
2.4	Components of the applied forces acting on a track structure	29
2.5	Vertical load distribution through a track structure	29
2.6	Traditional form of the wheel and rail contact	30
2.7	Concrete sleeper dimensions	31
2.8	Typical drainage configuration and water sources	32
2.9	Horisontal and vertical track planes	33
2.10	Track superelevation or cant	34
2.11	Track gauge	35
2.12	Overview of the components in a 3 kV DC traction circuit	35
2.13	Pantograph	36
2.14	Contact wire stagger	40
2.15	Points blade layout	41
2.16	Points equipped with an electric points machine	42
2.17	Examples of semaphore signals	43
2.18	Examples of colour light signals	44
2.19	Example of a route indicator	45
2.20	Illustration of an unoccupied track circuit	46
2.21	Illustration of an occupied track circuit	46
2.22	Block joint	47
2.23	Insulated track sections	47

LIST OF FIGURES

2.24	An axle counter layout	48
2.25	An example of axle counter heads	49
2.26	An example of an axle counter insert	50
2.27	An example of an axle counter evaluator	51
2.28	Communication network between the CTC and interlocking equipment .	51
2.29	CTC push button panel and diagram	52
2.30	CTC control desk	53
2.31	Overview of railway telecommunications infrastructure	55
2.32	A hypothetical rail demand forecast based on origin-destination groupings	57
2.33	A hypothetical example of transportation demand transferred onto the rail network	58
2.34	Rail demand over time	59
2.35	Categories of rail traffic	61
2.36	Illustrations of regular bulk traffic and direct bulk shipments	62
2.37	An example of a train diagram	66
2.38	Typical priorities and concerns of a physical asset management framework	70
2.39	Sustainable physical asset management	71
2.40	Interventions to improve asset effectiveness	72
2.41	PAM can improve RONA	74
2.42	Division of maintenance costs by origin	74
2.43	Optimum reliability cost	76
2.44	Contributions of the three elements of maintenance cost reduction . . .	77
2.45	Maintenance types	79
2.46	Evolution of equipment management	80
2.47	Cost advantages of maintenance types	81
2.48	The basic maintenance process	85
2.49	The RCM process	86
2.50	RCM task selection logic	87
2.51	Six common failure profiles and their occurrence percentage in commer- cial aircraft	88
2.52	The Total Productive Maintenance process	92
2.53	Failure modes	93
2.54	The basic FMECA process	96

LIST OF FIGURES

2.55	Interdependence of railway operations and maintenance	98
2.56	The track deterioration cycle	99
2.57	Track deterioration, rehabilitation and replacement	100
2.58	Track deterioration with maintenance interventions	101
2.59	Track deterioration with higher initial quality	101
2.60	Track deterioration when maintenance is deferred	102
2.61	High-level asset lifecycle process of railway infrastructure	103
2.62	The contribution of maintenance to the overall lifecycle cost of a railway	104
2.63	IM2000 railway inspection vehicle	107
2.64	Ultrasonic measuring car	107
2.65	The balance of reactive and proactive maintenance costs	109
2.66	Railway maintenance tactics	112
2.67	Decision framework for preventive maintenance	112
2.68	The run-to-failure maintenance tactic	112
2.69	On-track maintenance vehicle for overhead track equipment	114
2.70	Ballast tamper	115
2.71	Sleeper replacement machine	115
2.72	Rail grinding machine	116
2.73	Ballast screener	116
2.74	Classification of railway problems	122
3.1	Railway from Bellville to Wellington	133
3.2	Macroscopic representation of the railway from Bellville to Wellington .	134
3.3	The monthly maintenance planning process of Bellville depot	136
3.4	Overview of the maintenance possession process at Transnet Freight Rail	140
3.5	The monthly possession planning process of Bellville depot	141
3.6	The possession countdown process of Bellville depot	142
3.7	The possession countdown escalation process of Bellville depot	143
3.8	The NCC approval process for maintenance possessions	146
4.1	Definition of block sections	149
4.2	Elements of a microscopic station layout	149
5.1	The solution of Experiment 1	165

LIST OF FIGURES

5.2	The solution of Experiment 2	166
5.3	The solution of Experiment 3	167
5.4	The solution of Experiment 4	168
5.5	The solution of Experiment 5	169
5.6	The solution of Experiment 6	170
5.7	The solution of Experiment 7	171
5.8	The solution of Experiment 8	172
5.9	The solution of Experiment 9	173
5.10	The solution of Experiment 10	174
5.11	Feedback from senior technical experts of train scheduling, rail network maintenance and capital projects	176

List of Tables

2.1	Financial results of Transnet 2018	14
2.2	Volumes transported by Transnet 2018	15
2.3	Industrial engineering skills applied by IEs in the various TFR functions	23
2.4	Functional condition measurement parameters	106
2.5	Structural condition measurement parameters	106
2.6	Technical challenges of railway infrastructure maintenance	117
2.7	List of references used for the possession scheduling literature analysis .	121
4.1	Set definitions	151
4.2	Parameter definitions	152
4.3	Decision variables definitions	152
5.1	List of experiments	162

Nomenclature

Acronyms

ABL	Agriculture and Bulk Liquids
AC	Alternating Current
ADMM	Alternating Direction Method of Multipliers
CAB	Containers and Automotive Business
CAO	Chief Admin Occupations
CAS	Condition Assessment Systems
CBM	Condition-Based Maintenance
CGR	Cape Government Railways
CP	Capital Program
CP	Constraint Programming
CPM	Corrective Preventive Maintenance
CSAR	Central South African Railways
CTC	Centralised Traffic Control
DC	Direct Current
DE	Depot Engineer
DoT	Department of Transport

Nomenclature

DPE	Department of Public Enterprises
EBITDA	Earnings Before Interest, Tax, Depreciation and Amortisation
EBIT	Earnings Before Interest and Tax
ET	Engineering Technician
FCC	Feeder Catenary Contact
FMECA	Failure Modes, Effects and Criticality Analysis
ICTM	Information and Communication Technology Management
IDC	Industrial Development Corporation of South Africa Limited
IEBoK	Industrial Engineering Body of Knowledge
IE	Industrial Engineer
ILP	Integer Linear Programming
IM	Infrastructure Manager
IOM	Iron Ore and Manganese
IP	Integer Programming
ITCMS	Integrated Train Condition Monitoring System
ITP	Integrated Train Plan
LP	Linear Programming
MAB	Major Breakdown Maintenance
MIB	Minor Breakdown Maintenance
MILP	Mixed Integer Linear Programming

Nomenclature

MIP	Mixed Integer Programming
MMC	Mineral Mining and Chrome
MM	Maintenance Manager
MMP	Maintenance Manager Planning
MPS	Maintenance Possession Scheduler
MTS	Master Train Schedule
NCC	National Command Centre
NWB	Next Week's Business
NZASM	Netherlands South Africa Railway Company
OCC	Operational Control Centre
ODP	Organisational Development and Performance
OFC	Optical Fibre Cable
OHTE	Overhead Track Equipment
PAM	Physical Asset Management
PCB	Primary Circuit Breaker
PESP	Periodic Event Scheduling Problem
PM	Production Manager
RCA	Root Cause Analysis
RCM	Reliability Centred Maintenance
RN	Rail Network
RONA	Return on Net Assets
RPM	Routine Preventive Maintenance

Nomenclature

RSA	Republic of South Africa
SAA	South African Airways
SAC	Steel and Cement
SAIIE	South African Institute of Industrial Engineers
SAR&H	South African Railways and Harbours
SATS	South African Transport Services
SCS	Supply Chain Services
TBM	Time-Based Maintenance
TCO	Train Control Officer
TE	Transnet Engineering
TFR	Transnet Freight Rail
TGC	Transnet Group Capital
TI	Track Inspector
TNPA	Transnet National Ports Authority
TPL	Transnet Pipelines
TPM	Total Productive Maintenance
TPT	Transnet Port Terminals

Chapter 1

Introduction

Chapter 1 introduces the thesis topic by providing the background and describing the problem. It then sets the research aim and scope. This is followed by the research objectives, methodology and deliverables.

1.1 Background

Transportation systems are vital to society's social, economic and environmental well-being. The expected increase in transportation demand, urban congestion, greenhouse gas emissions and the use of limited energy resources all necessitate the development of efficient logistics systems.

Transportation systems are made up of various transport modes such as air, road, rail, pipelines and water. The different modes have unique performance characteristics with regards to speed, reliability, flexibility, volume and cost. Intermodal transportation involves two or more transport modes. The strengths of each mode must be exploited to alleviate the national transportation demand.

Railways are an essential part of the national transportation system. The low rolling resistance and fuel consumption combined with the high power output of electrical and diesel engines make railways efficient at transporting high volumes of passengers and freight (Ebersöhn & Gräbe, 2008a). Railways are the most economical mode of land transport for large volumes over long distances. This is of particular benefit to South Africa where the high volume routes span long distances between mining areas, ports and cities. In urban areas, railways can move high volumes of daily commuters on

1.1 Background

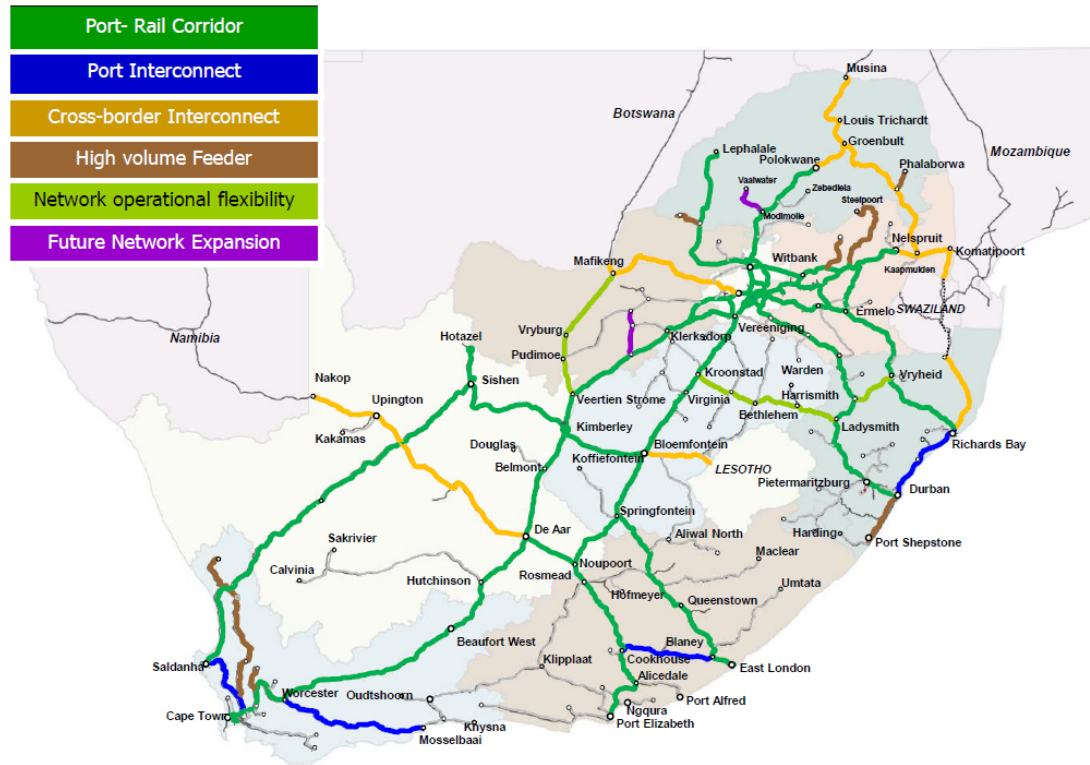


Figure 1.1: TFR core rail network (Gräbe, 2017)

much less land than road, while producing less air pollution and more safely (Ebersöhn & Gräbe, 2008a).

Transnet is a publicly listed South African logistics company with the South African government as its sole shareholder. Transnet Freight Rail (TFR) is the division of Transnet that owns, maintains and operates South Africa's freight railways. During 2018, TFR moved 226.3 million tonnes of freight. This was 90.8 million tons of general freight, 77 million tons of export coal and 58.5 million tons of export iron ore (Transnet website, 2018a).

TFR's railway network is shown in Figure 1.1. The network has a total route distance of 20 953 km which adds up to a total track length of 30 400 km. The total track length includes the lengths of parallel railway tracks on the same route. 7 100 km of the network is electrified. The iron ore line runs from Sishen to Saldanha while the coal line runs from Broodsmynersplaas to Richards Bay.

Maintenance is necessary to preserve the functional integrity of railway infrastruc-

1.2 Problem Description

ture: too little maintenance and the system breaks down, too much and it interferes with service delivery. Maintenance also accounts for a significant portion of total operational expenditure and therefore needs to be contained. The objective of maintenance managers is to plan the right amount and type of maintenance to ensure that the railway network is kept in a reliable, available, affordable and safe condition (Duvel, 2017).

Maintenance *possessions*, also known as track *occupations*, are required to execute maintenance on the rail network. Possessions are official authorities given to maintenance personnel that allow them to work on a track section.

Possessions are central to the maintenance scheduling problem since it impacts on available railway capacity and structures further maintenance scheduling (Lidén, 2015). It is also where there exists opportunities to save maintenance costs by sharing resources among the different engineering disciplines.

1.2 Problem Description

The *train schedule* allocates time and space on the rail network. A *train slot* is a specific series of sequential time and space allocations on the train schedule. The number of train slots on a route is determined and limited by the rail network layout.

Train slots are known by their origin, destination, planned departure and arrival times. Each train slot is assigned a train number. In a scheduled railway system, trains are planned to run in these train slots.

The *possession schedule* determines when track sections are taken out of service for maintenance activities. These possessions may interfere with the planned train service, *i.e.* the train slots.

There are three types of maintenance possessions: major possessions, minor possession and unplanned possessions. Major possessions do take track sections out of service and therefore requires that movement restrictions are applied to trains. Minor possessions do not take track sections out of service and therefore do not require that movement restrictions are applied to trains. Unplanned possessions, also known as emergency possessions, are taken when there is a breakdown which impacts on the train service.

1.2 Problem Description

Operations is the department responsible for planning and executing the movement of trains. *Infrastructure* is the department responsible for planning and executing maintenance of the fixed infrastructure.

A conflict exists between scheduling trains and maintenance possessions. Since the ultimate goal of the railway is to generate profit, the train service takes priority, but maintenance possessions must be scheduled for the railway to remain operational. If maintenance possessions are not scheduled the infrastructure will continue to deteriorate and the likelihood of breakdowns and service disruptions will increase. But on the other hand, if maintenance possessions are planned poorly it will increase interference with train service and reduce available capacity.

Maintenance depots have to apply for possessions from an Operational Control Centre (OCC). These possession requests are then approved or cancelled based on their operational impact. Train planners have three options to accommodate possessions on the train schedule: (1) depart trains earlier or later, (2) redirect trains on an alternative path if such a path exists, or (3) cancel the train. If the operational impact of the possession request is deemed unacceptable the request may be denied. In which case alternative dates and times are proposed to the requester. The process will then repeat until a suitable date and time is agreed.

From a train service perspective there are possession scheduling considerations. Firstly, the number of major possessions have a fundamental impact on network availability since major possessions take track sections out of service. The number of major possessions must therefore be minimised. This can be done by combining major possessions within and among the different engineering disciplines. This also has the benefit of reducing maintenance execution and travel costs by sharing resources. Other considerations from a train service perspective include avoiding possessions during the metro peak times and that certain trains, *e.g.* luxury long distance passenger trains, may not be delayed or cancelled.

From a maintenance planning perspective there are also possession scheduling considerations. Firstly, certain on-track maintenance machines are scheduled for maximum utilisation on the national network. Therefore, possessions for those on-track machines are prioritised. Secondly, certain on-track machines and maintenance work need to be scheduled in a specific sequence for technical reasons. From a maintenance execution point of view, few possessions of a long duration are preferred over many possessions

1.3 Research Aim

of a short duration since each possession will incur set up and travel costs. It is also preferred that possessions are scheduled sooner rather than later. Since, if the planned work is complete and the budget is spent it supports the motivation for the following year's maintenance budget.

The possession scheduling problem is to find a possession schedule that minimises service disruptions while scheduling all the required possessions in line with train service and maintenance planning preferences.

1.3 Research Aim

The *research aim* is to

develop a maintenance possession scheduler (MPS) for Transnet Freight Rail that can be used to analyse and recommend possession schedules of rail tracks.

At the core of the MPS is an *analytic model* of the *train schedule*. A *model solution* *allocates space and time* on the *network* to trains and maintenance possessions. The combined schedule is for a planning period, *e.g.* a day, on a *railway route* with a *specified layout*, *e.g.* a single line with loops or double line with crossing points. The *objective* of the model is to *minimise deviations* of the train service, while scheduling the necessary number of maintenance possessions.

1.4 Scope

It is intended that the MPS be for general use, although it will be applied on the Bellville to Wellington route. Furthermore, the focus will be on incorporating maintenance possessions onto the train schedule, not on determining what maintenance jobs are necessary. Only possessions that impact the train service, *i.e.* major possessions, will be considered. Only infrastructure maintenance possessions will be considered not rolling stock, *i.e.* locomotives and wagons, maintenance. Furthermore, the MPS will only schedule trains and maintenance possessions. The MPS will not schedule train crews and rolling stock resources.

1.5 Research Objectives

The research objectives are to:

1. *Understand the railway maintenance possession scheduling problem.*
2. *To obtain adequate knowledge and understanding of the railway systems, maintenance, scheduling, optimisation and programming theory required to solve the problem.*
3. *Develop and validate a model of the problem.*
4. *Develop a computer application that can be used to solve the problem and thus be used as a decision support system.*

1.6 Research Methodology

To complete the research objectives the following methodology is proposed:

1. Investigate the real-world problem at the Bellville infrastructure maintenance depot by conducting interviews with maintenance managers and possession planners while observing the current maintenance planning and possession scheduling processes.
2. Conduct a literature study of railway infrastructure and operations, maintenance in general, maintenance specific to railway infrastructure, scheduling in general, and possession scheduling.
3. Develop a model that can be used to study the problem.
4. Rigorously validate and verify the developed model. Ensure that the model is an adequate representation of the real world, and that the model's end users have confidence in the model's outputs. Further, identify and repair all internal model errors.
5. Solve the model to produce a near-optimal possession schedule for a planning period.
6. Analyse the model outputs.

1.7 Deliverables

7. Prove the advantages of the developed model.
8. Recommend practical possession schedules for maintenance execution.

1.7 Deliverables

The deliverables are:

1. Master's thesis describing the project.
2. Maintenance possession scheduler, *i.e.* an electronic decision support system, that can be used by a technical person to search for near-optimal maintenance possession schedules.

1.8 Thesis Organisation

Chapter 2 provides an overview of Transnet Freight Rail, railway infrastructure and operations, maintenance of railway infrastructure, scheduling in general and possession scheduling. The case where the maintenance possession scheduler is applied is described in Chapter 3. The maintenance possession scheduling model is presented in Chapter 4. This is followed by the experiments and results discussed in Chapter 5. Finally, Chapter 6 concludes with the project summary, appraisal and recommendation for future works.

Chapter 2

Literature Study

This chapter opens with an overview of Transnet and the role of industrial engineering in Transnet Freight Rail. It then continues to present a literature study that covers railway infrastructure and operations, maintenance in general, maintenance specific to fixed railway infrastructure, scheduling in general and closes with railway possession scheduling.

2.1 Overview of Transnet

This section provides an overview of Transnet. It first describes how various independent railways were combined to form a national railway, then, how it transformed from a government organisation into company, and then presents Transnet as it is today. It concludes with the role of industrial engineering in Transnet Freight Rail.

2.1.1 The Transition from Trailblazing Railways to an Unified Government Railway

The history of Transnet is linked to the historical development of South Africa. The discovery of diamonds and gold was the catalyst of South Africa's early economic development but it was rail that unlocked the potential through rapid transportation ([TFR website, 2018b](#)). Rail was then, as it is now in developed countries, and how it could be in the future for South Africa, the fastest, most reliable transport mode for freight and passengers ([DoT website, 2017](#)) .

2.1 Overview of Transnet

The first train ran in Durban 1860 while the first train in Cape Town ran in 1862 ([Spoornet, n.d.](#)). The discovery of diamonds in Kimberley, in 1869, created an urgent need for cheaper and faster transport to the centre of country. Rail was the efficient mass transport mode needed to exploit the newly discovered mineral wealth ([TFR website, 2018b](#)).

The Cape Government Railways (CGR) formed in 1873 when the Cape government exercised its right, written into the original financing agreements, to purchase and operate the private railway lines around Cape Town. The government was determined to extend the railway line over the mountains to secure trade with the mining camps ([Martin, 2004](#)).

The Cape and Natal lines initially used the standard British gauge of 1435 mm. However, a committee of the Cape parliament decided, in 1873, to settle on a narrower 1065 mm gauge. The rationale was that a narrower track would make it easier to cut through the mountains on the way to the inland plateau ([TFR website, 2018b](#)) and lower construction costs by an estimated one-third ([Martin, 2004](#)). The narrower gauge became known as Cape gauge and became the standard for railway development in Southern Africa ([TFR website, 2018b](#)).

The rail network developed on an inland course from Cape Town to Johannesburg with routes diverging to Port Elizabeth and East London. Branchlines were constructed off of these mainlines to open up rural areas to agriculture and mining ([DoT website, 2017](#)). The Cape line reached Kimberley in 1885 ([TFR website, 2018b](#)).

In 1886, major gold discoveries had been made on the Witwatersrand in the Transvaal ([TFR website, 2018b](#)). By arrangement with the Orange Free State Government, in 1889, the CGR were allowed to extend the line into the Free State ([Spoornet, n.d.](#)). The Cape line reached Bloemfontein in 1890. In 1896, the Orange Free State Government decreed to take over all the railways within its boundaries ([TFR website, 2018b](#)).

The railway line reached Johannesburg in 1892 ([Martin, 2004](#)). Thus, connecting Johannesburg to Cape Town, Port Elizabeth and East London ([Spoornet, n.d.](#)). The line between Durban and Johannesburg was completed in 1896 ([Martin, 2004](#)). In 1887, the Netherlands South Africa Railway Company (NZASM) was founded to construct and operate a railway line between Pretoria and Maputo ([TFR website, 2018b](#)). The line to Maputo opened in 1898.

2.1 Overview of Transnet

During the Anglo-Boer War from 1899 to 1902 the Imperial Military Railways took control of the two republican railways. After the declaration of peace these railways became known as the Central South African Railways (CSAR) (Spoornet, n.d.).

The Central South African Railways, the Cape Government Railways and the Natal Government Railways were officially unified with The South Africa Act of 1909 as the South African Railways and Harbours (SAR&H) (DoT website, 2015). The railways gradually merged until they were centralised in Johannesburg in 1916. SAR&H was the first version of the organisation that would become Transnet in 1990 and the beginning of the modern era for South African transport (Martin, 2004).

2.1.2 The Transition from a State Department to a Commercial Company

Most of this subsection is based on Martin (2004). Other references are cited in the text. The South Africa Act of 1909 set out the strategic direction of SAR&H. It specified that the railways, ports and harbours of South Africa should enable the development of agriculture and industry by providing cheap transport. It further stipulated that the revenues generated by SAR&H should not exceed what it needs to operate and service its debts (DoT website, 2015). The focus of the organisation was therefore not on generating profit but on serving the developmental goals of the state.

More than 8 800 km of track was added to the network between 1910 and 1924 (Spoornet, n.d.). Most of these lines were branchlines, constructed to serve the agricultural sector (TFR website, 2018b). By 1930, most of the country's railway infrastructure had been laid. Few network expansions took place until 1973.

SAR&H grew into a multi-modal transport organisation. In 1912, it introduced road transport services to complement the railway service, in 1934, it took control of a private airline company and formed South African Airways (SAA), and in 1965, it commissioned its first pipeline for petroleum products from Durban to Gauteng (Spoornet, n.d.).

In order to provide affordable transport services to developing sectors and passengers, SAR&H cross-subsidised low income generating services with high income generating services. *E.g.*, cheap rates were offered to farmers, remote mining ventures and passengers, while rates were inflated for commodities such as manufactured goods,

2.1 Overview of Transnet

liquor and packaged foods. Port revenues were also used to cover railway costs. This practice of cross-subsidisation allowed SAR&H to break even over the long term.

The weakness of cross-subsidisation was that the customers of overpriced services would abandon rail once more economical transport options became available. This was the case when diesel trucks arrived in the 1920's. The competition posed by road transport threatened to collapse the pricing system that allowed the railways to remain in business.

The conflict between road and rail became a transport policy crisis. In 1929, a commission of inquiry was established to study the problem. The commission resolved to protect the interests of rail through transport regulation.

Road transport regulation was introduced by the Motor Carrier Transport Act of 1930. It prohibited companies from transporting goods on road more than 50 km from their premises without a permit. These permits were only issued for express shipments and commodities that the railways could not handle well, *e.g.* cement pipes. Industry and business interests protested against the transport legislation every time the SAR&H budget came up in parliament. They were upset for being overcharged when shipping high value goods on rail. This tension was managed for decades.

The transport regulation act did however leave room for discretion when granting permits for road transport. As time passed these permits were issued more frequently. As a result the railways' estimated share of the national freight market, measured in ton-km, fell from 61 percent in 1957 to 51 percent in 1971.

In 1975, a series of National Transport Studies reviewed international trends of transport deregulation. These studies found that local transport policies were a constraint to economic growth. At the time, public opinion also started shifting away from state control of industries towards a free market philosophy.

The first step towards transport deregulation was relaxing the application requirements for road permits with the Road Transport Act of 1977. This act also increased the permissible road vehicle mass to 42 tons. As a result, rail's share of the freight market fell further to 37 percent by 1981.

In 1981, the name of SAR&H was changed to South African Transport Services (SATS) to reflect its multi-modal capabilities ([DoT website, 2015](#)). The organisation was also restructured into divisions for railways, harbours, road transport, aviation and pipelines ([TFR website, 2018b](#)).

2.1 Overview of Transnet

The 1980's were a period of economic strain for South Africa. There were severe capital shortages as a result of international economic sanctions against apartheid and increasing military costs because of civil wars in neighbouring Angola and Mozambique. As a result, capital spending in all government departments came under scrutiny (DoT website, 2015).

Against the backdrop of economic pressure, Wim de Villiers, a former chairman of Sasol, was tasked with benchmarking SATS against similar enterprises around the world. He highlighted that investments had been made in sectors where the railways were unable to compete with other modes and warned against growing losses in the railways (DoT website, 2015). The railways had lost more than R1.5 billion between 1983 and 1985.

The following recommendations were made (DoT website, 2015):

1. SATS should cut back on new rail investment and rather focus on increasing utilisation of existing assets.
2. SATS should restructure to separate the major modes namely railways, harbours, airways and pipelines.
3. Suburban passenger services should be separated from the rest of SATS and should be subsidised directly by Government.
4. SATS should be allowed flexibility to set tariffs that would provide adequate returns.
5. SATS should operate like a private investor-owned company, required to make a profit by functioning as a commercial enterprise under government ownership whilst earning an appropriate return on capital by cutting costs and managing assets better.

In 1986, parliament accepted the recommendations of the de Villiers report. SATS was directed to reorganise as a private corporation within four years, with the government as its sole shareholder.

The corporatisation of SATS required that fundamental changes be made to its procedures, organisation and culture. The management philosophy changed from pursuing technical excellence through capital investment to maximising return on assets.

2.1 Overview of Transnet

Marketing capabilities were strengthened and management were reorganised into business units according to market segments. To improve competitiveness, costs had to be lowered. 50 Percent of costs were staff. The headcount was reduced by natural attrition, *i.e.* not replacing staff that resigned or retired, and so avoided involuntary retrenchments. The number of employees went down from a peak of 245 000 in 1982, to 165 000 in 1990.

Government started reducing the capital allocated to SATS. Prior to 1986, SATS received up to R2 billion capital a year. In 1986, it was reduced to R1.44 billion. In 1988, only R699 million was allocated. By 1990, zero capital was allocated (DoT website, 2015).

Total deregulation of road transport was implemented with the Transport Deregulation Act of 1988, which meant that SATS was competing in a free market for the first time. On top of that, the Road Traffic Act of 1989 increased the permissible road vehicle mass further from 42 to 45 tons. These measures caused the road transport sector to expand significantly (DoT website, 2015).

In 1990, SATS ceased to exist and a new corporate entity, Transnet Limited, was registered (Spoornet, n.d.). At the same time, passenger rail assets and operations were transferred to the newly established Rail Commuter Corporation (DoT website, 2017). After 80 years as an arm of government, Transnet was now obliged to function like a business, *e.g.* make a profit, pay taxes and be subject to private auditing (Spoornet, n.d.).

2.1.3 Transnet Limited a State Owned Company

Transnet is a publicly listed South African logistics company with the South African government as its sole shareholder (TFR website, 2018a). The board of directors reports to the minister of public enterprises (DPE website, 2018).

Transnet owns, operates and maintains South African railways, ports and pipelines (Transnet website, 2018c). The company is made up of the following divisions (Transnet website, 2018b):

- Transnet Freight Rail (TFR) owns, maintains and operates South Africa's freight railways (TFR website, 2018a).

2.1 Overview of Transnet

Table 2.1: Financial results of Transnet 2018 ([Transnet website, 2018a](#))

Revenue	R72.9 billion
Operating Expenses	R40.4 billion
EBITDA	R32.5 billion
Net Profit	R4.9 billion

- Transnet Engineering (TE) is focussed on the manufacture, upgrade, repair and maintenance of railway locomotives and wagons ([Transnet website, 2018d](#)).
- Transnet National Ports Authority (TNPA) provides the port infrastructure and marine services at eight commercial ports in South Africa ([Transnet website, 2018f](#)).
- Transnet Port Terminals (TPT) operates on the South African ports where it handles containers, mineral bulk, agricultural bulk and Ro-Ro, *i.e.* roll-on roll-off, sectors ([Transnet website, 2018h](#)).
- Transnet Pipelines (TPL) transports petroleum and gas products through its long distance pipelines ([Transnet website, 2018g](#)).

The Transnet group is also supported by company wide specialised units. These are Transnet Group Capital (TGC), Transnet Property and Transnet Foundation ([Transnet website, 2018c](#)). TGC provides consulting, engineering and project management services to the Transnet divisions, Transnet Property manages the Transnet property portfolio, and Transnet Foundation is responsible for corporate social investment.

A summary of Transnet's annual results for the financial year ending in 2018 are provided in Table 2.1. R21.8 billion worth of capital investments were made during the 2018 financial year. The total capital expenditure over the past six years were R165.6 billion ([Transnet website, 2018a](#)). Although Transnet is wholly owned by the state it does not receive government subsidies nor sovereign guarantees for its debt. Investments are funded from the company's balance sheet alone ([DoT website, 2015](#)). A summary of the transported volumes during the 2018 financial year is shown in Table 2.2.

TFR is the largest division of Transnet and contributed 59 percent of the EBITDA 2018 ([Transnet website, 2018a](#)). The company is vertically integrated in that it not

2.1 Overview of Transnet

Table 2.2: Volumes transported by Transnet 2018 ([Transnet website, 2018a](#))

General Freight	90.8 million tonnes
Export Coal	77 million tonnes
Export Iron Ore	58.5 million tonnes
Total Rail	226.3 million tonnes
Containers (TPT)	4.664 million TEUs
Petroleum	16.345 billion litres

only owns, operates and maintains the rolling stock needed to transport goods, but also owns and maintains the infrastructure that it operates on ([DoT website, 2015](#)).

TFR owns a vast railway network across South Africa. An overview of the network is shown in Figure 2.1. The network has 30 400 km of track and 20 953 route km. The core network is 12 801 route km.

TFR has two flagship export railway lines. These are the Iron Ore and Coal lines. The Iron Ore line runs from Sishen in the Northern Cape to Saldanha bay in the Western Cape. The line is famous for its 342 wagon production trains that are 3.8 km long. The Coal line is fed from mines in the Waterberg, Ogies and surrounding areas through Ermelo to Richards Bay. Trains of up to 200 wagons run on the Coal line.

Organisationally TFR is organised in commodity centred business units to promote customer centricity, while operationally it functions in channels to maximise total production volumes. The business units are ([Transnet website, 2018e](#)):

- Agriculture and Bulk Liquids (ABL)
- Containers and Automotive Business (CAB)
- Coal
- International Business
- Iron Ore and Manganese (IOM)
- Mineral Mining and Chrome (MMC)
- Steel and Cement (SAC)

2.1 Overview of Transnet



Figure 2.1: TFR rail network ([Transnet website, 2017](#))

2.1 Overview of Transnet

The channels are the Thaba, Phala, Coal, Natcor and Cape channels. The Thaba channel is the north western part of the railway system. It connects the borders of Botswana and Zimbabwe to Gauteng, while it also flows from Lephalale through Thabazimbi to Gauteng. The Phala channel consists of a mainline flowing from Gauteng through Komatipoort to Maputo. It also connects Zimbabwe to South Africa from Beitbridge through Phalaborwa and then Swaziland to Richards Bay. The Natcor channel is focussed on container traffic and transports goods between Gauteng and Durban. The Cape channel consists of the Iron Ore line, the Manganese line running from Hottazel to Port Elizabeth, and the Cape line running from Cape Town in a northern direction through Beaufort West and De Aar. The Cape channel also includes the rail network in the Eastern Cape. Quick facts of TFR ([TFR website](#), 2018a):

- Moves 17 percent of South Africa's freight annually
- 100 percent of Export coal
- 100 percent of iron ore
- 30 percent of the core network carries 95 percent of freight volumes
- Annual revenues of over R37 billion
- Over 31 000 employees

2.1.4 The Role of Industrial Engineering in Transnet Freight Rail

Engineering bridges the gap between science and human needs. It involves the use of intellect, energy, materials, systems, money and labour to develop, operate and maintain undertakings that aim to conserve and make optimum use of natural resources, to provide the means of communication and transport, as well as produce a great variety of things ([Sperotto, 2015](#)).

Engineering specialisations developed in response to the different problems arising from the wide range of human activities. For example, engineers who focussed on the construction of bridges and buildings became known as civil engineers and those who focussed on machines became known as mechanical engineers. A multitude of engineering specialisations have developed to deal with the complexity of technology ([Darwish & Van Dyk, 2016](#)).

2.1 Overview of Transnet

Given the separate development of each discipline the need for integration arose. The need to integrate (i) technologies within systems, and (ii) systems with their external social, economic and natural environments. This type of integration requires a unique combination of engineering and non-engineering knowledge (Darwish & Van Dyk, 2016).

Industrial engineers (IEs) are specialists of integration. They have the technical skills, capabilities, and problem-solving mind-set of engineers, while also having the ability to stand back and see the whole picture. IEs can evaluate the parts of a system in relation to the whole. They aim to maximise the overall value obtained from a system for the good of humanity, the environment and the economy (Darwish & Van Dyk, 2016).

The need for industrial engineering in South Africa started in the 1920s when major industries, *e.g.* electricity and steel production, were established. Industrialisation accelerated during World War II when there were restrictions on imports. This led to the introduction of more sophisticated mass production and manufacturing technologies. In 1940, the Industrial Development Corporation of South Africa Limited (IDC) was set up by government. The main objective of the IDC was to establish and develop industrial enterprises on sound business principles. The IDC led the way with several industries and projects, *e.g.* SASOL's oil-from-coal project (Sperotto, 2015).

The South African Association of Production Engineers formed in 1943 and later evolved into the South African Institute of Industrial Engineers (SAIIE) in 1981 (Sperotto, 2015). The profession developed from the problems in industry, interdisciplinary integration, growth, balance and connectedness. During the 20th century the focus was on work studies, standardisation, and production lines. Industry needed solutions to the problems that originated from the complex interactions between workers, productivity and resources (Darwish & Van Dyk, 2016).

Today, IEs are responsible for the analysis, design, planning, implementation, operation, management and maintenance of integrated systems consisting of people, money, material, equipment, information and energy (University of Stellenbosch website, 2018). IEs optimise these integrated systems to improve quality and productivity and to reduce costs (Schutte *et al.*, 2016).

Industrial engineering has a broad knowledge base, skillset and application areas (Darwish & Van Dyk, 2016). On top of the fundamental engineering knowledge areas of

2.1 Overview of Transnet

mathematics and science such as Calculus, Statistics, Probability, Chemistry, Physics and the Engineering Sciences, the Industrial Engineering Body of Knowledge (IEBoK) includes the following knowledge areas ([Institute of Industrial and Systems Engineers, 2019](#)):

- Work design and measurement
- Operations research and analysis
- Engineering economic analysis
- Facilities engineering and energy management
- Quality and reliability engineering
- Ergonomics and human factors
- Operations engineering and management
- Supply chain management
- Safety
- Information engineering
- Design and manufacturing engineering
- Product design and development
- Systems design and engineering

The broad knowledge base equips IEs with a unique thinking style that is influenced by systems, processes, creative and entrepreneurial approaches. Industrial thinking starts with identifying value for the customer, it is needed to close the gap between state of the art research and deployed technologies, and it continuously improves systems used for manufacturing and supply. Industrial thinking considers supply and demand, optimisation, standardisation, end goal, scale, work and value. IEs use systems thinking to understand the complexity of an interconnected world, while industrial thinking is used to extract value from it ([Darwish & Van Dyk, 2016](#)).

2.1 Overview of Transnet

The unique thinking style of IEs enables them to identify connections between technologies, knowledge fields and specialisations. These connections may unlock value by taking advantage of new industry sectors, sub-specialities or synergies (Darwish & Van Dyk, 2016). Consequently, IEs enhance the standard of human life by increasing productivity, improving the quality of products and services, and improving work environments (Schutte *et al.*, 2016).

Industrial engineering differs from industry to industry and within companies in different departments. IEs may use fundamentally different tools and methods depending on the work they do (Darwish & Van Dyk, 2016). Industrial engineering jobs may vary from very technical engineering-based work, to softer business- and law-based consulting services (Schutte *et al.*, 2016).

IEs work across all major industries. Although, there is low number of IEs working in general government services in the Republic of South Africa. In the RSA, the top five industries where IEs work are (Schutte *et al.*, 2016):

1. Manufacturing
2. Finance, insurance and business services
3. Transport, storage and communication
4. Mining and quarrying
5. Wholesale and retail trade, and accommodation

Transnet is the second biggest employer of IEs in the RSA. The top five are (Schutte *et al.*, 2016):

1. Sasol
2. Transnet
3. Eskom
4. Imperial Logistics Group
5. LTS Consulting

Within Transnet Freight Rail (TFR) IEs work in the following departments:

2.1 Overview of Transnet

- **National Command Centre (NCC)**: is responsible for train service design, service planning, service execution management and emergency response.
- **Organisational Development and Performance (ODP)**: is responsible for process standardisation, process adherence, continuous improvement through lean six sigma projects, performance measuring and monitoring, and the implementation of governance systems.
- **Rail Network (RN)**: plans and executes maintenance of the fixed railway infrastructure so that it remains in a reliable, available, affordable and safe condition. RN minimises infrastructure breakdowns and invests in infrastructure to retain the installed capacity.
- **Information and Communication Technology Management (ICTM)**: enables business performance through process integration and automation, expanding digital connectivity, developing digital platforms, rationalising the existing ICTM landscape, and reducing the total cost of information and communication technology.
- **Capital Program (CP)**: is responsible for delivering the TFR capital portfolio from strategic planning to post implementation review. This is done by developing long term plans for capacity requirements, developing business cases, analysing engineering design problems, developing practical solutions to create capacity, developing revised or new operating models, developing operational readiness plans, managing operational readiness activities when new assets or services are deployed, monitoring and reporting on project progress, and tracking project benefits.

Typical examples of work that IEs do in TFR include:

- Redesign of train service operating models to increase capacity or enhance operational efficiencies
- Management of operational readiness activities when deploying new services or assets

2.1 Overview of Transnet

- Lean and six sigma process improvement projects such as bottleneck identification and cycle time reduction in train operations and departmental processes
- Data analysis for projects and operational performance management
- Improve integration between departments
- Internal consulting and process auditing
- Developing automated spreadsheets
- Planning and scheduling of maintenance jobs
- Root cause analysis of railway infrastructure faults or train service delays
- Multi-criteria analysis for prioritising Copex and Opex maintenance budgets
- Managing compliance to standards
- Project management

Table 2.3 shows which industrial engineering skills are applied by IEs in the various departments of TFR. The table shows that a large subset of the industrial engineering skills are applied by IEs in TFR irrespective of their departments. Whereas, some other industrial engineering skills, such as Enterprise Engineering, Operations Research, Rapid Prototyping and Reliability Engineering, are only applied in specific functions.

There are also some industrial engineering skills that are not applied in TFR at all. These include Machine Learning, Design of Experiments, Ergonomics, Facilities Management, Mechatronics and Automation, Robotics and Simulation Modelling. A reason for Machine Learning could be that its applications within the railway industry are still in the research and development phase and not widely adopted in practice. IE skills such as Ergonomics and Facilities Management are applied by specialists or other departments within TFR. And for Simulation Modelling the reason is that the frequency of simulations needed is low and therefore the business prefers to outsource that skill.

Other departments in TFR where IEs could work, but where they are not typically employed include:

2.1 Overview of Transnet

Table 2.3: Industrial engineering skills applied by IEs in the various TFR functions

Industrial Engineering Skills	NCC	ODP	RN	ICTM	CP
Business Analytics	Yes	Yes	Yes	Yes	Yes
Business Process Modelling	Yes	Yes	Yes	Yes	Yes
Change Management	Yes	Yes	Yes	Yes	Yes
Data Analysis	Yes	Yes	Yes	Yes	Yes
Financial Analysis	Yes	Yes	Yes	Yes	Yes
Innovation Management	Yes	Yes	Yes	Yes	Yes
Interface Design	Yes	Yes	Yes	Yes	Yes
Leadership Roles	Yes	Yes	Yes	Yes	Yes
Needs Analysis	Yes	Yes	Yes	Yes	Yes
Project Management	Yes	Yes	Yes	Yes	Yes
Quality Management	Yes	Yes	Yes	Yes	Yes
Requirements Analysis	Yes	Yes	Yes	Yes	Yes
Risk Management	Yes	Yes	Yes	Yes	Yes
Root Cause Analysis	Yes	Yes	Yes	Yes	Yes
Strategy Development	Yes	Yes	Yes	Yes	Yes
Systems Design	Yes	Yes	Yes	Yes	Yes
Technology Management	Yes	Yes	Yes	Yes	Yes
Conceptual Design	Yes	Yes	No	Yes	Yes
Environment, Health and Safety Management	Yes	Yes	Yes	No	Yes
Facility Design	Yes	Yes	Yes	No	Yes
Lean Six Sigma	Yes	Yes	Yes	Yes	No
Statistical Modelling	Yes	Yes	Yes	No	Yes
Supply Chain Design	Yes	Yes	Yes	No	Yes
Time Studies	Yes	Yes	Yes	No	Yes
Information Systems Development	No	Yes	Yes	Yes	No
Operations Management	Yes	Yes	Yes	No	No
Enterprise Engineering	No	Yes	No	No	No
Operations Research	Yes	No	No	No	No
Rapid Prototyping	No	No	No	Yes	No
Reliability Engineering	No	No	Yes	No	No
Machine Learning	No	No	No	No	No
Design of Experiments	No	No	No	No	No
Ergonomics	No	No	No	No	No
Facilities Management	No	No	No	No	No
Mechatronics and Automation	No	No	No	No	No
Robotics	No	No	No	No	No
Simulation Modelling	No	No	No	No	No

2.1 Overview of Transnet

1. **Operations:** is responsible for planning and executing the movement of trains to deliver on customer commitments and minimise service variations. IEs support Operations through the NCC and ODP functions.
2. **Technology Management:** is responsible for short, medium, and long term technology strategy, specifications, research and development, evaluation and approval of components and systems, and technical audits. Technology Management is responsible for the technology life cycle, whereas RN is responsible for the asset life cycle.
3. **Risk Management:** ensures that TFR identifies both strategic and operational risks and that measures to manage these risks and their impacts are implemented, whilst simultaneously, taking advantage and harnessing inherent opportunities. This is done by embedding pro-active risk management into business processes to facilitate informed decision making, integration of business continuity and resilient models into the organisation to improve business process recoveries and resumptions, management of environmental risk exposure, effective handling of hazardous and dangerous goods, and insurance management.
4. **Supply Chain Services (SCS):** deals with the inbound supply chain, *i.e.* the procurement of goods and services for TFR. Also, the transportation and storage of these goods. SCS is core part of the TFR strategy and the procurement policies are based on the principles of fairness and transparency, social equity, and value for money.
5. **Commercial:** is the central function responsible for commercial management, commercial strategy and planning, market research, customer experience management, real estate, and new business development. Each business unit has its own sales function.

Even in departments where IEs do not work they can be used as internal consultants to increase productivity, improve the quality of products and services, and improve work environments. Departments where IEs could perform this function includes:

- **General Counsel:** is responsible for legal, governance, compliance and fraud risk management.

2.1 Overview of Transnet

- **Finance:** is responsible for creating a platform for customer transactions, driving profitability and growth, ensuring that the correct assets are purchased at the right time at the right cost, managing cash flow and ensuring that financial reporting and governance responsibilities are delivered.
- **Human Capital:** is responsible for making sure the employees are well trained and motivated to enable the company to deliver on its promises, to attract and retain talent, and to make TFR an employer of choice.

A major capability of IEs is their interconnectedness to other fields of study (Sperotto, 2015). They are aware of where the solution to a problem might lie even though they are not necessarily subject matter experts of the problem. They are able to find the relevant tools, knowledge fields, and individuals that are necessary to solve a problem. If they are not qualified to apply a specific tool or knowledge field, they become responsible for the integration and quality management of the work produced by specialists (Darwish & Van Dyk, 2016).

The 4th industrial revolution will have a major effect on how IEs are utilised. This is reflected in the industrial engineering research trends. Research related to production systems are decreasing, while research related to intelligent systems, supply chain management, and information technology are on the increase. It is expected that future industrial engineering research on topics such as optimisation and quality will also increase (Schutte *et al.*, 2016).

The evolution of industry brings about a rapid increase in the need to deal with complexity and with it a significant increase in the amount of roles requiring industrial engineering capabilities. Currently, demand for IEs exceeds the supply: industrial and production engineering was placed at No. 8 on the scarce skills list of the Department of Higher Education in 2014 (Schutte *et al.*, 2016). Given the growing need for industrial engineering skills in the railway industry and the contributions IEs make, by extracting value from systems through integration, systems thinking, quality and productivity improvements, reducing costs, innovation, and overall versatility, TFR needs to attract and retain IEs to be competitive in the 4th industrial revolution.

2.2 Railway Infrastructure and Operations

2.2 Railway Infrastructure and Operations

Transportation systems consists of fixed infrastructure, moving infrastructure and control systems (Gräbe, 2017). The fixed infrastructure of a railway is made up of the track, bridges, tunnels, power distribution equipment, signalling and telecommunication systems. The moving infrastructure, known as rolling stock, consists of the locomotives and wagons that are combined to form trains. The control systems are the technologies used by the Train Control Officers (TCOs) and train crews to authorise and execute the safe movement of trains.

2.2.1 Fixed Railway Infrastructure

Fixed railway infrastructure is made up of different technologies that integrate with one another to function as a safe pathway for trains. An overview of the main railway components is shown in Figure 2.2. A summary of the main features will be discussed in the following subsections.

2.2.1.1 Track Structure

The track structure components are grouped into the superstructure and substructure. Refer to Figure 2.3. The superstructure is made up of the rails, sleepers, fasteners, and crib ballast. Whereas, the substructure is made up of the top ballast, bottom ballast, subballast, placed soil and natural ground (Marutla, 2017).

The purpose of the track structure is to provide a stable guide-way for trains to operate safely and economically while withstanding the train loads and environmental factors. The track structure must therefore be able to sustain the applied forces in size and quantity for a reasonable time and maintenance requirement to provide an acceptable standard of operation (Ebersöhn & Gräbe, 2008b).

2.2.1.2 Components of Forces Acting on a Track Structure

The applied forces on the track structure can be divided into vertical, lateral and longitudinal components (Marutla, 2017). These components are shown in Figure 2.4.

The high vertical loads are concentrated at the wheel-rail contact surface. The rail distributes these loads over the sleepers and down through the ballast to the subgrade. Figure 2.5 illustrates the vertical load distribution through a track structure.

2.2 Railway Infrastructure and Operations

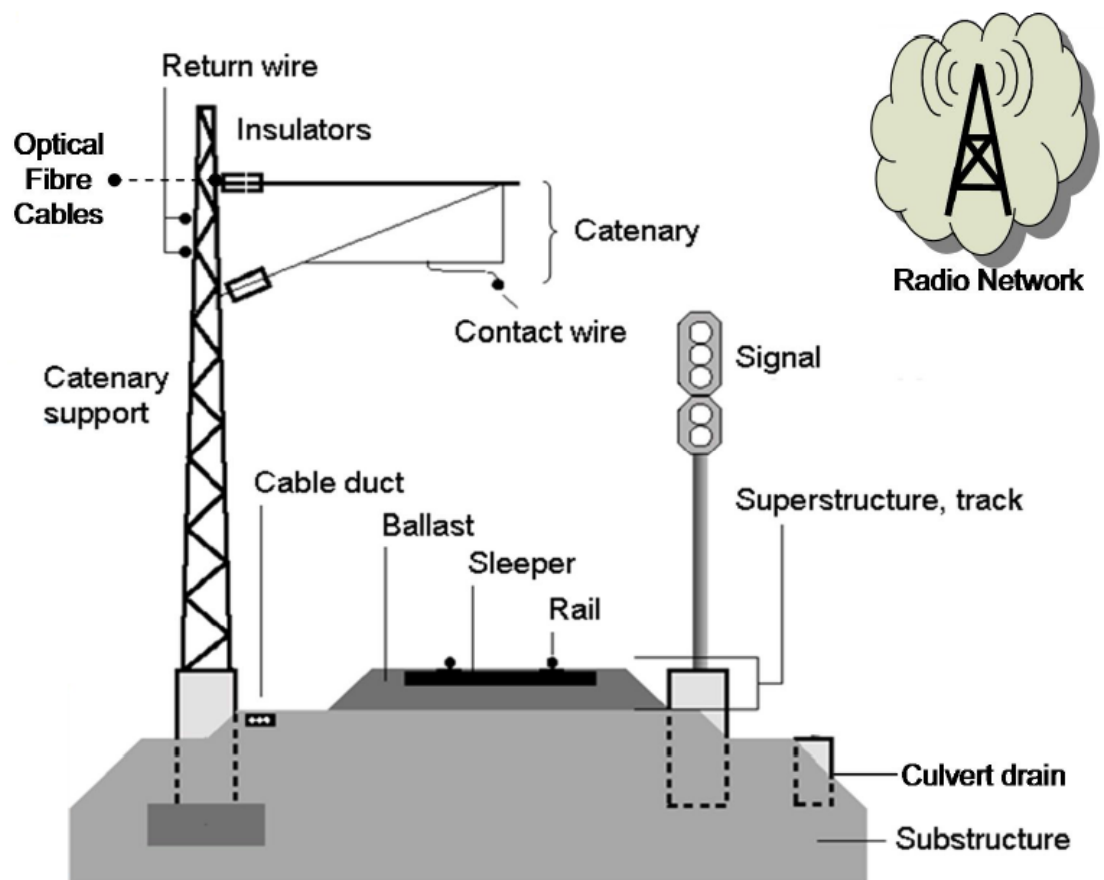


Figure 2.2: Overview of fixed railway infrastructure components (Marutla, 2017)

2.2 Railway Infrastructure and Operations

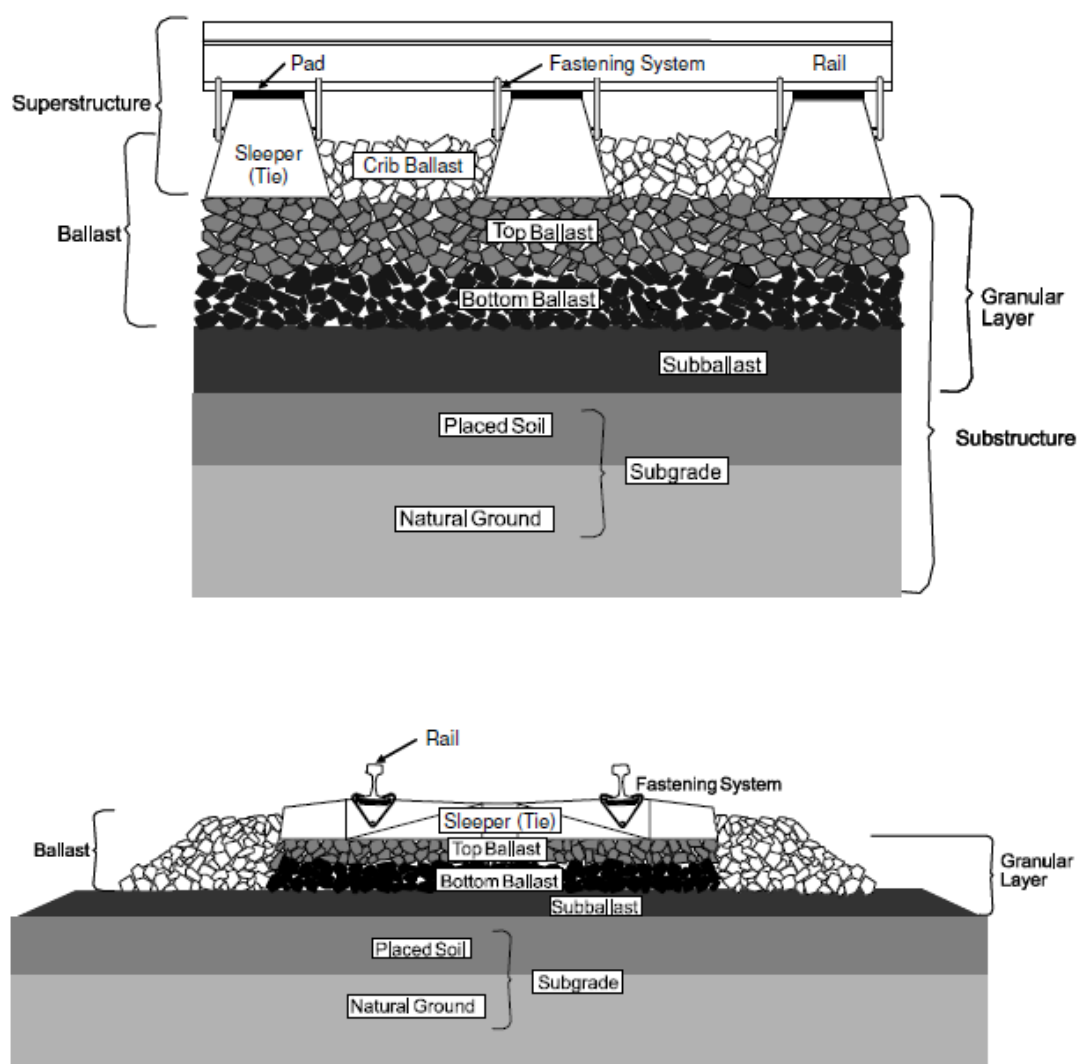


Figure 2.3: Terminology of track components (Ebersöhn & Gräbe, 2008b)

2.2 Railway Infrastructure and Operations

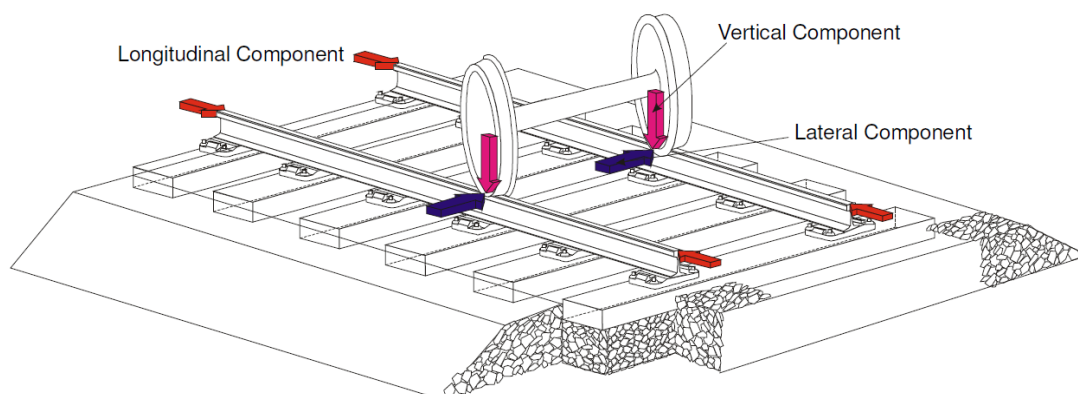


Figure 2.4: Components of the applied forces acting on a track structure (Ebersöhn & Gräbe, 2008b)

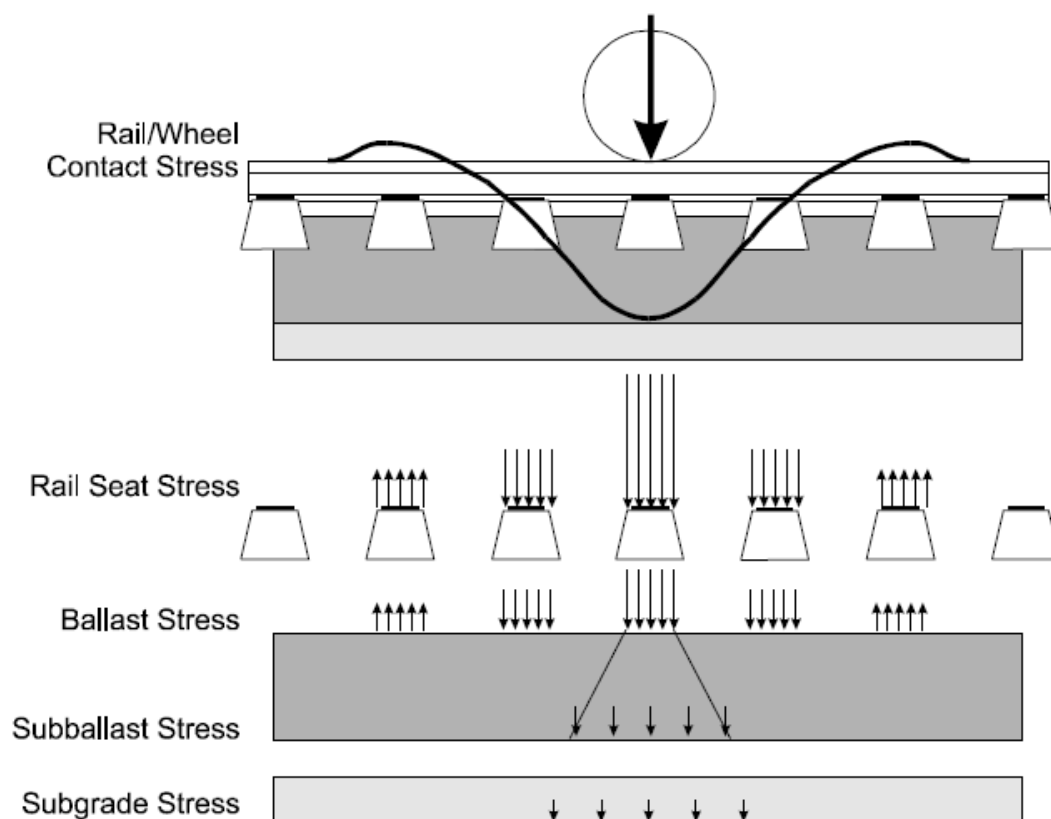


Figure 2.5: Vertical load distribution through a track structure (Ebersöhn & Gräbe, 2008b)

2.2 Railway Infrastructure and Operations

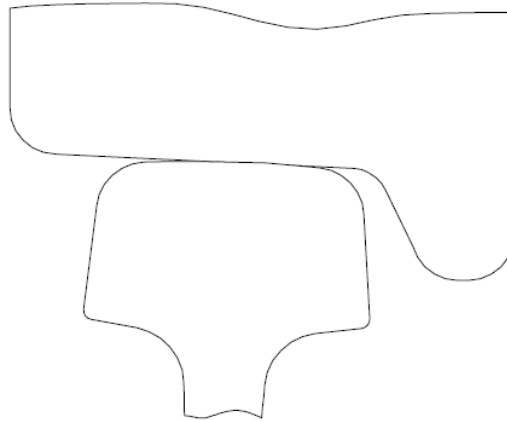


Figure 2.6: Traditional form of the wheel and rail contact (Ebersöhn & Gräbe, 2008b)

The lateral stiffness of the track superstructure and friction on the side and bottom of the sleepers resist forces pushing the track sideways. Whereas, the weight of the crib ballast and the sleepers anchored in the crib ballast resist longitudinal forces that push the track forwards (Ebersöhn & Gräbe, 2008b).

2.2.1.3 Main Track Components and Functions

Rails guide the train wheels, support the train loads and convey electric currents. The profile of the wheel and rail contact surface is designed for the wheels to roll evenly and continuously in a manner that minimises the wear and fatigue rate (Marutla, 2017). Refer to Figure 2.6 for an illustration of the traditional form of the wheel and rail contact. Rails further support and distribute the concentrated vertical loads and serve as electrical conductors in the electric locomotive and train detection circuits (Ebersöhn & Gräbe, 2008b). Traction power and train detection circuits are discussed in subsections 2.2.1.5 and 2.2.1.7, respectively.

The fastening system retains the rails against the sleepers. It resists the vertical, lateral, longitudinal and overturning movements acting on the rail. These forces are caused by the train wheels and temperature changes in the rail (Ebersöhn & Gräbe, 2008b).

Pads are installed between the rails and the concrete sleepers as shown in Figure 2.3. These pads damp the wheel induced vibrations and prevents the gradual wearing down of the rail sleeper contact (Marutla, 2017).

2.2 Railway Infrastructure and Operations

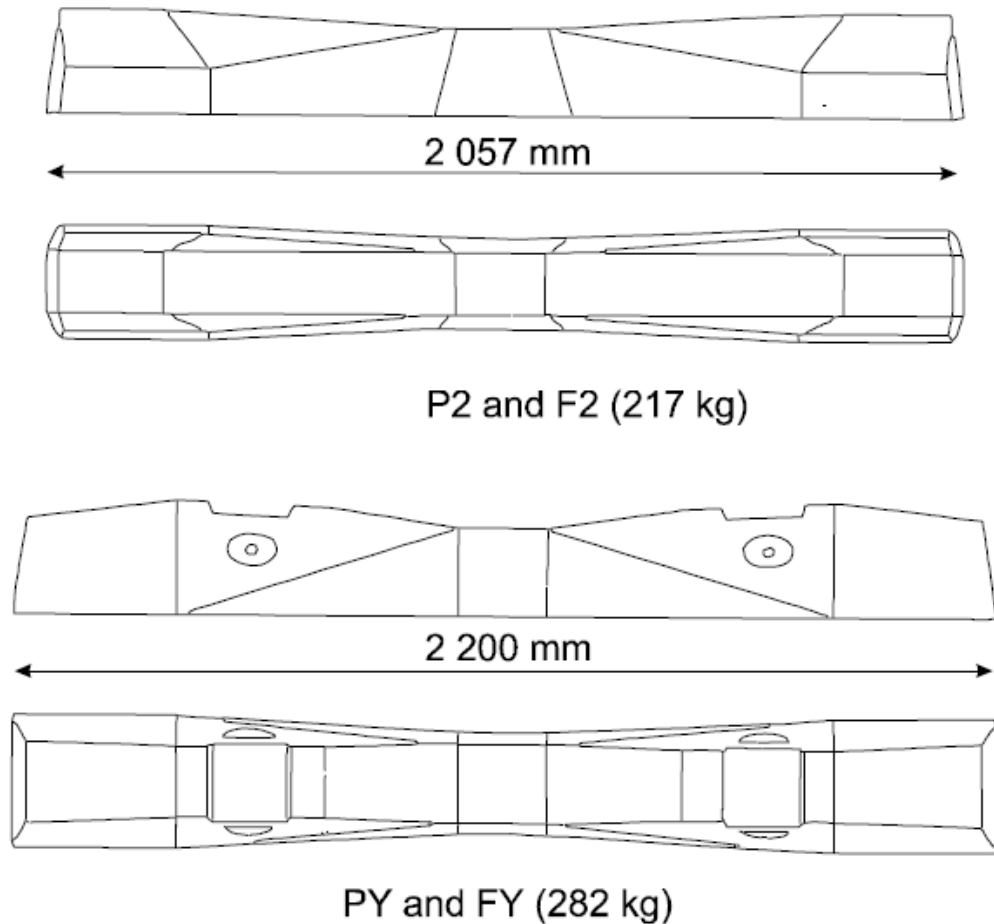


Figure 2.7: Concrete sleeper dimensions (Ebersöhn & Gräbe, 2008b)

Reinforced concrete sleepers are the most common type of sleepers used in South Africa. Although, timber or steel sleepers are also used in some areas. Figure 2.7 shows concrete sleeper dimensions (Ebersöhn & Gräbe, 2008b). Sleepers have several functions (Marutla, 2017):

1. Sleepers receive and distribute the load from the rails over the ballast.
2. Sleepers hold the fasteners to maintain the right distance between the rails, *i.e.* the track gauge.
3. Concrete sleepers provide the required angle to the rails for proper wheel-rail contact.

2.2 Railway Infrastructure and Operations

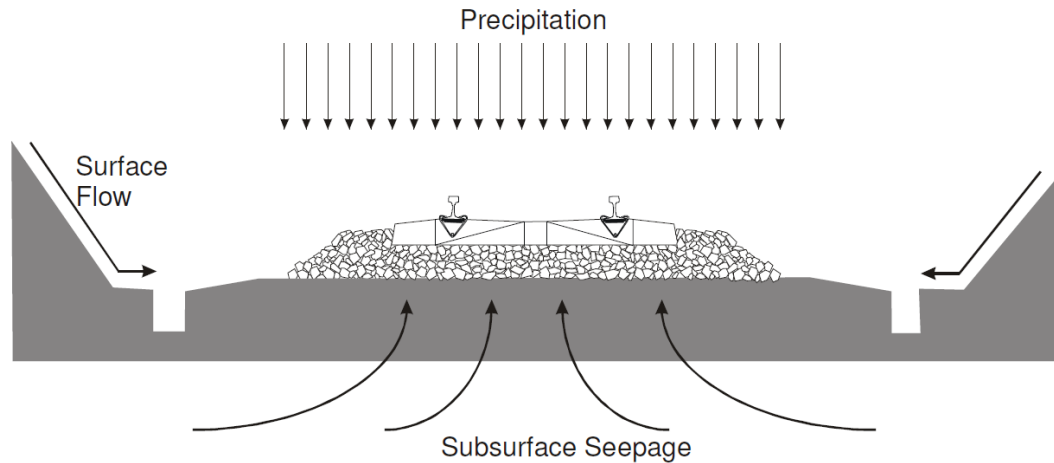


Figure 2.8: Typical drainage configuration and water sources (Ebersöhn & Gräbe, 2008b)

4. Sleepers resist rail movements by anchoring the superstructure in the ballast.

Ballast is the top layer of substructure. Ballast resists forces applied to the sleepers to retain the track in the required position, provides drainage for water falling onto the track and reduces the stress levels for the underlying material (Ebersöhn & Gräbe, 2008b).

Subballast separates the ballast and the subgrade. Subballast reduces the stress transferred to the subgrade, prevents the ballast from wearing down the subgrade surface and directs water to the drainage channels (Marutla, 2017).

The subgrade is the base of the track structure. It consists of placed soil on natural ground. The purpose of the subgrade is to form a stable foundation for the subballast and ballast (Ebersöhn & Gräbe, 2008b).

A typical drainage configuration and water sources are shown in Figure 2.8. Moisture influences the properties of the substructure and therefore drainage contributes to the performance of the track structure. The drainage system intercepts subsurface water entering the track substructure from below, it intercepts surface water approaching the track structure from the sides, and removes water running through the ballast from the top of the track structure (Ebersöhn & Gräbe, 2008b).

2.2 Railway Infrastructure and Operations

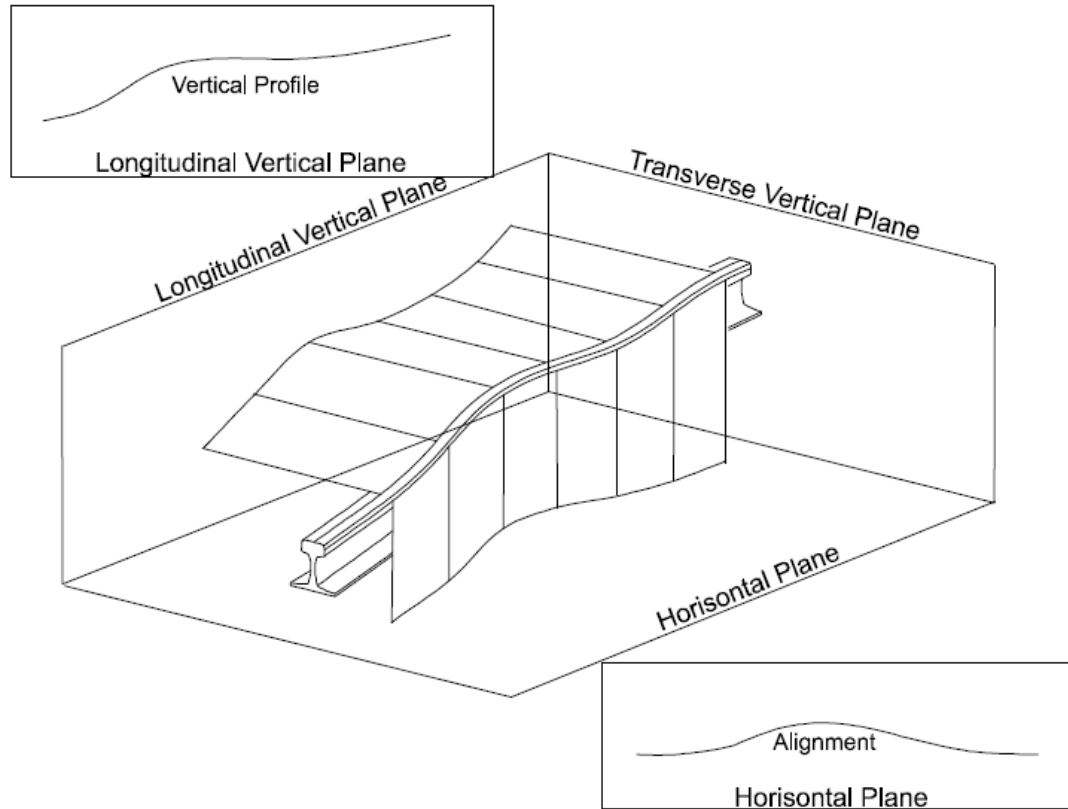


Figure 2.9: Horizontal and vertical track planes (Ebersöhn & Gräbe, 2008b)

2.2.1.4 Track Geometry

This subsection is based on Marutla (2017). Track geometry refers to the position of each rail in space. The longitudinal vertical, transverse vertical and horizontal reference planes are shown in Figure 2.9. Track geometry is measured to determine the condition of the track so that it can be maintained effectively.

Alignment refers to the sideways or lateral position of the rail. Whereas, profile refers to the vertical position of the rail.

Superelevation or cant refers to the difference in elevation between the top of two adjacent rails. Refer to Figure 2.10. Cant is applied in curves to distribute centrifugal force more evenly.

Track gauge is the distance between the inside of the rails, as shown in Figure 2.11. Standard gauge is 1435 mm, while Cape gauge is 1065/7 mm. Cape gauge is installed

2.2 Railway Infrastructure and Operations

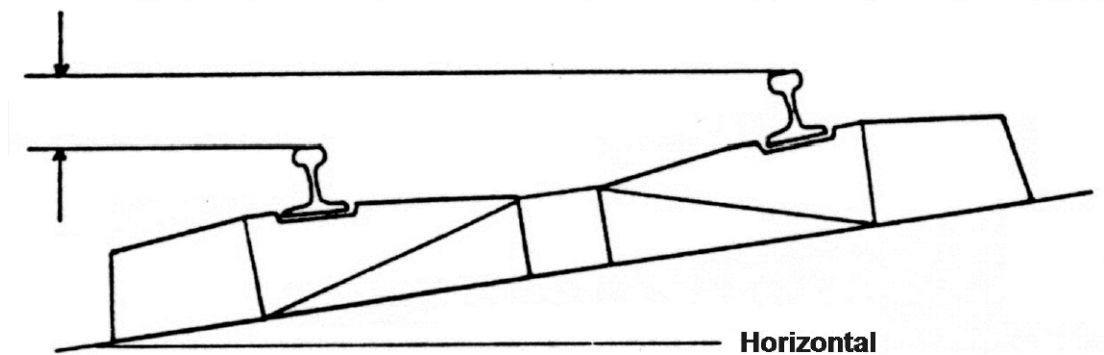


Figure 2.10: Track superelevation or cant (Marutla, 2017)

on most of the South African railways.

2.2.1.5 Electrical System

The primary function of the electrical system is to supply power to the traction motors on board electric locomotives (du Plessis, 2017a). In South Africa, 3 kV DC, 25 kV AC and 50 kV AC traction circuits are used (du Plessis, 2017b). There are also electrical subsystems that supply 11 kV AC to the railway signalling circuits. An overview of 3 kV DC traction circuit components is shown in Figure 2.12.

In most cases, the electricity provider supplies 33 kV AC, 66 kV AC or 132 kV AC to the substations (du Plessis, 2017b). The substation then transforms the incoming power to the required voltage and current for the traction motors (TFR, 2012). The power is then fed through the overhead track equipment (OHTE) to the locomotive. The locomotive collects the current from the contact wire with a pantograph (du Plessis, 2017a). Figure 2.13 shows a pantograph. The current then continues through the traction motor on board the locomotive (du Plessis, 2017b). Finally, the current returns to the substation via one of the rails and auxiliary conductors to complete the traction circuit (Kruger, 2012).

A summary of the functions of the major components of a 3 kV DC traction circuit will be discussed next. Most of the summary is based on TFR (2012). Other references are cited in the text.

2.2 Railway Infrastructure and Operations

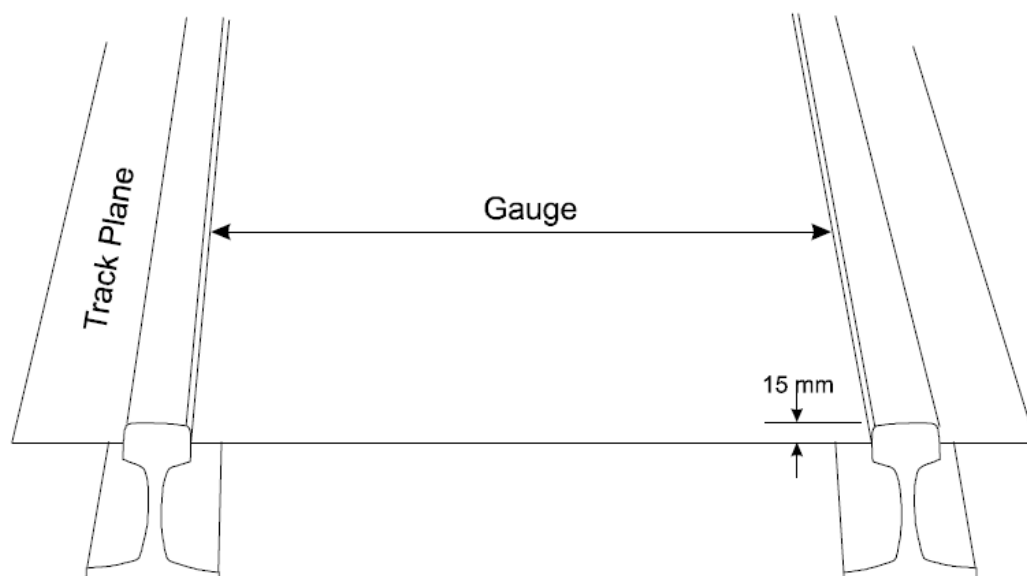


Figure 2.11: Track gauge (Ebersöhn & Gräbe, 2008b)

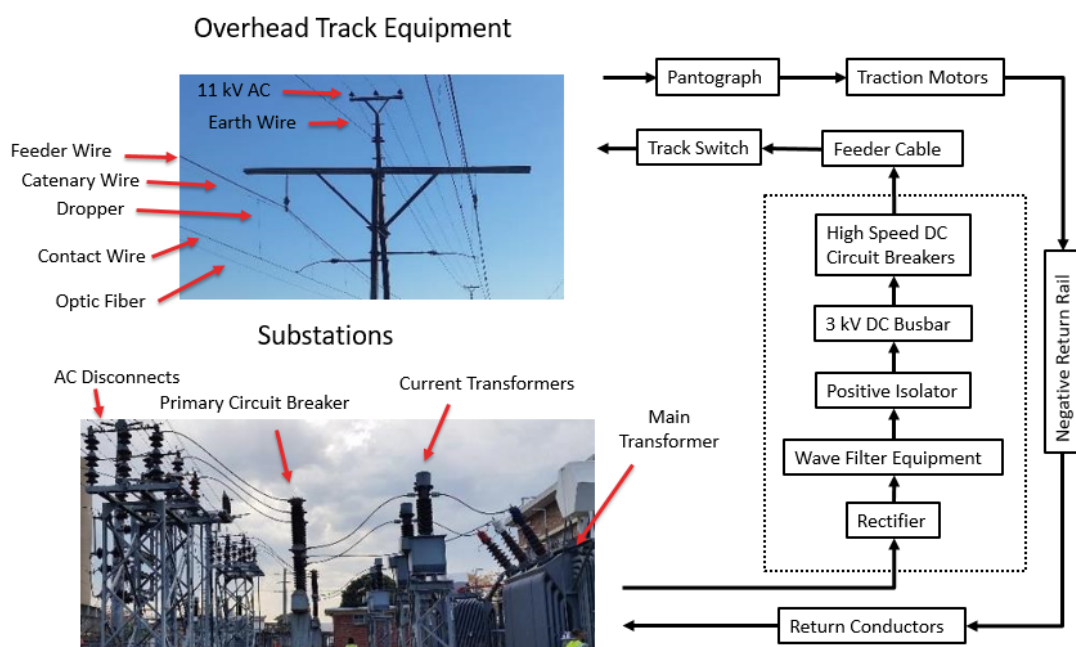


Figure 2.12: Overview of the components in a 3 kV DC traction circuit

2.2 Railway Infrastructure and Operations



Figure 2.13: Pantograph (du Plessis, 2017a)

Lightning Arrestors

- Lightning arrestors are installed between the incoming supply and the AC disconnecting switches. The lightning arrestors protect the substation equipment by opening the circuit in case of electric surges.

AC Disconnects

- The disconnecting switches are used to isolate the substation from the incoming supply to carry out maintenance work.

Primary Circuit Breaker (PCB)

- The PCBs are installed before the main transformer, on the primary side, to protect the transformer from overloading and high fault currents.

Current Transformers

- Current transformers are installed for power metering and equipment protection purposes. It steps down the incoming current to a readable value for the metering

2.2 Railway Infrastructure and Operations

equipment and also trips if the incoming current is too high.

Main Transformer

- The main transformer steps down the supplied voltage to 3 kV AC and outputs 6 phases to the rectifier. It is also the most expensive piece of equipment in a substation. That is why there is a series of equipment installed in front of it that offer protection against overloading and high fault currents.

Rectifier

- The rectifier is connected to the six phases on the secondary side of the main transformer. It converts alternating current to direct current.

Wave Filter Equipment

- This equipment removes harmonic voltages that may interfere with telephone lines, signalling circuits and control circuits on some locomotives.

Positive Isolator

- The positive isolator is used to isolate and earth the output of the rectifier for maintenance work.

3 kV DC Busbar

- The busbar is a high current carrying conductor and provides points to tap into the circuit.

High Speed DC Circuit Breakers

- The high speed DC circuit breakers are installed between the 3 kV DC busbar and OHTE circuits to protect the OHTE from overloading and high fault currents.

Feeder Cable

- The feeder cable supplies the 3 kV DC to the OHTE switch structure.

2.2 Railway Infrastructure and Operations

Track Switch

- The purpose of the track switch is to isolate a section of the OHTE from power so that maintenance work can be done safely in the track section (du Plessis, 2017b).

Feeder Wire

- The feeder wire is the main current carrying conductor of the OHTE (du Plessis, 2017a).

Catenary Wire

- The catenary wire supports the contact wire that is suspended from it with vertical wires known as droppers. The catenary also carries current. The amount of current it carries is proportional to the ratio of the contact and catenary wire sizes (du Plessis, 2017a).

Contact Wire

- The contact wire carries electric power to the pantograph of the locomotive. The pantograph continuously taps of power from the contact wire (du Plessis, 2017a).

Feeder Catenary Contact (FCC) Jumpers

- The FCC jumpers are not shown in Figure 2.12. They are vertical wires that collect current from the feeder wire and distributes it to the catenary and contact wires (du Plessis, 2017a).

Return Wire

- The return wire provides a secondary path for the return current in case there is a fault on the main return conductors (du Plessis, 2017a).

11 kV AC Wires

- The 11 kV AC wires provides power to the electronic signalling system.

2.2 Railway Infrastructure and Operations

Optic Fibre

- The optic fibre is used for telecommunications and signalling systems (Marutla, 2017).

2.2.1.6 Overhead Track Equipment Geometry

This subsection is based on du Plessis (2017a). The purpose of the OHTE is to supply power to the locomotive and to enable uninterrupted current collection by the pantograph during motion. This is achieved by the profile of the contact and catenary wires, the constant pressure exerted by the pantograph on the contact wire, and keeping the condition of the track and supporting structures within construction tolerances and specifications.

OHTE geometry measurements include stagger and contact wire height. The contact wire is suspended slightly to the left and right at successive supports so that the contact wire slides from side to side on the pantograph. This is known as stagger. Refer to Figure 2.14. Stagger enables the pantograph to wear down more evenly.

2.2.1.7 Signalling System

Most of this subsection is based on Ostrofsky (2014). While, other reference are cited throughout the text. Signalling systems are used to authorise train movements. There are several train authorisation methods available ranging from elementary to advanced. Train authorisation methods include telegraph order, wooden staff, van Schoor token, pilot working, radio train order, track warrant, communication-based authorisation systems and lineside signalling. Lineside signalling is the most common authorisation method used on South African mainlines and will therefore be the focus of this subsection.

Signalling is the communication system between a Centralised Traffic Control (CTC) office and a train driver. A Train Control Officer (TCO), based in the CTC, will be in control of a specific railway section. The signals provide visual information to the train driver. Signals inform the driver if a track section is available, which route to take and what the safety of the way ahead is. The purpose of signalling is to ensure that there is a safe distance between trains, safeguard movements at crossings, regulate the movement of trains and provide railway capacity.

2.2 Railway Infrastructure and Operations



Figure 2.14: Contact wire stagger (du Plessis, 2017b)

2.2 Railway Infrastructure and Operations

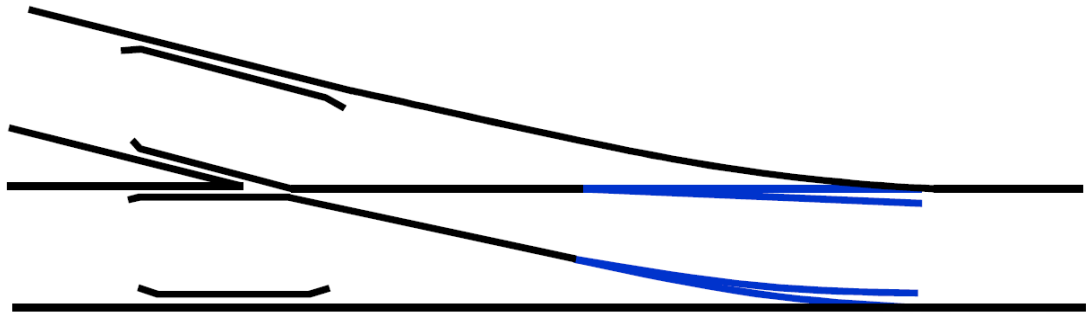


Figure 2.15: Points blade layout ([Ostrofsky, 2017](#))

The elements of a signalling system are:

- Points, also known as turnouts or crossings
- Signals
- Train detection
- Interlocking
- Control mechanisms, *i.e.* man-machine interfaces

Points

Points enable the movement of trains from one track to another. They can be operated by hand, mechanically, electrically or hydraulically. On contemporary signalling systems the points are controlled remotely from the CTC. A typical points layout is shown in Figure 2.15. An electric points machine is shown in Figure 2.16.

Signals

Signals provide information to the train driver that enable him to decide how to handle the train. There are two types of signals: Semaphore, *i.e.* mechanical, signals and colour light signals. Examples of semaphore signals are shown in Figure 2.17.

The indications of colour light signals are known as aspects in railway terminology. Figure 2.18 shows examples of colour light signals. A route indicator is shown in Figure 2.19. Route indicators show to which route or track platform a train is being directed. Many other signals and combinations are in use.

2.2 Railway Infrastructure and Operations



Figure 2.16: Points equipped with an electric points machine (Ostrofsky, 2014)

2.2 Railway Infrastructure and Operations

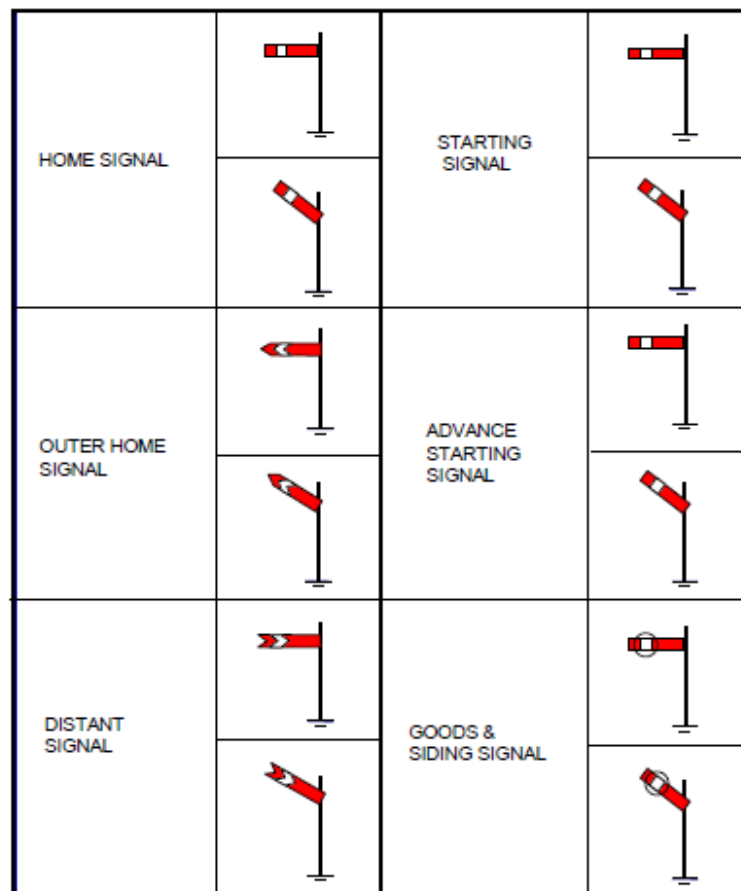


Figure 2.17: Examples of semaphore signals ([Ostrowsky, 2014](#))

2.2 Railway Infrastructure and Operations

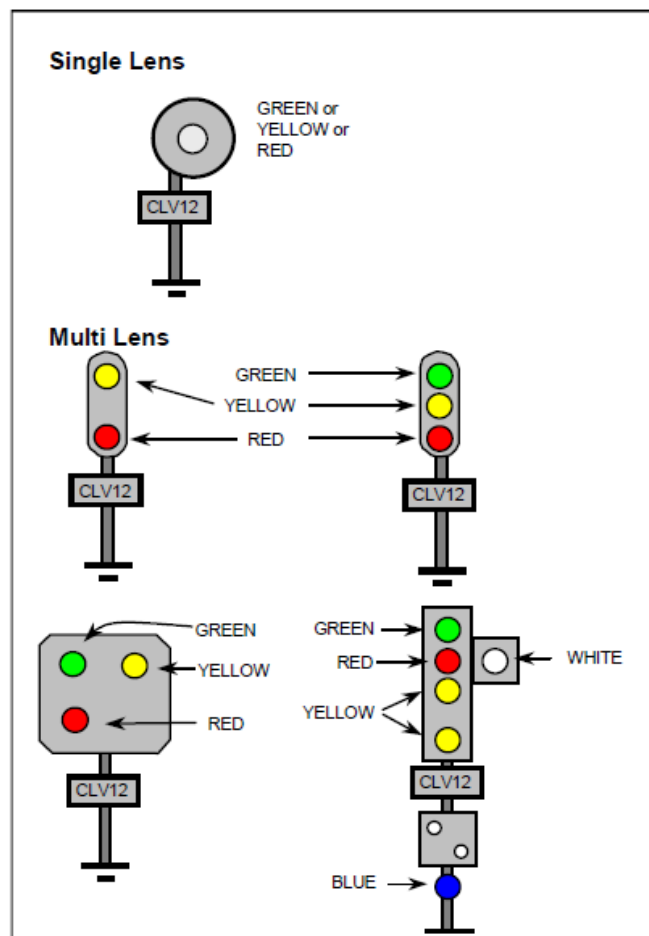


Figure 2.18: Examples of colour light signals (Ostrowsky, 2014)

2.2 Railway Infrastructure and Operations

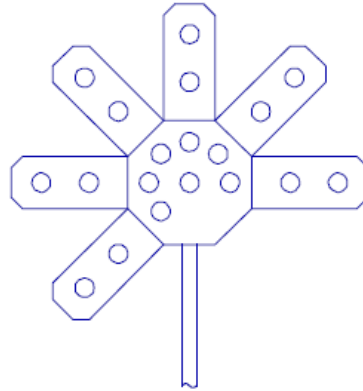


Figure 2.19: Example of a route indicator ([Ostrofsky, 2014](#))

Train Detection

Track vacancy or train detection in a railway section is determined by track circuits or axle counters. A basic illustration of a track circuit is shown in Figure 2.20. Current flows through both rails to pick up a relay when the section is unoccupied. The relay is energised under normal conditions and therefore any circuit failure, such as a rail break, can be detected immediately.

When a train is in the section, the wheels and axles form a low resistance path in parallel with the relay and so the relay de-energises and drops, as shown in Figure 2.21. This indicates that the railway section is occupied.

Insulated rail joints, known as block joints, insulate different track circuits from one another to form track sections. Refer to Figure 2.22 and Figure 2.23.

Axle counters can be used instead of track circuits. Axle counters detect the completeness of trains, counts wheels and determines the direction of movements. An axle counter layout is shown in Figure 2.24.

Rail mounted sensors, known as axle counter heads, are installed on either end of a track section. The axle counter heads generate a magnetic field around the rail. As the train wheel passes over the axle counter head, the phase of the induced voltage reverses and the electronics counts this as a train wheel. The insert converts this input to a frequency and transmits it to an evaluator. The number of pulses generated corresponds

2.2 Railway Infrastructure and Operations

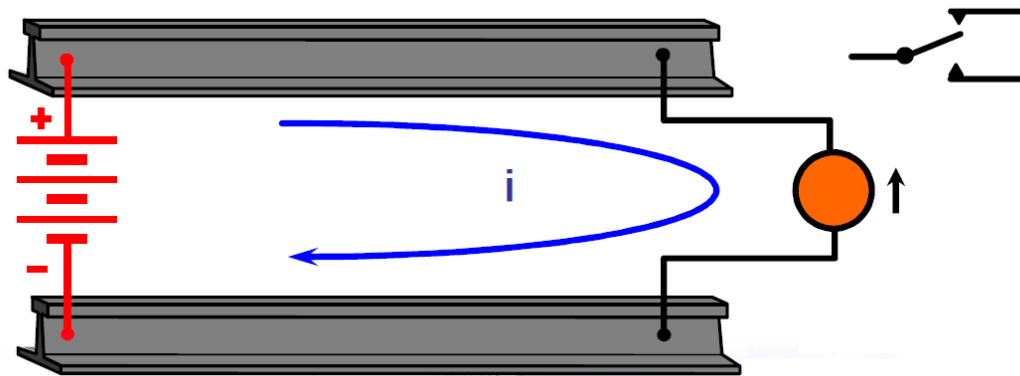


Figure 2.20: Illustration of an unoccupied track circuit (Ostrofsky, 2017)

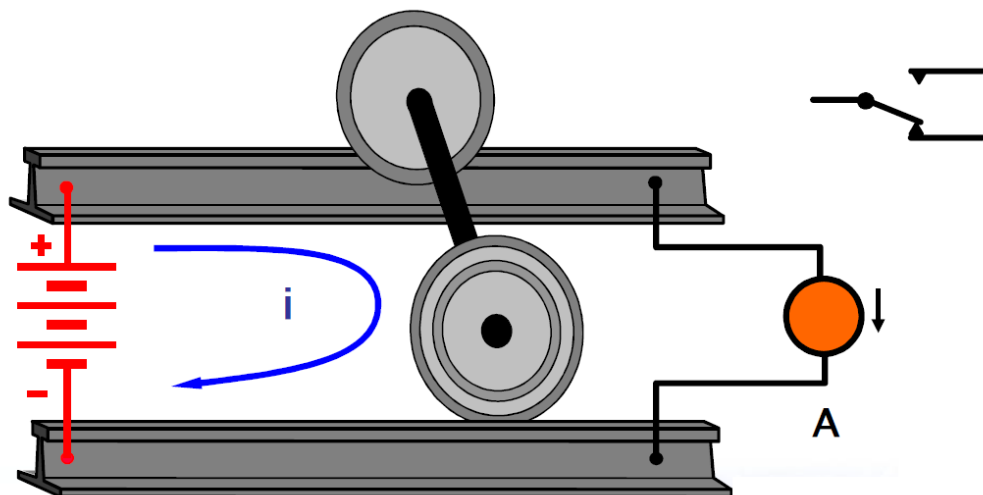


Figure 2.21: Illustration of an occupied track circuit (Ostrofsky, 2017)

2.2 Railway Infrastructure and Operations



Figure 2.22: Block joint (Ostrowsky, 2017)

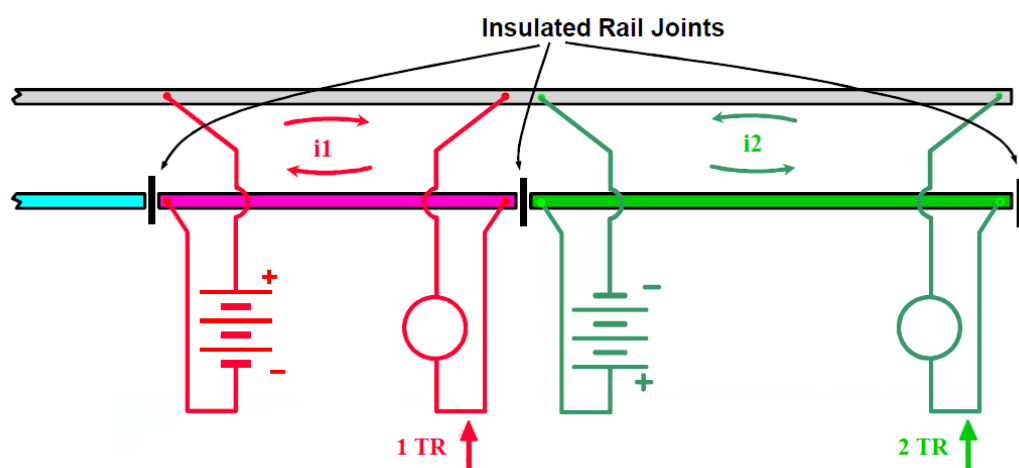


Figure 2.23: Insulated track sections (Ostrowsky, 2017)

2.2 Railway Infrastructure and Operations

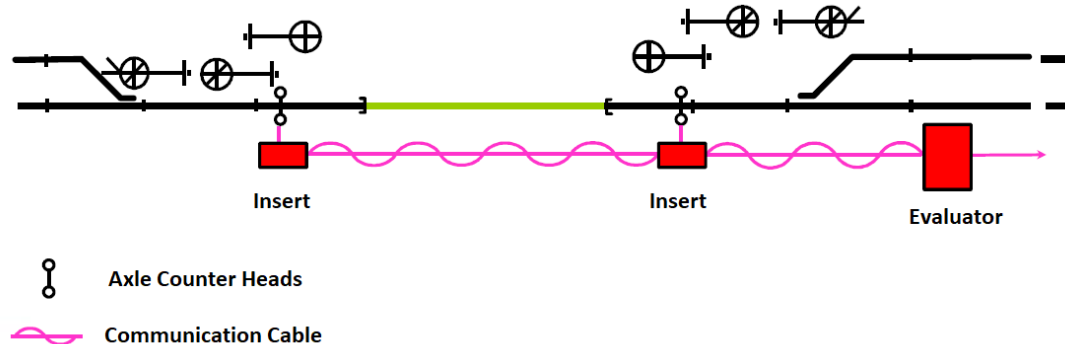


Figure 2.24: An axle counter layout (Ostrofsky, 2017)

to the number of axles passing the axle counter head. The axles are counted as the train enters the section and again when it exits. The evaluator compares the two counts and if they are equal, the section is declared unoccupied. Axle counter heads are shown in Figure 2.25, an insert is shown in Figure 2.26 and an evaluator is shown in Figure 2.27.

Interlocking

Interlocking is the equipment that combines the points, signals, track detection and other equipment to prevent dangerous or conflicting train movements from taking place. Interlocking is designed to:

- Ensure that no signals can display a proceed aspect unless all the required sets of points are detected in the correct positions and locked.
- Prevent opposing signals from being operated simultaneously.
- Prevent sets of points from being thrown until the train has moved clear of the relevant set of points.
- Expedite train movements with maximum safety.
- Increase capacity of the line.

Interlocking types include mechanical, relay, hybrids of mechanical and relay, and electronic.

2.2 Railway Infrastructure and Operations



Figure 2.25: An example of axle counter heads ([Ostrofsky, 2014](#))

Control Mechanisms

TCOs are in control of train movements. Therefore, there needs to be an interface between the TCO and the signalling system. These control mechanisms can be mechanical lever frames, push button panels or computers.

When the TCO operates the control mechanism it sends a signal to the interlocking. If the instruction is allowable, it will be executed on the train network. Figure [2.28](#) shows an illustration of the communication network between the CTC and the interlocking equipment.

A CTC diagram displays the state of the points, signals and track sections to the TCO. The diagram gives the TCO an overall picture of all the train movements under his control. Figure [2.29](#) shows a CTC push button panel and diagram and Figure [2.30](#) shows a contemporary CTC control desk.

2.2 Railway Infrastructure and Operations

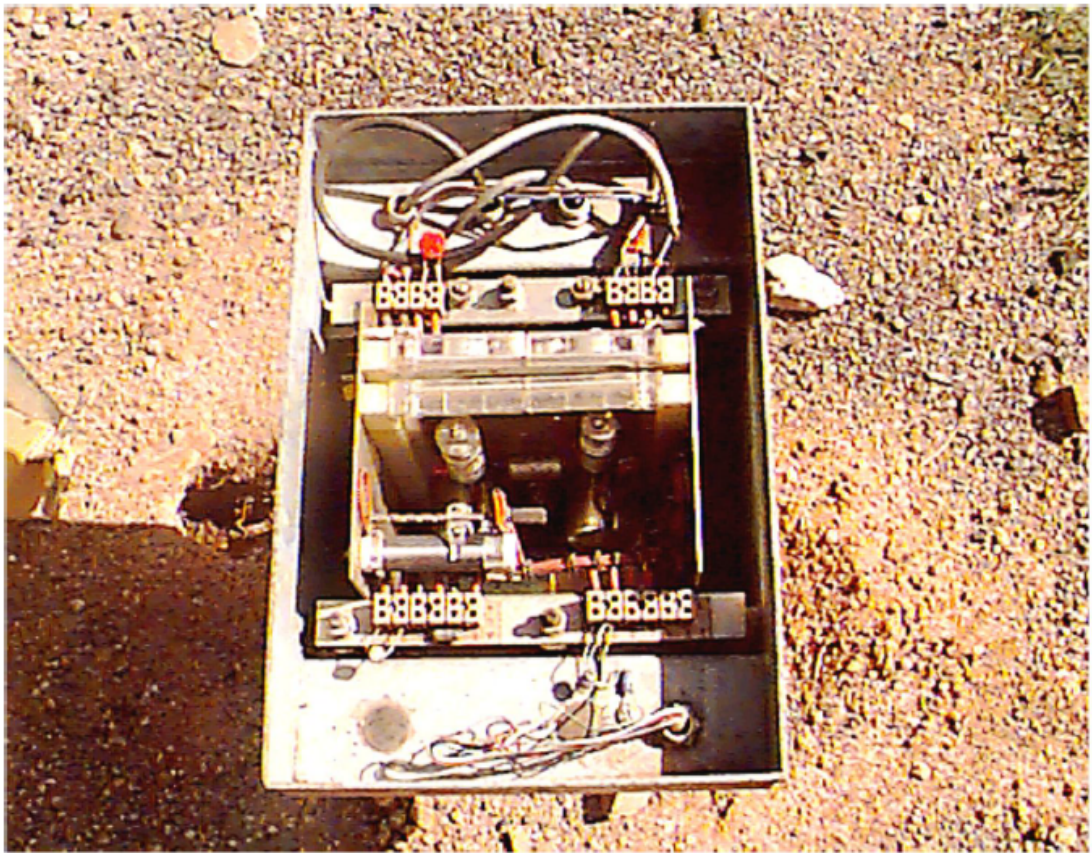


Figure 2.26: An example of an axle counter insert ([Ostrofsky, 2014](#))

2.2 Railway Infrastructure and Operations

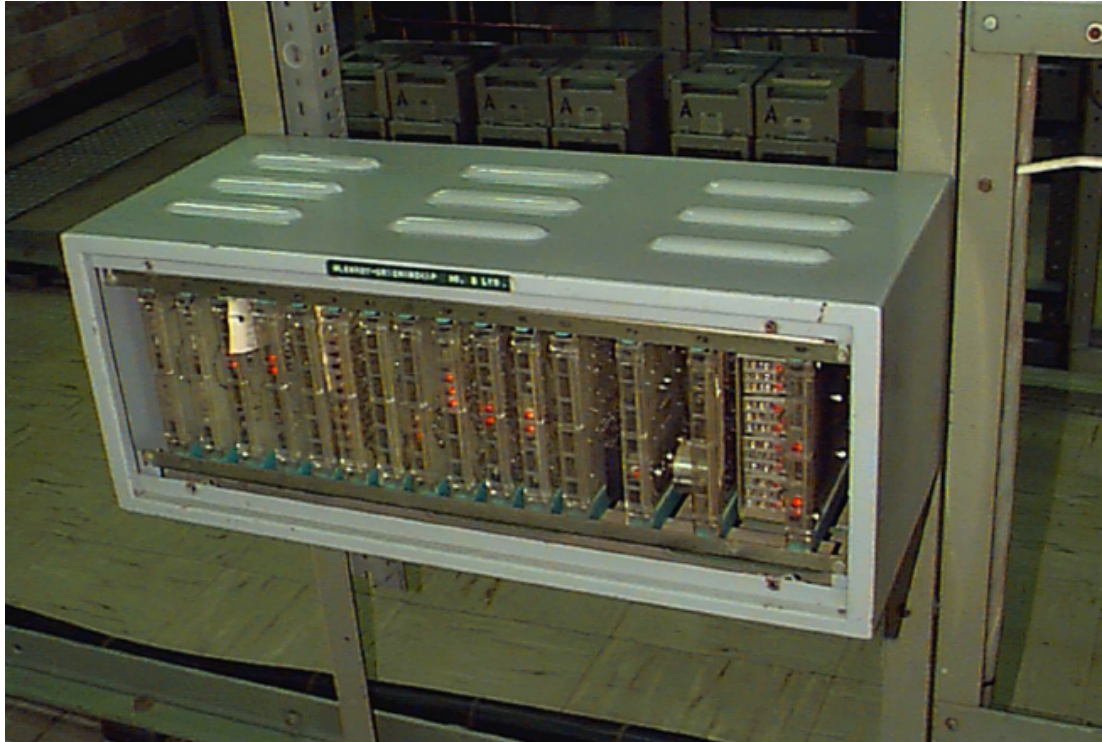


Figure 2.27: An example of an axle counter evaluator (Ostrofsky, 2014)

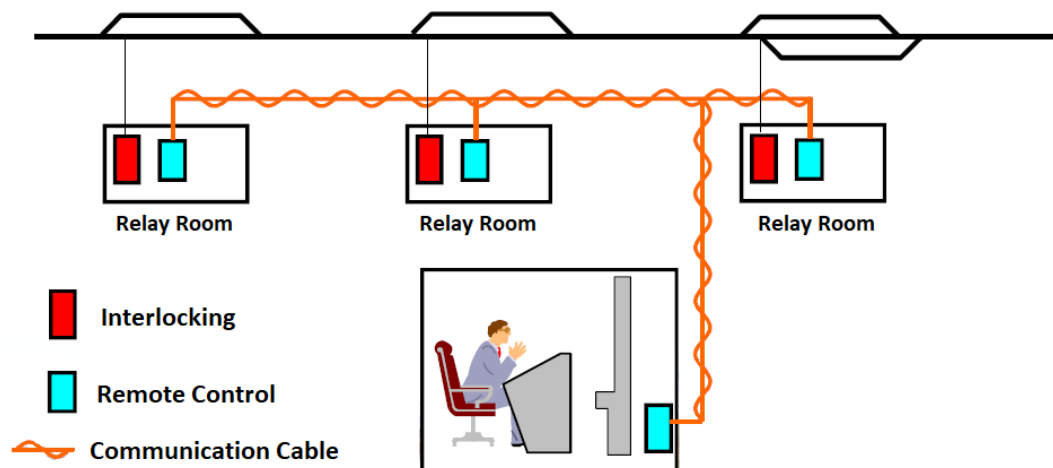


Figure 2.28: Communication network between the CTC and interlocking equipment (Ostrofsky, 2017)

2.2 Railway Infrastructure and Operations



Figure 2.29: CTC push button panel and diagram (Ostrofsky, 2017)

2.2 Railway Infrastructure and Operations



Figure 2.30: CTC control desk (Ostrofsky, 2017)

2.2 Railway Infrastructure and Operations

2.2.1.8 Condition Assessment Systems

This subsection is mostly based on [van Niekerk \(2012\)](#). Other references are cited in the text. Condition assessment system (CAS) are early warning systems installed along the track to measure infrastructure and rolling stock parameters. These systems generate alarms when predetermined limits are violated. The CAS systems also obtain operational and maintenance information that can be used for predictive maintenance.

In TFR, the CAS systems are integrated with the Integrated Train Condition Monitoring System (ITCMS). The purpose of the ITCMS is to initiate emergency procedures when hazardous conditions are detected and provide information on the condition of rolling stock ([van Niekerk, 2017](#)).

Examples of infrastructure monitoring systems are:

- Rail break detectors
- Rail stress measurement systems

Examples of track side rolling stock monitoring systems are ([van Niekerk, 2017](#)):

- Hot bearing detectors
- Wheel temperature measurement systems
- Acoustic bearing detection systems
- Lateral force measurement systems
- Weighbridges
- Weigh-in-motion systems
- Wheel profile monitoring systems
- Load profile monitoring systems
- Automatic vehicle identification systems

2.2 Railway Infrastructure and Operations

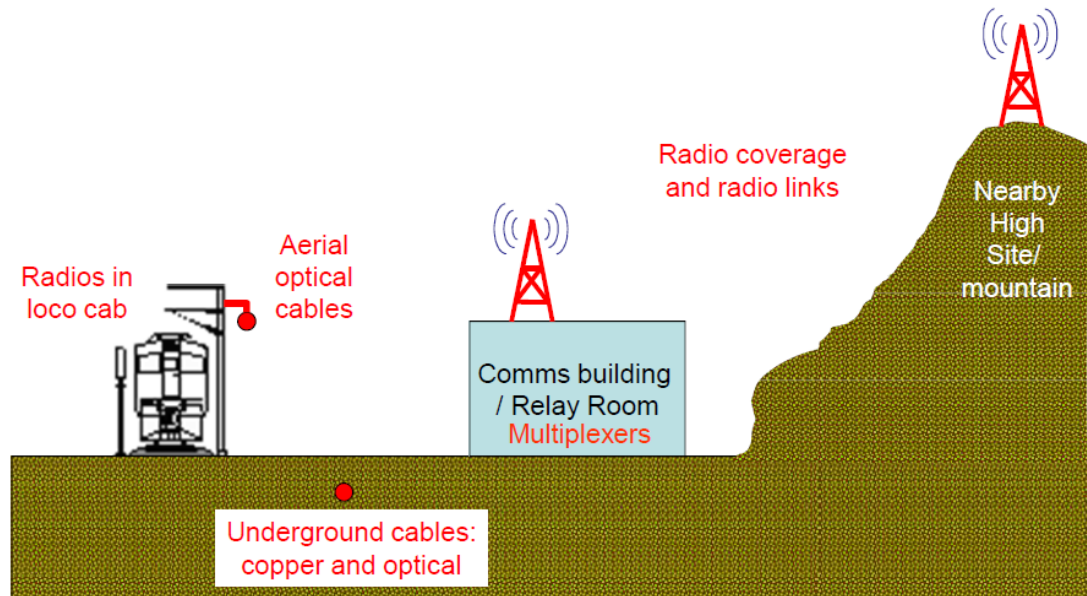


Figure 2.31: Overview of railway telecommunications infrastructure (Matseke, 2016)

2.2.1.9 Railway Telecommunication

An overview of railway telecommunications infrastructure is shown in Figure 2.31. The railway telecommunications infrastructure provides the mediums necessary for train authorisation systems, condition assessment systems, remote control of electric substations, voice and data services (Matseke, 2016).

Railway telecommunications can broadly be divided into two categories: wired and wireless. Wired technologies include: copper cables, optical fibre cable (OFC), transmission line equipment and multiplexers. Whereas, wireless technologies include: radio systems, consisting of base radios, handheld radios and repeaters, as well as microwave links between high sites (Botha *et al.*, 2016).

2.2.2 Railway Operations

Most of this subsection is based on Lombard (2010). Other references are cited in the text. A transportation service can be defined as: safely moving clients' cargo from origin to destination to arrive at an agreed time in a complete and undamaged fashion at an acceptable price in the market place, whilst making a fair profit. Railway transportation services must be designed and planned effectively, so that railways can

2.2 Railway Infrastructure and Operations

operate efficiently and make a profit. Train service planning also determines customer service levels and railway asset utilisation.

Railway operations can be divided into *train service planning* and *train operating technology*. Train service planning includes macro service planning, transportation demand analysis, rail operating strategies, wagon fleet management, locomotive fleet management, train crew management and train service scheduling. Whereas, train operating technology includes train design elements such as locomotive types, tractive effort characteristics, wagon types, train composition, braking characteristics, train speeds, and the affect of track topography, such as gradients, on train dynamics and locomotive loads.

Train service planning is directly related to the thesis topic, *i.e.* improving the integration between train service schedules and fixed infrastructure maintenance schedules. Whereas, train operating technology is not directly related to the thesis topic and will therefore not be discussed. The elements of train service planning related to train service scheduling will be discussed in the following subsections. While, the resource management elements such as crew, locomotive and wagon fleet management fall outside the thesis scope and will also not be discussed.

2.2.2.1 Macro Service Planning

Macro service planning is normally done on an annual basis to determine the expected activity levels and resource requirements. From the activity levels and resource requirements, the expected income and expenditure can be determined. These forecasts then form the main inputs for the annual budget. The macro service planning process includes the following steps: rail demand forecasting, transferring the demand onto the rail network, determining the network activity requirements and finally, resolving the network and resource constraints.

Rail Demand Forecasting

The rail demand forecast forms the base of the macro service planning process. Historical transportation trends, transportation needs provided by customers, macro-economic indicators, and other factors that may be applicable are used to forecast the expected rail transportation demand. The demand forecast is based on origin-destination groupings that are independent of the railway network layout and its capacity constraints. A

2.2 Railway Infrastructure and Operations

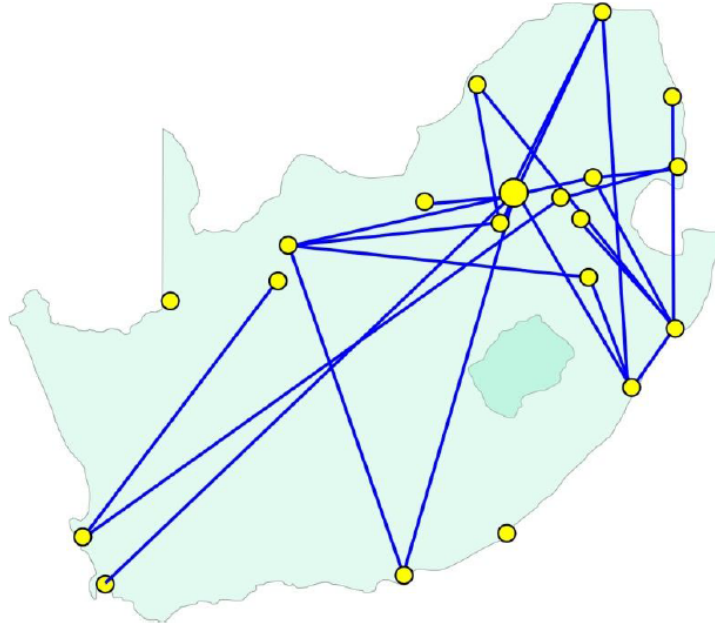


Figure 2.32: A hypothetical rail demand forecast based on origin-destination groupings (Lombard, 2010)

hypothetical rail demand forecast based on origin-destination groupings is illustrated in Figure 2.32.

Transferring the Demand onto the Rail Network

The rail demand forecast is transferred onto the rail network to determine the expected transportation demand on each railway route. The demand is transferred in line with the operating strategies of the railway. A hypothetical example is shown in Figure 2.33. In the figure, line widths reflect the demand volumes.

Determining Network Activity Requirements

From the demand forecast per route, the network activity requirements are determined for the production elements. These elements include mainline trains, train marshalling activities, and maintenance. Mainline train requirements include the number and type of trains, the number of locomotives and wagons, as well as the number of train crews. Marshalling activities refer to the staff and equipment requirements in train compilation

2.2 Railway Infrastructure and Operations

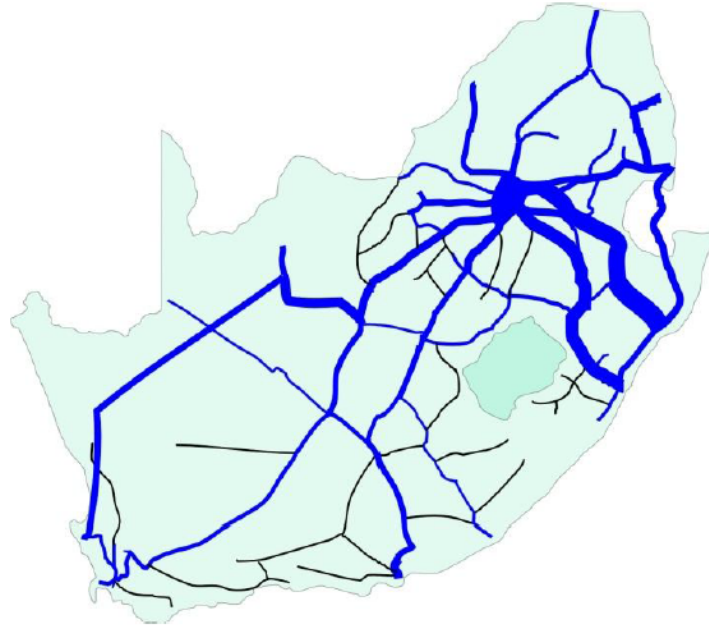


Figure 2.33: A hypothetical example of transportation demand transferred onto the rail network (Lombard, 2010)

yards. Whereas, maintenance refers the infrastructure and rolling stock maintenance workload.

Resolving the Network and Resource Constraints

Capacity constraints can be identified by comparing the network activity forecast to the available resources. These constraints may include infrastructure, rolling stock and train crew capacity. Some of the tactical options to resolve these constraints include rerouting traffic to less congested lines, reallocating rolling stock resources to different areas or changing the crewing strategies. These constraints will also inform the long-term network planning decisions related to network expansion, rolling stock acquisition and employment strategies.

2.2.2.2 Characteristics of Rail Transportation Demand

The nature of rail transportation demand must be understood to successfully plan the train service. The ideal rail demand would be constant over time. In reality, the

2.2 Railway Infrastructure and Operations

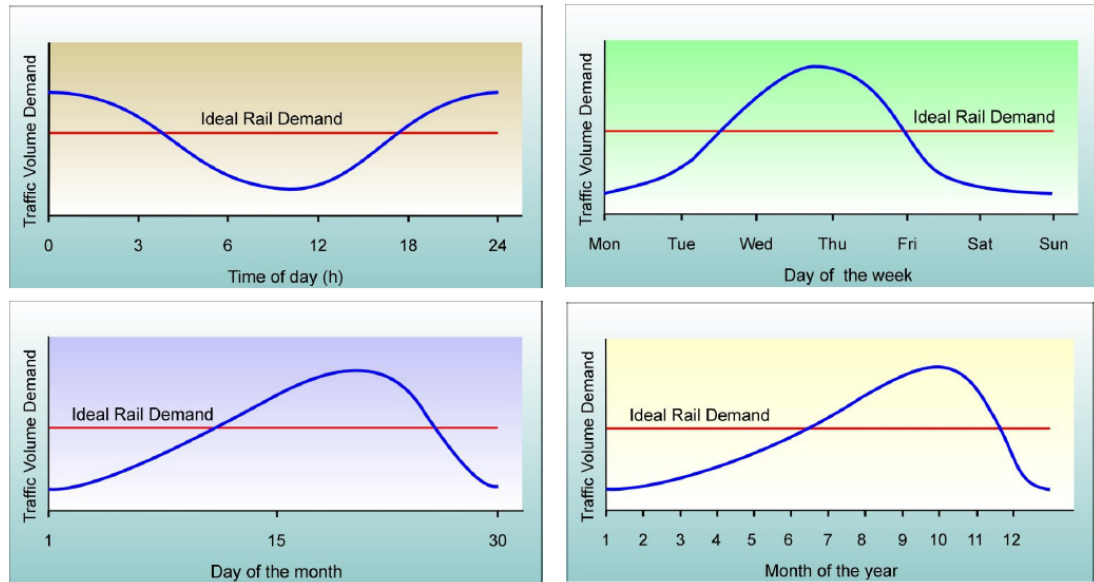


Figure 2.34: Rail demand over time (Lombard, 2017)

demand is not constant over time. One of the challenges facing railways is to plan its rail services to meet the variations in demand.

Rail demand varies with the time of day, day of the week, day of the month and month of the year. These demand variations are illustrated in Figure 2.34. Time of day variations are caused by the working hours of freight customers, while passenger trains peak during the commuter rush hours. Day of the week variations are caused by the workweek of rail customers. Monthly variations are caused by the financial processes of commercial entities as well as the disposable income of passengers. Lastly, yearly demand rises steadily from the middle of January until late in November.

Other factors such as seasonal trends and international markets also influence demand characteristics. Seasonal trends influence the demand for agricultural produce and passenger travel. Whereas, the demand for export commodities are influenced by changes in international markets such as exchange rate fluctuations, politics, wars and technological developments. Exceptional climate related events could influence demand for commodities. For example, droughts may lead to food shortages and increase the demand for grain products in the affected areas.

2.2 Railway Infrastructure and Operations

2.2.2.3 Transportation Requirements

Specific customers or commodities can have specific transportation requirements. For example:

- High value manufactured goods may require rapid transit times and adherence to strict delivery schedules as a part of just-in-time production systems.
- Perishable goods, such as wine or agricultural produce, may need to travel overnight to avoid the high daytime temperatures.
- Hazardous materials may have to be routed to avoid densely populated areas.
- Special security measures are necessary for transporting explosives.

2.2.2.4 Rail Traffic Categories

Rail traffic can be divided into the following categories: passenger, bulk and general traffic. Examples of the three categories and container traffic are shown in Figure 2.35.

Passenger Trains

Regular passenger trains are operated on fixed schedule and run irrespective if they are occupied to full capacity or not. These regular passenger trains form a stable element of the overall train service. Irregular passenger trains include seasonal trains during holidays and special trains for large events. Practical passenger train lengths are determined by the historical demand. While, the physical train lengths are limited by the passenger platform lengths and technical limitations of the train operating technology.

Bulk Traffic

Bulk traffic refers to the transportation of large volumes of raw or semi-processed materials. These commodities are transported between fixed loading and offloading locations. After offloading, the trains normally return empty to their origin.

Bulk traffic can be regular or irregular. Regular bulk traffic shows little variation over time which results in a constant demand for rail resources. This scenario is ideal because rail resources can be planned for maximum utilisation. There are normally stockpile facilities at the origin and destination that allows the trains to load and offload

2.2 Railway Infrastructure and Operations



Figure 2.35: Categories of rail traffic (Lombard, 2017)

2.2 Railway Infrastructure and Operations

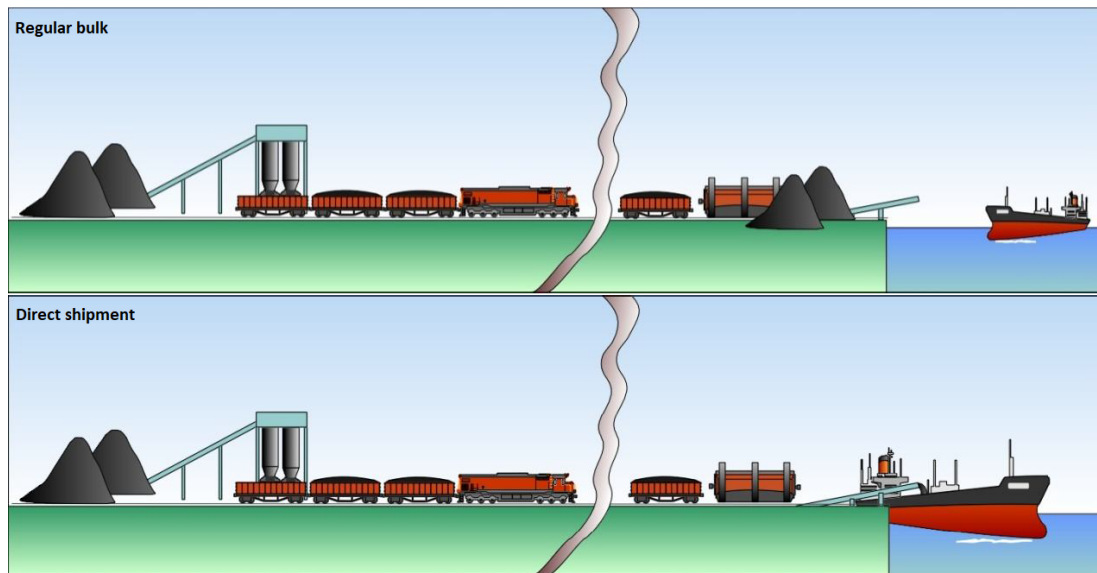


Figure 2.36: Illustrations of regular bulk traffic and direct bulk shipments (Lombard, 2010)

independently from the producers production process. It is also possible to bypass the harbour stockpiles and load directly onto a cargo ship via the mechanical offloading equipment and conveyor systems of the port terminals. Illustrations of regular bulk traffic and direct bulk shipments are shown in Figure 2.36.

Irregular bulk traffic, for example seasonal produce such as grain, shows large variations in volume over time. The demand for trains peak when a cargo ship is being loaded in the port. Rail resources are put under pressure in periods of high demand, but are underutilised in periods of low demand. This situation is made worse when purpose built wagons are needed to transport the commodity.

Bulk trains are usually designed for maximum mass and length within the technical limitations of the train operating technology and infrastructure. There must be a good fit between the train size and the capacity of the train loading and offloading equipment to ensure that the overall process is optimal.

General Traffic

General traffic is broad classification for all traffic other than passenger and bulk traffic. Customer parcel sizes are normally less than a full train load. Trains are formed by

2.2 Railway Infrastructure and Operations

coupling different parcels and wagon types together, according to the prevailing train configuration on a main route. At consolidation yards, these trains are again marshalled into different train configurations depending on the next destinations or delivery points. The train handling activities are time consuming and make the transit times relatively long as compared to other forms of transportation.

General traffic train lengths are determined by the demand, and may vary from day to day. The maximum train length is constrained by the limits of the train operating technology and infrastructure.

2.2.2.5 Railway Operating Strategies

In many cases, railway operating strategies evolve over time. In some cases, custom built railways are built for a specific operating plan. Factors such as the network layout, position of stations, marshalling yards and maintenance depots limit the operating options. The production plan must be tailored to the installed technologies, asset conditions, maintenance capabilities and supporting infrastructure such as information systems.

The major components of a rail operating strategy are the train marshalling, train movement, rolling stock deployment and the train crew deployment strategies. Train marshalling and movement will be discussed next. Rolling stock and train crew deployment fall outside the thesis scope and will not be discussed.

Train Marshalling

Train marshalling, or shunting, refers to the organisation of traffic for consolidation and distribution purposes. Marshalling is costly because it is labour intensive, utilises expensive equipment and increases the train throughput times. Three common marshalling strategies are pick-up trains, organised blocks and hub-and-spoke.

Pick-up trains depart from a major depot, then collect, and sets out the wagons for each consignment at the intermediate stations and yards on the way to the next major depot. Pick-up trains operate on routes with low traffic volumes and frequencies.

Organised blocks marshalling is similar to pick-up marshalling, except that the wagons are pre-organised at the intermediate stations and yards into customer consignment blocks per destination. This reduces the marshalling requirements on the way.

2.2 Railway Infrastructure and Operations

In hub-and-spoke marshalling, the rail network is organised into major hubs per area depending on the traffic density. Consignments originating near, or between, hubs first travel to the nearest major hub. There, the consignments blocks are combined to form trains that travel between hubs. To reach further destinations, consignment blocks are switched between trains at the hubs.

The objective of hub-and-spoke marshalling is to minimise train marshalling on the way, but it can increase the overall locomotive and wagon kilometres. It may lead to consignments travelling past their origins.

Train Movement

Train movement refers to the methods used to move trains through a network. These methods include locomotive optimisation, *ad hoc* planning, traffic push, scheduled and traffic corridors. Combinations of these methods are also used.

In the case of locomotive optimisation, trains are designed so that locomotives are utilised optimally. This implies that the train size is adapted from the origin, and potentially at intermediate stations, to ensure that the locomotives work at optimum loads. Since trains could be broken into smaller wagon blocks, the possibility exist that a customer consignment arrives in fragments.

The *ad hoc* train movement strategy is reactive in nature. Only once a transportation request is received from a client are the resources planned and the service executed. *Ad hoc* movements are normally used by small railways with basic information systems, low traffic volumes and low service levels.

The traffic push strategy is used in cases where there are high traffic volumes, high capacity and where the precise estimation of arrival times is not critical. The marshalling strategy governs the movements from the origins to the intermediate stations. From the intermediate stations, trains are compiled of wagons that have to travel in the same direction. The process continues from hub to hub until each wagon reaches its final destination. The expected transit times for the push strategy are determined statistically.

In scheduled train movement, all train and wagon activities are planned. Demand fluctuations are accommodated by scheduling contingency capacity on the production plan. The daily train plan is derived from the base plan in advance. The daily train

2.2 Railway Infrastructure and Operations

plan links customer orders to wagons, wagons to trains, and trains to the schedule. Locomotives and train crews are then allocated to trains on the schedule.

A scheduled train service requires information systems, good communication between stakeholders and disciplined execution. The train service is predictable and therefore delivery times can be guaranteed. Clients are also required to load and offload trains within a specified time. Intermediate stations can absorb minor variations in the plan. Scheduled train services lead to a high level of asset utilisation.

Train movements can also be planned in traffic corridors. Major traffic flows are identified between major commercial centres. Unit trains, *i.e.* trains transporting one consignment, are compiled at the origin and run through regional and geographic borders, without marshalling on the way, straight to the final destination. These trains operate uninterrupted for long distances that reduce the overall costs and improve the resource utilisation.

Railways enter into contracts and service level agreements with its clients. These agreements assist the railway to achieve the desired train turnaround-times. Turnaround-time is the time elapsed from when a train is loaded and ready for departure, moved to the offloading destination, offloaded, moved to the next loading destination, placed for loading, loaded and ready to depart again. The agreements prescribe the loading times, offloading times, schedule and penalties for non-performance.

2.2.2.6 Train Diagrams

A train diagram is a tool used in train service design, planning and monitoring. It is a visual representation of the position of trains over time on particular route. An example of a train diagram is shown in Figure 2.37.

The diagram heading indicates the route, in this case stations A to I. The horizontal axis depicts time with equally spaced vertical lines. Normally, train diagrams represent a 24 hour period. The vertical axis shows distance according to scale with schematic layouts of track features such as stations, sidings, crossing loops, single and double lines. Other information such as point-to-point running times and train numbers may also be included. Lines on the diagram show the movement of trains.

Figure 2.37 depicts a two hour period. Train number 0011 is travelling from Station A towards Station I at a speed of 60 km/h. It departs from Station A at 00:00 and

2.2 Railway Infrastructure and Operations

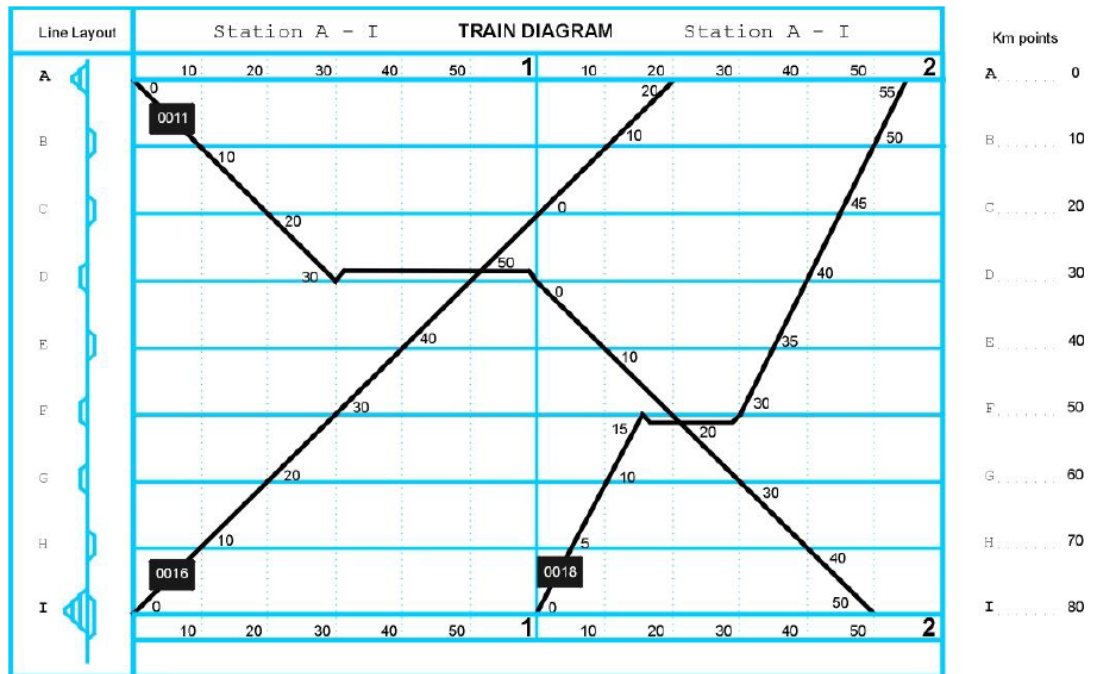


Figure 2.37: An example of a train diagram (Lombard, 2010)

arrives at Station I at 01:50. At 00:30, train 0011 arrives at Station D and waits for train 0016 to pass. At 01:00, train 0011 departs again for Station I.

Train number 0016 is travelling from Station I towards Station A at 60 km/h. It departs from Station I at 00:00 and arrives at Station A at 01:20.

Train number 0018 is travelling from Station I to Station A at 120 km/h. It arrives at Station F at 01:15, waits for train 0011 to pass, and then continues to Station A at 01:30. Finally, train 0018 arrives at Station A at 01:55.

2.2.2.7 Rail Capacity

Rail capacity can be defined as the number of trains that can be run on a section of a track or a rail network over a defined period. An understanding of rail capacity is important to optimise train designs and service planning.

The capacity of a railway is influenced by the individual characteristics and capacities of the rail infrastructure, rolling stock, human resources and auxiliary equipment. The factor with lowest capacity will limit the overall capacity.

2.2 Railway Infrastructure and Operations

The infrastructure capacity is determined by the characteristics of the following features, for example:

- Axle loads, line speeds and layout of the track.
 - The line layout, *i.e.* whether it is a single or double line, has a major impact on the line capacity. The capacity of a double line is dependant on the section with the maximum time between trains, *i.e.* the headway between trains, on a particular route. Whereas, the capacity of a single line is determined by the section with the longest train running time between crossing loops on a particular route.
 - From the previous point, it follows that the number, distance between and length of crossing loops influence capacity.
- Location, capacity and facilities of terminals and depots
- Voltage and capacity of the overhead track equipment
- Capacity, speed and integrity of the movement control and authorisation systems and procedures. These systems determine the following distance between trains, and also the time it takes to issue an authority and trains to cross.

The rolling stock capacity is determined by:

- The size of the locomotive and wagon fleets
- Axle loading
- Power of locomotives
- Payloads of wagons
- Availability of enabling technologies such as radio distributed power or electronically controlled pneumatic brakes

Human resource capacity is determined by the quantity, proficiency and location of train crews, train control officers, technical support and safety personnel. Whereas, the capacity of auxiliary equipment is determined by the quantities of equipment such

2.2 Railway Infrastructure and Operations

as telemeters or tarpaulins that may be required to operate a particular service. Differences in the speed of trains also have an effect on line capacity, especially on single lines.

2.2.2.8 Train Service Scheduling

Train service scheduling is the process of combining all the relevant production plan factors such as the demand forecasts, traffic requirements, train marshalling, train movement and resource management strategies into a train schedule. Traditionally, train schedules are developed by drawing train diagrams by hand. In recent years, computer based support has been developed for the scheduling task, but the traditional method is still common.

The magnitude of the scheduling task depends on the number of trains, size and complexity of the network, operating strategies, interdependency of different schedules and the availability of the required information. The following information is required to develop the train schedule:

- A network description in the form of a schematic layout of the rail network that includes all sidings, stations, marshalling yards and rail links with their positions and distances.
- Track information such as gradients, kilometre points, signal positions, the positions of crossings, single and double lines.
- Point-to-point train running times. Typical point locations include the centre of stations, the position of signals or crossover points.
- The train authorisation method and the affect it has on the train movement.
- Train design information, such as the length of trains to schedule crossings.
- Train priorities for scheduling purposes. Train priorities will depend on the type and purpose of every train. Train prioritisation affects the transit times, crew and rolling stock utilisation as well as line capacity.

2.3 Maintenance

Most of this section is based on Mitchell (2014). Other references are cited in the text. Maintenance forms part of the Physical Asset Management (PAM) framework that is concerned with maximising asset effectiveness. Therefore, the principles of PAM are introduced in Subsection 2.3.1 before maintenance specific concepts are presented from Subsection 2.3.2 onwards.

2.3.1 Physical Asset Management

The term *Asset Management* is used in the financial, real estate and construction fields, amongst others, and it generally refers to maximising the utilisation and return on assets, usually financial assets. While, the term *Physical Asset Management* (PAM) is used by manufacturing and service organisations to refer to the lifetime and performance management of physical assets. Mitchell (2014) prefers the term *Physical Asset Optimisation* when referring to PAM.

The Institute for Asset Management defined a framework for PAM. It is known as the Publicly Available Specification 55. The framework defines PAM as “*systematic and coordinated activities and practices through which an organisation optimally and sustainably manages its assets and asset systems, their associated performance, risks and expenditures over their whole life cycle for the purpose of achieving the organisational strategic plan*”. The typical priorities and concerns of a PAM framework are shown in Figure 2.38 (Vlok, 2014).

Mitchell (2014) defines Physical Asset Optimisation as a “*comprehensive, fully integrated strategic programme directed to safely gaining and sustaining greatest lifetime value, utilisation, productivity, effectiveness, profitability and return on assets from physical manufacturing, production, operating and infrastructure assets*”.

2.3.1.1 Asset Lifetime

Asset lifetime can be defined as the span of time over which the asset is designed, acquired and utilised to fulfil its intended purpose, including end-of-life disposal. Asset lifetime is normally divided into four stages:

1. specification, design and procurement

2.3 Maintenance

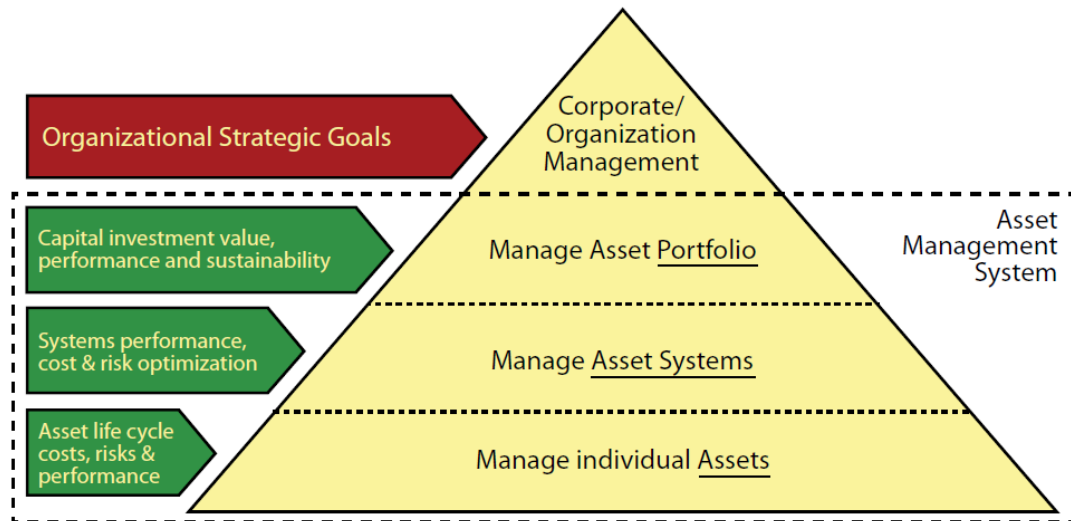


Figure 2.38: Typical priorities and concerns of a physical asset management framework (Vlok, 2014)

2. construction, installation and commissioning
3. operation
4. disposal

Lifetime cost is the total cost of ownership over the expected lifetime of an asset from specification to disposal. The first, second and fourth stages of asset life are only cost. While, the third stage is the productive and revenue generating stage of asset life. The operation stage is also the longest stage of asset life.

During the first two stages, Engineering must make sure that the asset is specified, designed, procured and installed for optimum operating life in the operation phase. Engineering for optimum operating life includes:

- Ensure that the design has optimum intrinsic reliability.
- Ensure that the correct materials are used.
- Ensure that the asset is operable.
- Ensure that the asset is designed for efficient maintenance by including common parts, ease of access and disassembly.

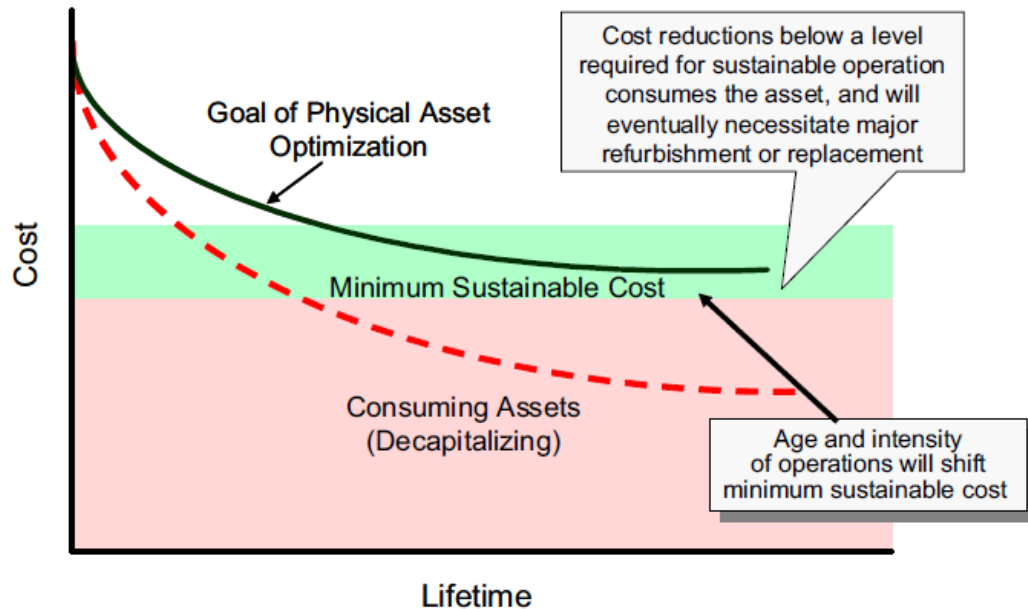


Figure 2.39: Sustainable physical asset management (Mitchell, 2014)

Industry leaders form a team of Engineering, Production, and Maintenance personnel to audit the design, construction and installation stages. This is to ensure that maintenance and reliability considerations are designed and built into new assets.

2.3.1.2 The Objective of Physical Asset Management

The objective of PAM is to maximise the financial return delivered by the physical assets that generate revenue. The objective is achieved by managing the asset lifetimes for optimum life and minimum sustaining costs.

Figure 2.39 shows that there is a minimum cost necessary to operate physical assets sustainably. Cost reductions below this minimum will lead to premature refurbishment or replacement. For example, reducing costs by deferring minor maintenance such as sealing small cracks, painting or corrosion protection may lead to the failure of a structure long before its designed end of life. Figure 2.39 further indicates that the age and intensity of operations will shift the minimum level of sustaining costs. For example, operating above specified production rates may increase equipment stress and reduce asset life.

2.3 Maintenance

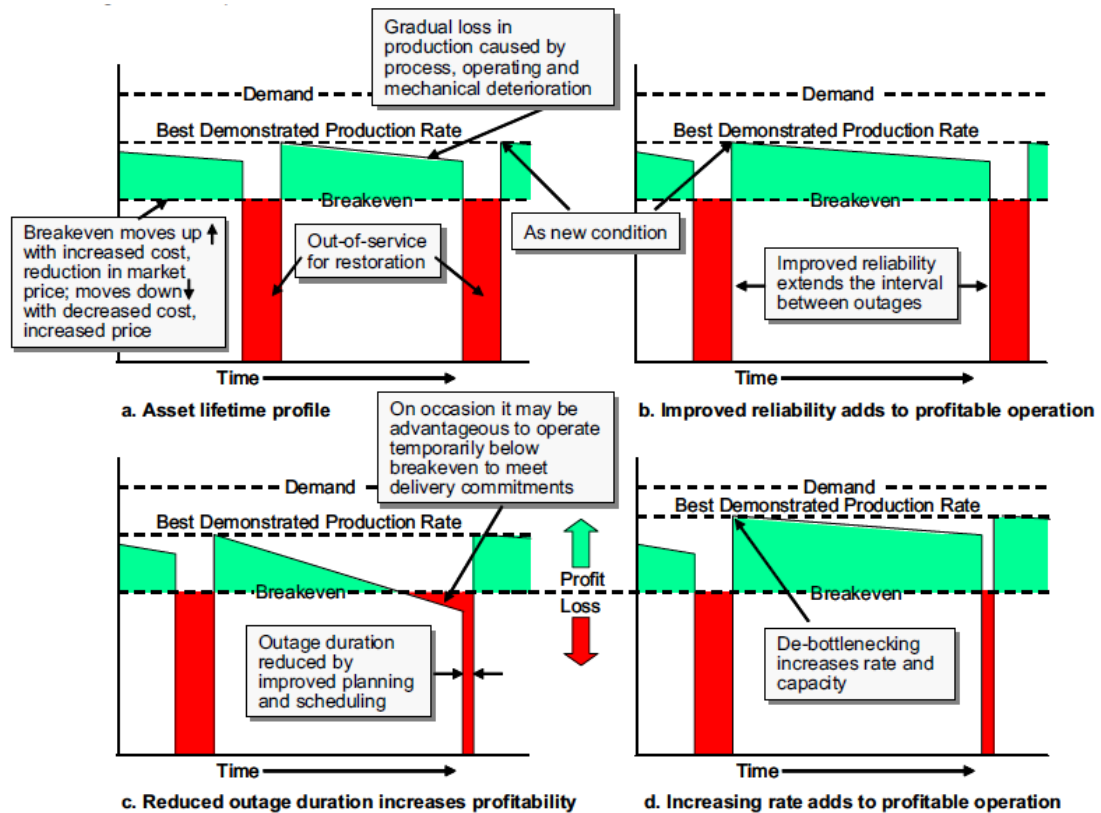


Figure 2.40: Interventions to improve asset effectiveness (Mitchell, 2014)

Deferring sustaining maintenance, such as sealing cracks or repairing insulation, often leads to short term cost reductions. The effects of deferring minor maintenance are gradual. The real costs and consequences may only become apparent once production is affected and a major failure occurs.

To maximise the profit gained from physical assets, a PAM programme must estimate the asset lifetimes and improve reliability by correcting asset defects. Asset life can be extended by improving designs, careful installation, correct operation and maintenance. Eventually, every asset must be removed from operations for maintenance.

Figure 2.40 shows the lifetime profile of typical operating and production assets. The shaded areas above the breakeven line indicate the ability of these assets to generate profit. The shaded areas below the breakeven line indicate interruptions, losses and restrictions that prevent the generation of profit.

Figure 2.40a shows that process, operating and mechanical deterioration lead to a

2.3 Maintenance

gradual decline in production rate and profit. It further shows that the breakeven point moves up or down when there are changes in the production cost or market price.

Figures 2.40b, c and d show three different interventions to increase the profits generated by assets. Figure 2.40b shows that improving reliability, by eliminating defects, extends the time between maintenance outages. Figure 2.40c shows that improving maintenance planning and scheduling can shorten the duration of maintenance outages. Figure 2.40d also shows that operating below breakeven may be justified in certain circumstances. For example, meeting delivery commitments to preserve customer relationships. Figure 2.40e shows that removing bottlenecks increases the production rate. Furthermore, improving the production rate may lead to greater capacity and profitability.

Return on Net Assets (RONA) is a financial ratio that measures corporate effectiveness. It measures how well a company is generating revenue from its assets. It can be defined as

$$RONA = \frac{EBIT}{TOTAL\ ASSETS - CURRENT\ LIABILITIES}, \quad (2.1)$$

where EBIT = Earnings before interest and tax (Correia *et al.*, 2013). Figure 2.41 shows that revenue, asset utilisation and cost have an influence on RONA. PAM can increase RONA by increasing asset utilisation and reducing cost.

2.3.1.3 Principles of Physical Asset Management

PAM strategy begins at design

Figure 2.42 shows the division of maintenance costs by origin. More than 20 percent of maintenance costs originate from engineering, design and construction.

Design decisions determine the inherent system reliability, availability and maintainability. Design characteristics, such as component life and ease of repair, have a fundamental impact on lifetime cost. Good design minimises problems, such as operating errors, and incorporates solutions to existing problems into new designs. Compromising on reliability, availability and maintainability during design results in systems that are costly to operate and maintain.

2.3 Maintenance

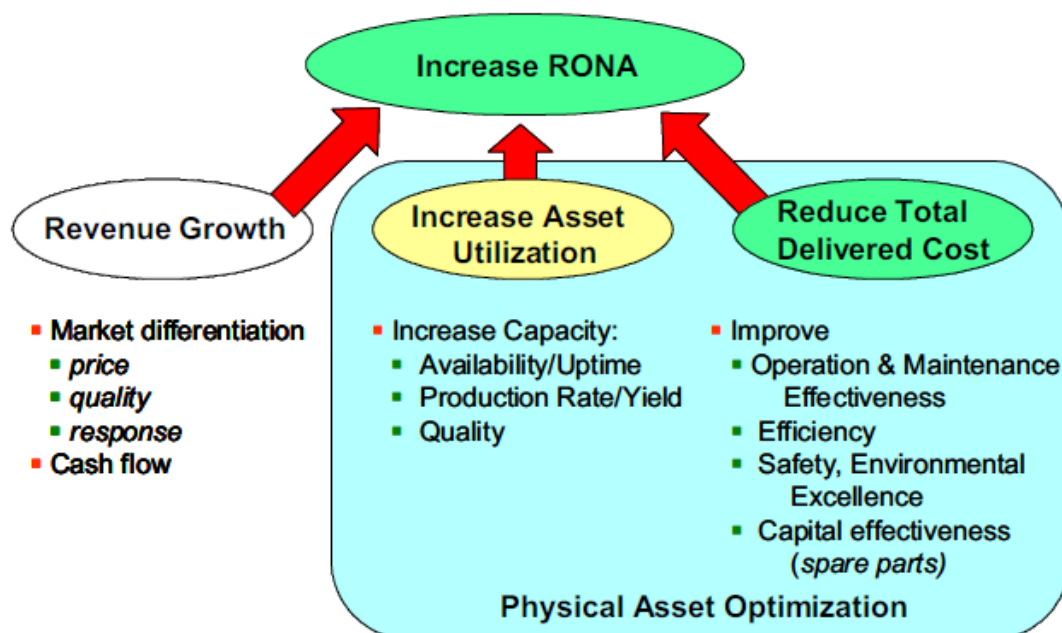


Figure 2.41: PAM can improve RONA (Mitchell, 2014)

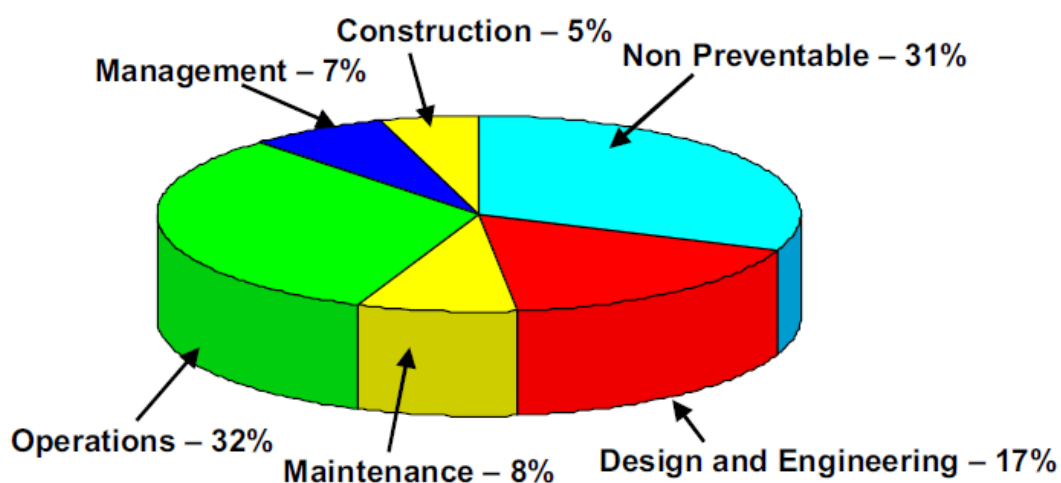


Figure 2.42: Division of maintenance costs by origin (Mitchell, 2014)

2.3 Maintenance

Procurement directed to optimising lifetime cost

Procurement decisions must be directed towards optimising lifetime cost. If procurement decisions aimed at reducing costs, for example relaxing material specifications, introduce unreliability, inefficiency, additional maintenance or lowers quality it can cause losses during operation. The long-term costs may exceed the initial savings.

Correct operation must be assured

Figure 2.42 shows that 32 percent of maintenance costs are a result of operating errors. Operating errors, and the associated maintenance costs, can be reduced by making sure that the correct operating procedures are used and by correcting off-design performance. Off-design operations can be corrected by matching design specifications, such as temperature, speed and power, to actual operating conditions.

Optimised maintenance is a necessity

Maintenance is discussed in detail in Subsections 2.3.2 to 2.3.7. PAM views maintenance as an investment in future profits. Maintenance adds value through capacity assurance, improved throughput, quality and reduced operating cost.

2.3.1.4 Reliability

Reliability is the basis of PAM. It is the probability that a system, device, component or product will perform its required functions in a satisfactory manner for a given period of time when used under specified operating conditions in a specified environment.

Reliability analysis includes lifetime prediction and failure analysis. Together, this provides the information necessary to extend asset life and avoid unplanned outages. Reliability is evaluated by statistically estimating the remaining asset life. While, failure analysis identifies the root causes of failures.

As stated in the previous subsection, design determines the maximum achievable reliability. High inherent reliability is easier to achieve with static equipment, such as piping, instrumentation and power distribution systems. Whereas, it is more difficult to achieve with mechanical equipment that can be exposed to lubrication contamination, erosion, unbalance or misalignment.

The main features of a reliability improvement programme are:

2.3 Maintenance

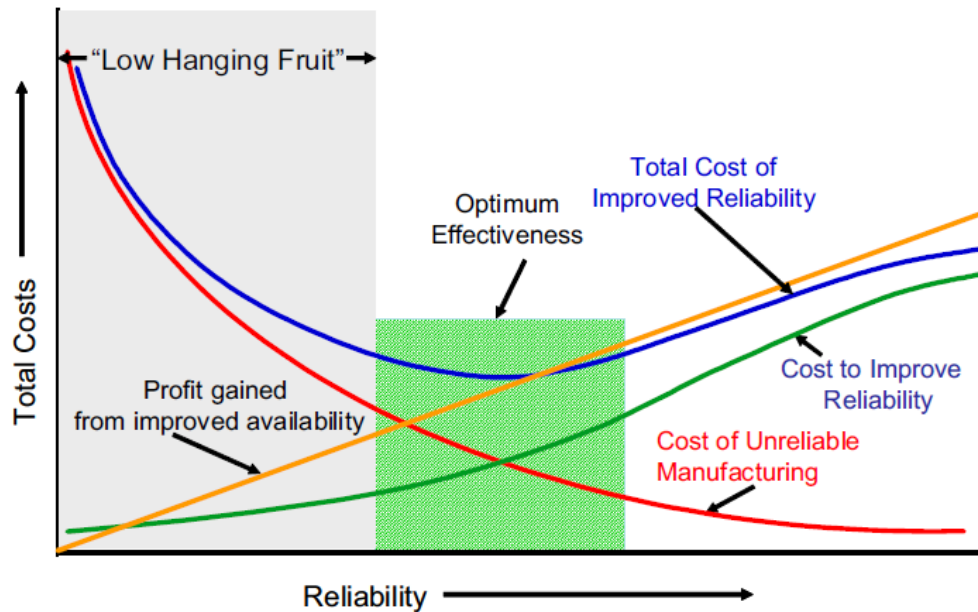


Figure 2.43: Optimum reliability cost (Mitchell, 2014)

1. Eliminate failures and the cause of failures.
2. Extend the operating lifetime of assets.
3. Reduce the cost of asset care by reducing the requirements for work.

The optimum reliability cost is shown in Figure 2.43. It is the minimum point that balances the cost of unreliable manufacturing and the cost of improved reliability. Every organisation should aim for the minimum reliability cost necessary to meet mission requirements.

Three ways to reduce maintenance spending are to:

1. Improve maintenance work management, planning, scheduling and processes. This can reduce maintenance costs by 15 to 20 percent.
2. Improve the management of maintenance spares. This can reduce maintenance costs by 10 percent.
3. Improve system, equipment and component reliability by eliminating defects. This reduces the need for maintenance labour and parts, and therefore reduces maintenance costs.

2.3 Maintenance

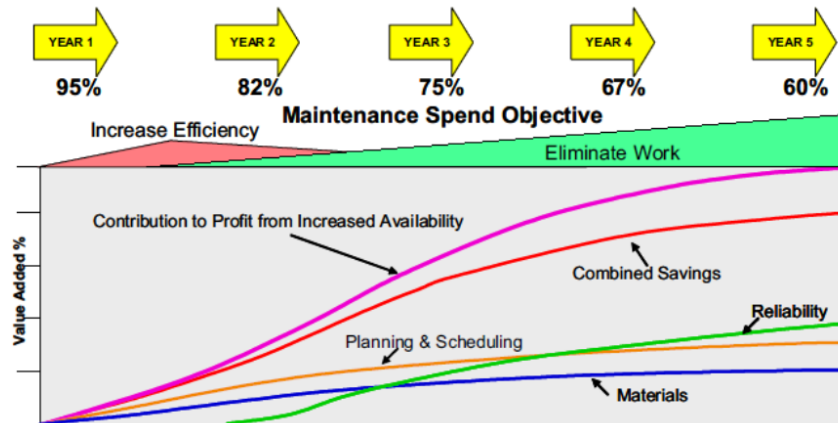


Figure 2.44: Contributions of the three elements of maintenance cost reduction (Mitchell, 2014)

The contributions of the three ways to reduce maintenance costs are shown in Figure 2.44. It shows that the early maintenance cost reductions are mostly achieved through improved efficiency as result of improvements in planning, scheduling and materials management. Whereas, the later benefits are achieved by eliminating the need for work through improved reliability. Cost reductions achieved through reliability improvements are sustainable over the long term. Whereas, cost reductions as a result of efficiency improvements slow in contribution over time.

Eliminating asset defects improves reliability. Better reliability reduces the need for maintenance labour and parts, and therefore reduces maintenance cost. Further benefits of better reliability include improvements in availability, production output and profit. Reliability is sustained by adhering to the standards and specifications, while reliability is enhanced by asset improvement projects.

2.3.2 Maintenance Introduction

Maintenance is a part of PAM. It is a core business activity that aims to obtain maximum asset effectiveness. In the past, it was viewed as inferior to production, but during the past few decades industry has started to recognise that maintenance is an essential part of production and operations (Vlok, 2014).

Maintenance is the activities involving the servicing, care, upkeep, repair or correction of an asset or components to (Duvel, 2017):

2.3 Maintenance

- Keep assets in the correct state or condition
- Prevent deterioration
- Restore performance
- Reduce the probability or consequences of failures
- Sustain assets to perform the required functions

Maintenance management (Duvel, 2017):

- Directs and integrates all maintenance functions
- Prioritises activities within a constrained budget
- Is strategic and systematic
- Involves various tactics to optimise asset performance at the lowest possible cost
- Delivers a cost-effective and safe service
- Ensures that the equipment is in a reliable, available, affordable and safe condition

Maintenance can be controlled and planned for maximum value, but it cannot be deferred for too long or be ignored. Deferred costs will return later with a greater financial and operational impact.

2.3.3 Maintenance Tactics

Maintenance actions are either planned or unplanned. Planned maintenance is scheduled prior to execution, for example a week before a task is done. Whereas, unplanned maintenance is reactive, done once there is a fault with the equipment.

Figure 2.45 shows the different maintenance tactics. *Corrective maintenance* can be unplanned or planned. *Time-based maintenance* (TBM), *condition-based maintenance* (CBM) and *proactive maintenance* are planned. TBM and CBM are also classified as *preventive maintenance*. Maintenance progress is achieved by moving from mostly unplanned maintenance to mostly planned maintenance and reliability improvement. The best practice is to plan 85 to 90 percent of maintenance, while the rest would be unplanned corrective maintenance.

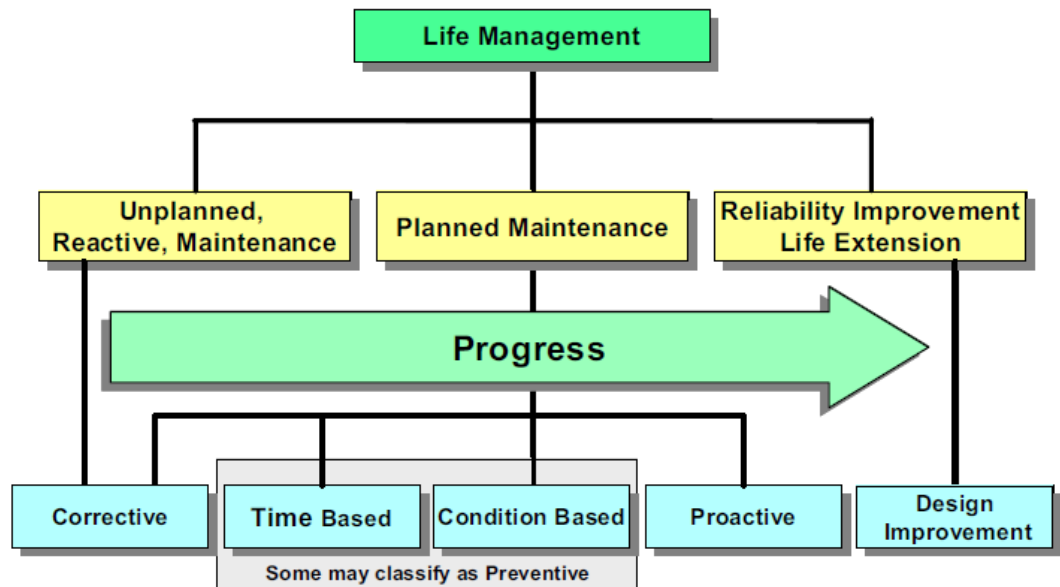


Figure 2.45: Maintenance types (Mitchell, 2014)

The aim of preventive maintenance is to prevent unexpected failures. Corrective maintenance is more expensive than prevention because of the costs of secondary damage to equipment, production losses, late delivery penalties, overtime labour and safety incidents that result from failures. For these reasons, preventive maintenance is generally more cost effective than corrective maintenance (Vlok, 2014).

Preventive maintenance intervenes with equipment before failure occurs and therefore limits the component or asset life. For this reason, the timing of preventive maintenance must be optimised. A trade-off exists between residual life and the risk of unexpected failure. Intervening too soon results in the waste of residual life. Whereas, intervening too late increases the risk of unexpected failure. The timing of the intervention must therefore be chosen so that an acceptably little amount of residual life is wasted, while the risk of unexpected failure does not increase too much (Vlok, 2014).

Preventive maintenance is not always the most effective maintenance tactic, even though it is the most effective tactic in most cases. The technical and economic feasibility of a maintenance tactic must be assessed before implementation. Methodologies such as *Reliability Centred Maintenance* (RCM) or *Total Productive Maintenance* (TPM) are used to determine which maintenance tactics apply. An overview of RCM and TPM are presented in Subsections 2.3.5 and 2.3.6, respectively (Vlok, 2014).

2.3 Maintenance

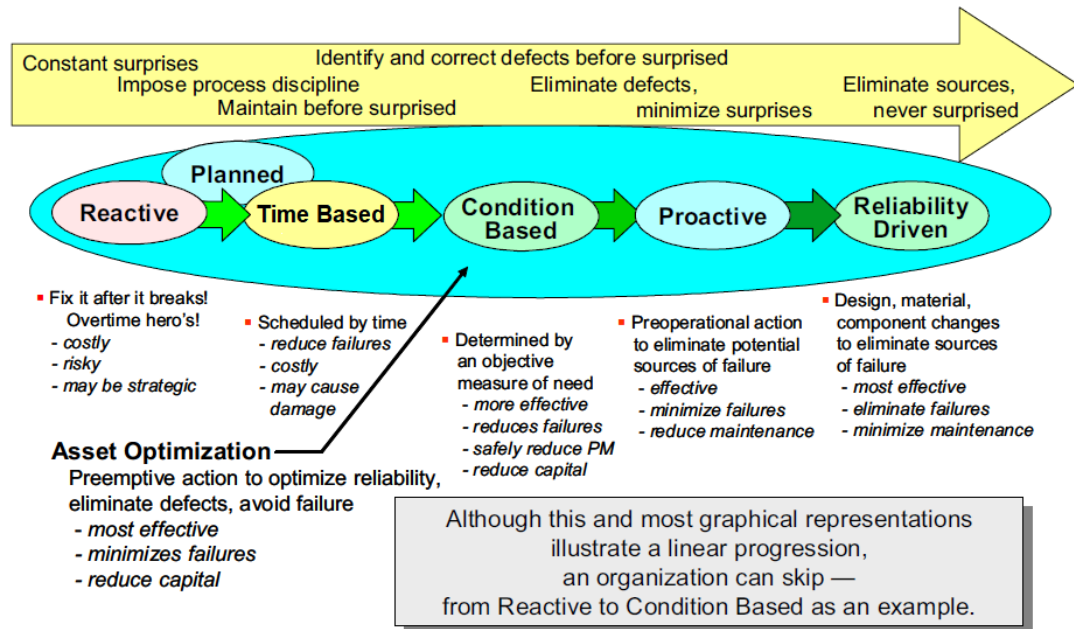


Figure 2.46: Evolution of equipment management (Mitchell, 2014)

Maintenance has evolved from a “fix it when it breaks” approach to planned, proactive and reliability driven maintenance as shown in Figure 2.46. Moving from reactive to condition based maintenance has the advantages of increasing effectiveness and reducing cost, as shown in Figure 2.47.

2.3.3.1 Corrective Maintenance

Corrective maintenance is corrective action taken on failure or obvious, unanticipated threat of failure. It consists of problems identified by Production, such as loose bolts, leaks, worn out insulation, excessive temperatures, noise or equipment that is out of calibration.

Corrective maintenance can be planned or unplanned. If the problem is not urgent, corrective maintenance is planned for later. However, if the problem is urgent it may require immediate intervention, *i.e.* unplanned work.

Unplanned corrective maintenance is the least effective and most expensive form of maintenance. It must be minimised for greatest effectiveness and least cost. Nevertheless, there are instances where it does make technical and economic sense to implement a run-to-failure maintenance tactic, as may be the case for easily accessible lightbulbs.

2.3 Maintenance

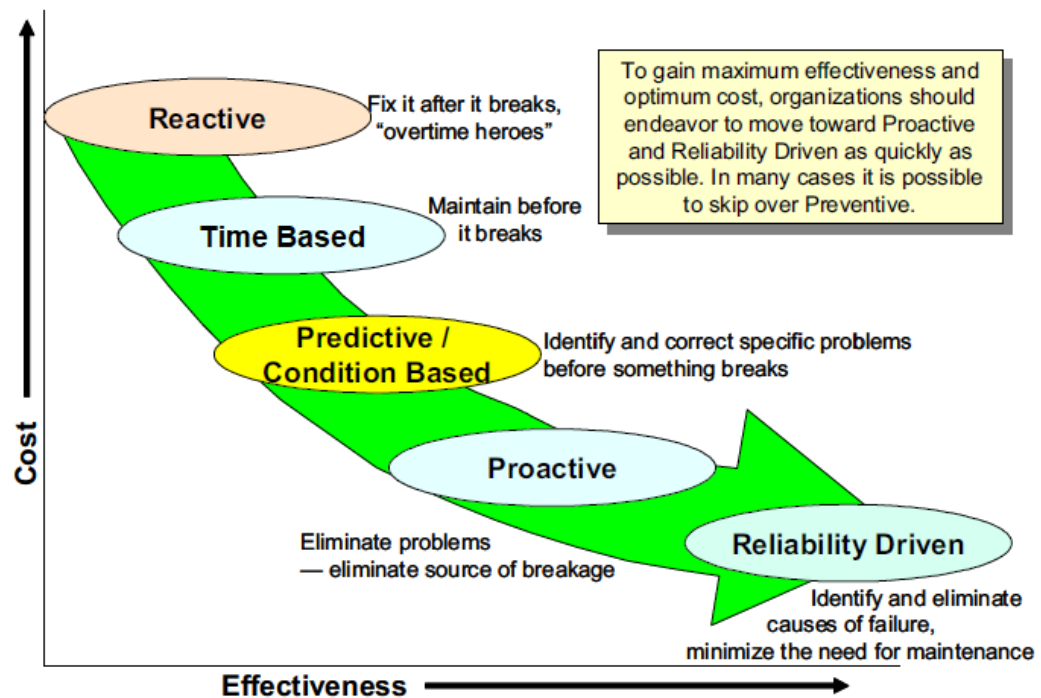


Figure 2.47: Cost advantages of maintenance types (Mitchell, 2014)

2.3 Maintenance

General side effects of unplanned corrective maintenance are that it is more expensive than prevention, potentially dangerous, interrupts production and may lead to secondary damage.

2.3.3.2 Time-Based Maintenance

Time-based maintenance (TBM) is maintenance tasks, including inspection, service or replacement, conducted at regular, scheduled intervals of calendar or operating time established to avoid failure based on average statistical or anticipated lifetime. Generally, TBM is invasive. It requires outage and disassembly for visual inspection, overhaul and replacement regardless of the condition.

The intervals of TBM are based on average life based on time, *e.g.* weekly, monthly or quarterly, or service time, *e.g.* hours of operation. The intention of TBM is to ensure that there is a high probability that performance and material condition are maintained within the required limits. However, unexpected failure may still occur when the difference between average and minimum life is large or external and localised conditions affect lifetime.

To minimise the risk of failure, those responsible for determining TBM intervals typically select intervals on the low side of average lifetime. This results in equipment being over maintained and components being replaced that are in good condition with substantial life remaining.

Intrusive TBM may force equipment back into an infancy failure regime and introduce errors from personnel, reassembly or procedural mistakes. For these reasons, TBM can be unnecessary, expensive and labour intensive. TBM intervals must be optimised or it must be replaced with CBM.

2.3.3.3 Condition-Based Maintenance

Condition-based maintenance (CBM) is maintenance tasks based on actual condition obtained from in-place non-invasive tests, as well as operating and condition measurements. CBM can usually identify abnormalities for correction early enough to minimise the risk and impact of operational interruptions. Best practice reliability improvement programmes combine CBM and Root Cause Analysis (RCA) to eliminate defects. This leads to reliability improvement and is essential to physical asset management.

CBM consists of three activities:

2.3 Maintenance

1. Condition measurement
2. Condition monitoring and assessment
3. Repair and maintenance actions

Condition measurement is non-invasive measurements that determine mechanical and operating condition. For example, measuring vibration, temperature, pressure, acoustic or electrical characteristics. These measurements may be recorded continuously or periodically. Measurements can also be taken individually or collectively depending on the equipment.

Condition monitoring and assessment is the comparison of individual and collective condition measurement values and operating data over time to determine the current condition, identify and analyse defects and estimate the remaining life of assets. Condition monitoring also identifies degradation mechanisms. Once the cause of degradation is understood, the degradation can be eliminated or controlled before the asset condition breaches specifications.

The repair and maintenance actions are based on the objective evidence gathered during condition monitoring and assessment. Furthermore, RCA is used to identify the root cause and necessary actions to eliminate defects. The CBM actions are only complete once the design, operations and maintenance recommendations from the RCA are implemented to prevent problems.

The predictive information gathered through CBM is valuable:

- The early warning of failures minimises the risk of failures, safety and environmental hazards.
- A reduction in failures reduces the costs of primary and secondary damage.
- Operating interruptions are anticipated in time to plan and schedule equipment shutdowns with minimum impact on operations.
- It provides the information and time necessary to manage spare parts, labour and tools logistics more effectively.
- It provides information to analyse system reliability.

2.3 Maintenance

2.3.3.4 Proactive Maintenance

Proactive maintenance is non-repetitive actions to prevent problems and gain reliability. Examples of proactive maintenance are precision shaft alignment and balancing or the opportunistic upgrade of equipment during regular maintenance, amongst others. Proactive maintenance aims to cost effectively improve safety, predictability, asset lifetime and performance.

Proactive maintenance begins with understanding the root causes of problems. Improvements to process, programmes and technology are then defined and implemented to eliminate defects and extend equipment life. The implementation of improvements are prioritised by cost and return.

A facility that implemented a pump improvement programme is an example of proactive maintenance. A set of proactive improvements were defined for the pumps. Then, whenever a pump went for repairs, the improvements were made. Within a few years, all the pumps were upgraded to the latest specifications with increased reliability.

Benefits of proactive maintenance include lower maintenance costs, lower risk of production interruptions, greater operating efficiency, greater labour productivity and lower demand for spare parts.

2.3.4 The Maintenance Process

Leading corporations regard the maintenance process as an essential part of production. They know that maintenance generates value by ensuring timely, efficient and effective delivery of their products. The basic maintenance process is shown in Figure 2.48.

2.3.5 Reliability Centred Maintenance

Reliability Centred Maintenance (RCM) is a systematic process that ensures safety and mission accomplishment by assuring continuity of system functions. The RCM process defines system boundaries and identifies system functions, functional failures and likely failure modes for the equipment within the specific operating context. It also identifies the causes and effects of functional failures. The RCM process produces a maintenance strategy that specifies applicable and cost effective maintenance tactics to reduce the

2.3 Maintenance

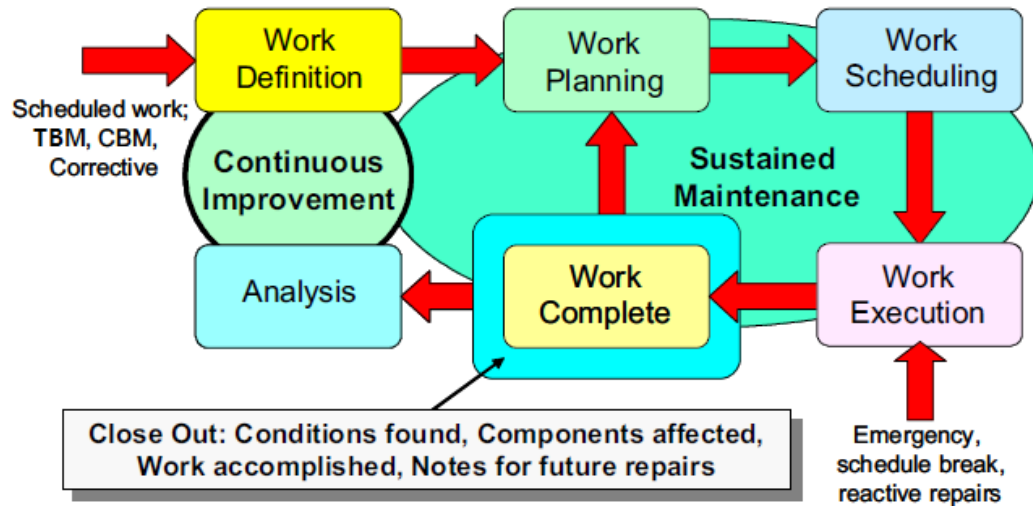


Figure 2.48: The basic maintenance process (Mitchell, 2014)

probability of functional failure. The RCM analysis extracts the experience and knowledge of operating and maintenance experts who understand how the equipment works, as well as its operating and maintenance deficiencies.

The RCM process has seven steps, as shown in Figure 2.49. In the first step, the boundaries and performance standards for all systems and assets are defined. The inputs and outputs for every system and asset are identified in block diagram format. In the second step, the functional failures are defined and the failure modes of every component that can lead to a functional failure are identified. The third, fourth and fifth steps consist of establishing what causes each functional failure, stating what happens when each failure occurs and establishing the consequences of each failure, respectively. In the sixth step, the most applicable and cost effective countermeasure is selected for each failure mode. Countermeasures include tasks that prevent, mitigate, detect the onset of or discover equipment failure modes. The final step determines what should be done if a suitable avoidance action is not practical. For example, a temporary measure may be used while a permanent solution is found and implemented.

RCM applies a task selection logic to determine the optimum maintenance plan for an asset. The logic is shown in Figure 2.50. The logic considers condition and time based tasks. Tasks are evaluated based on cost effectiveness. The most cost effective tasks are preferred. Most of the time, condition based tasks are more effective since

2.3 Maintenance

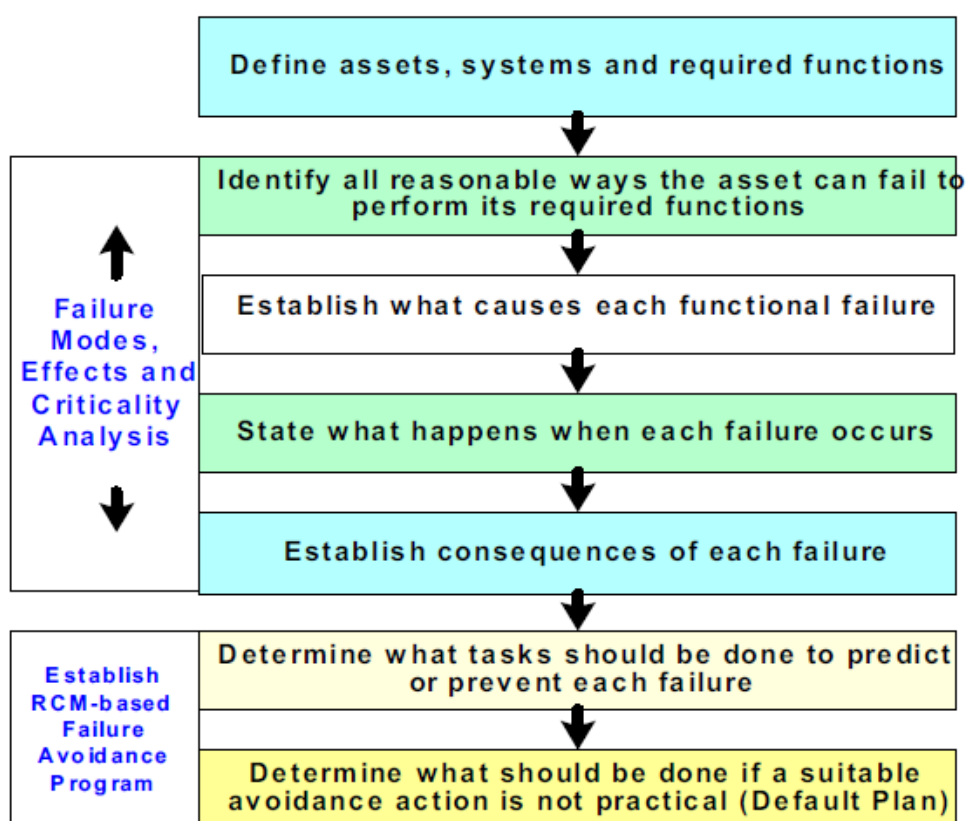


Figure 2.49: The RCM process (Mitchell, 2014)

2.3 Maintenance

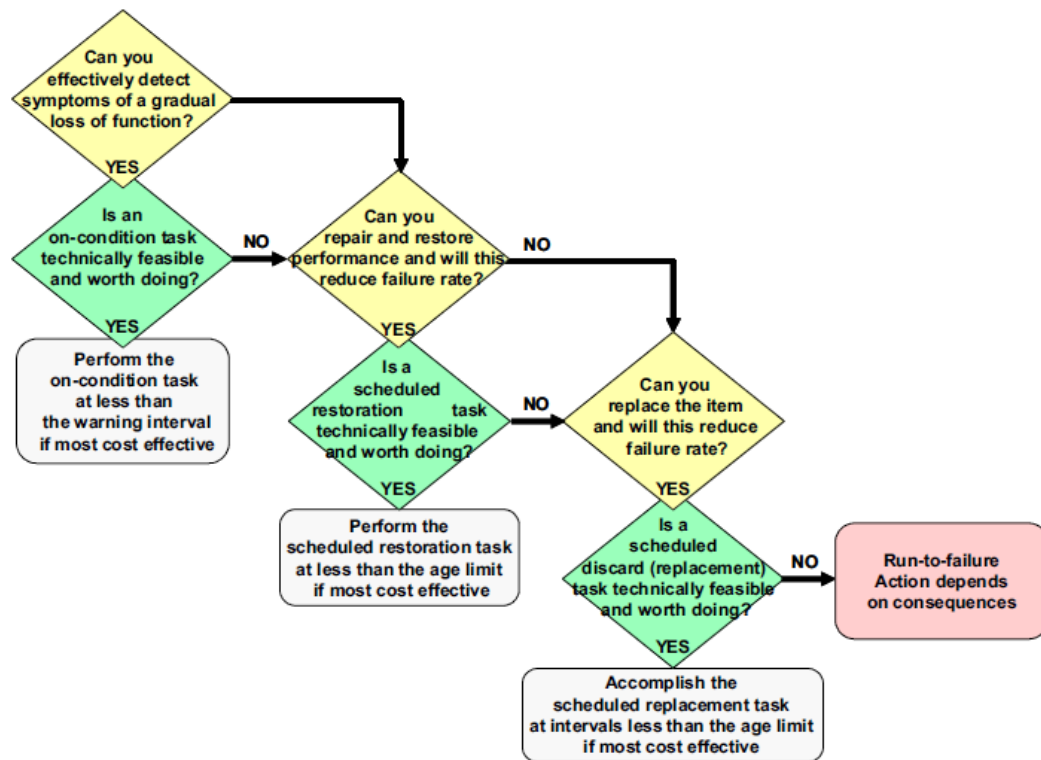


Figure 2.50: RCM task selection logic (Mitchell, 2014)

only monitoring is required until the asset condition needs correction.

If no applicable and cost effective task is identified, the only immediate option may be run-to-failure. If the failure mode has severe consequences the asset may have to be re-designed to eliminate or mitigate the consequences of a failure. Frequently, the failure mode can be managed with a low cost modification. For example, installing monitoring equipment that provide a warning at the onset of degradation before a functional failure occurs.

RCM recognises that equipment fail according to different failure modes. Figure 2.51 shows six common failure modes. Maintenance intervention may only improve reliability for those profiles that display a rising probability of failure toward the end of asset life, *i.e.* wears out. In Figure 2.51 only 11 percent of failures display this wear-out characteristic. For the majority of cases that do not display wear-out failure patterns, condition based maintenance is more applicable and cost effective than time based maintenance.

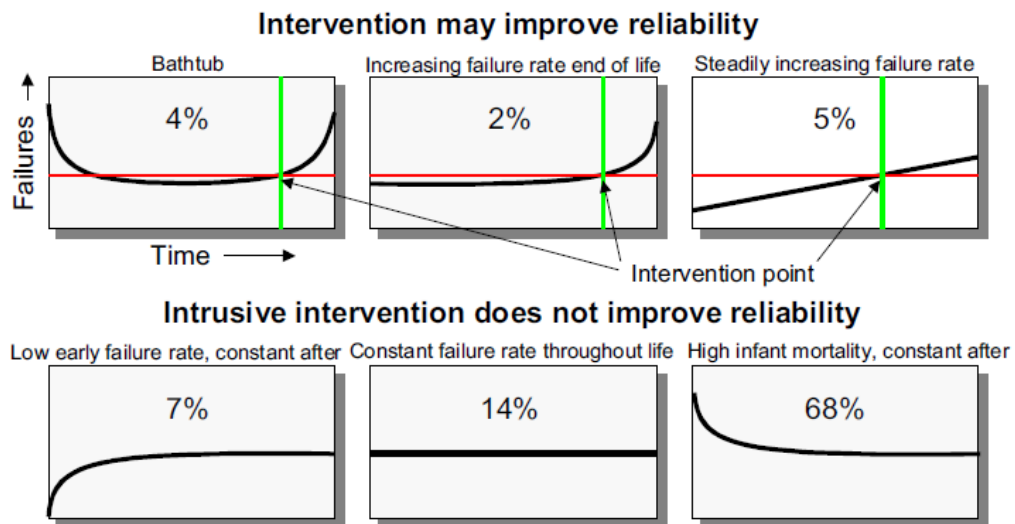


Figure 2.51: Six common failure profiles and their occurrence percentage in commercial aircraft (Mitchell, 2014)

Based on knowledge of the failure process RCM applies decision criteria to evaluate maintenance tasks. These decision criteria are the presence of a dominant failure mode, task applicability, and task effectiveness.

Regarding the dominant failure mode, one can say that:

- Maintenance interventions focus on the dominant failure modes. The failure modes must be specified and reasonably likely to occur.
- Operating and maintenance personnel are best suited to determine whether, or not, a failure mode is reasonably likely to occur.
- The failure history does not expose all the potential failures modes that may occur. All potential failure modes must be identified and the risk of each must be assessed.
- CBM or monitoring equipment must be able to provide positive answers to the following questions:
 - Does the technology monitor a specific failure mode?
 - If so, what is the failure mode?

2.3 Maintenance

- Is it reasonable to expect that this failure mode will occur during the lifetime of the equipment?

RCM applies rules to determine the applicability of maintenance tasks. These are:

- Every task must be technically feasible. That is, the task must enable a person to find, mitigate or prevent a failure or degrading condition that leads to a failure.
- Time based maintenance is applicable if the interval between failures is known with reasonable accuracy.
- Condition based maintenance is applicable if:
 - The measured parameters correlate to deterioration of the asset condition and the related failure modes.
 - The parameters are measurable, repeatable and stable enough over time to reliably indicate when corrective action is needed.
 - There is enough time between identifying a potential failure and an actual failure to take the appropriate corrective action.

The effectiveness of a maintenance task is determined by its ability to address the consequences of a failure. A task effectively addresses critical failures if it reduces the risk of failure to a tolerable level. Whereas, a task effectively addresses non-critical failures if it is cost effective. That is, the cost of the resources needed to prevent the failure is less than cost of the resources needed to repair the failure.

The RCM process provides a logical approach to determine what maintenance tactics will improve reliability. It identifies all the potential functional failures, their causes and failure modes and then specifies maintenance interventions based on applicability and effectiveness.

2.3.6 Total Productive Maintenance

Total Productive Maintenance (TPM) optimises production processes and results by establishing a partnership between Operations and Maintenance. Small teams develop and carry out a knowledge-based strategy. It is a multidisciplinary, team based plant improvement methodology.

TPM emphasises:

2.3 Maintenance

- Partnership and cooperation between Operations and Maintenance departments.
- Autonomous maintenance conducted by small multi-skilled teams consisting of operations and maintenance personnel. The teams accept and share responsibility for the cleanliness, performance and maintenance of their equipment.
- The aim of zero defects and loss operations.
- The aim to uphold cleanliness.

TPM results in rapid and continuous improvement of the production process through employee involvement, employee empowerment and closed-loop measurement of results. TPM is a long-term approach to production optimisation more than it is a short-term intervention.

TBM is based on the following five principles:

1. Improving equipment effectiveness
2. Autonomous maintenance performed by operators
3. Preventive maintenance done by the maintenance department
4. Training to improve operation and maintenance
5. Introducing reliability and maintainability into the design process to prevent problems during the start-up of new equipment

TPM implementations in Japan have emphasised:

- A culture of pride
- In-depth study of the TPM principles before implementation
- TPM as the primary maintenance focus with no other maintenance improvement programmes
- Starting TPM implementation by restoring the plant to an “as new” condition
- The use of TPM principles to allocate tasks. Typically, operators are responsible for operation and daily maintenance. While, the maintenance department is responsible for time based maintenance, repairing failures and improving maintainability.

2.3 Maintenance

Typical American implementations of TPM follow the process shown in Figure 2.52 and emphasise:

- Cross-functional multi-level steering teams that are composed of managers, supervisors, support staff and workers from departments such as Maintenance, Production, Engineering and Stores. Other departments such as Information Technology, Finance, Procurement and Admin are also involved in the TPM process.
- Multidisciplinary teams that consist of working level maintenance and production staff dedicated to operation, identifying and solving problems and the goals of zero accidents, zero defects and zero failures.
- A culture of initiative and ownership
- A published master plan that includes objectives, list of all activities, schedule, resource requirements, roles and responsibilities and criteria for measuring progress and success.
- Safe, clean and orderly workspaces
- Proactive problem discovery and prevention
- Operations that accept and share responsibility with maintenance for equipment condition. Operations do autonomous maintenance such as:
 - Development of cleaning and lubrication standards
 - Preventive cleaning
 - General inspections to identify and correct chronic problems
 - Minor maintenance and adjustments
 - Minimising waste
 - Building pride and ownership
 - Upholding process discipline
- Optimised time and condition based maintenance
- Proactive activities to increase equipment reliability and maintainability such as:
 - Work prioritisation

2.3 Maintenance

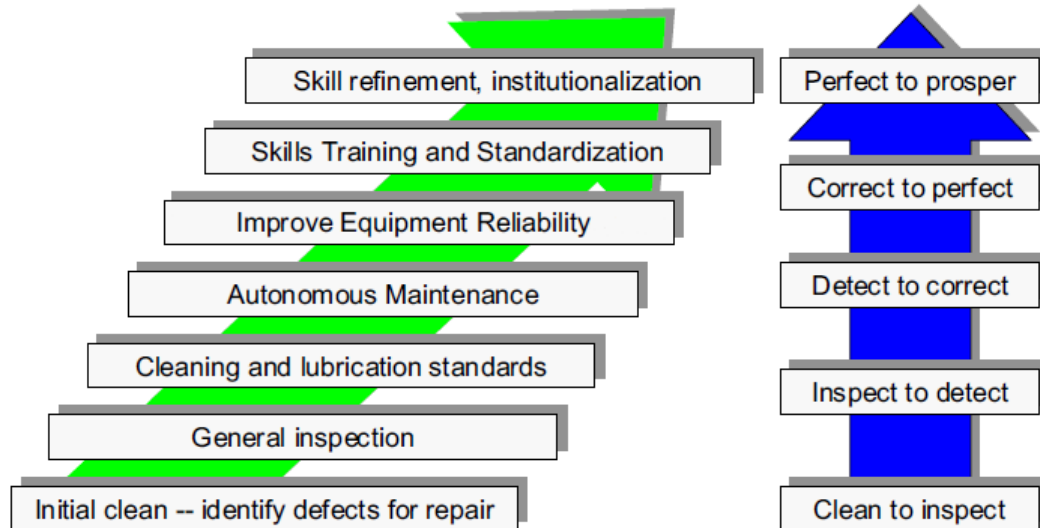


Figure 2.52: The Total Productive Maintenance process (Mitchell, 2014)

- Improved planning and scheduling
- Continuous improvement
- Skills training for operations and maintenance

2.3.7 Failure Analysis

A failure is the inability of a component or system to meet the specified operating characteristics. Failure analysis methods provide a structured approach to find a cause. The best failure analysis methods identify the root of a problem.

Figure 2.53 shows the normal state of an asset alongside the two types of failures that can occur. In the normal state, the capability of a component, asset or system is greater than or equal to the demand placed on it. A failure occurs once the demand exceeds the capability. There are two types of failures. The first type occurs when capability drops below the demand, as shown in the centre of Figure 2.53. If the capability drops but some functionality remains, then a partial failure has occurred. If the asset cannot function at all, a total failure has occurred. The second type of failure occurs if the demand exceeds the capability of an asset. This may result in risky behaviour such as pushing equipment beyond its specified capabilities.

2.3 Maintenance

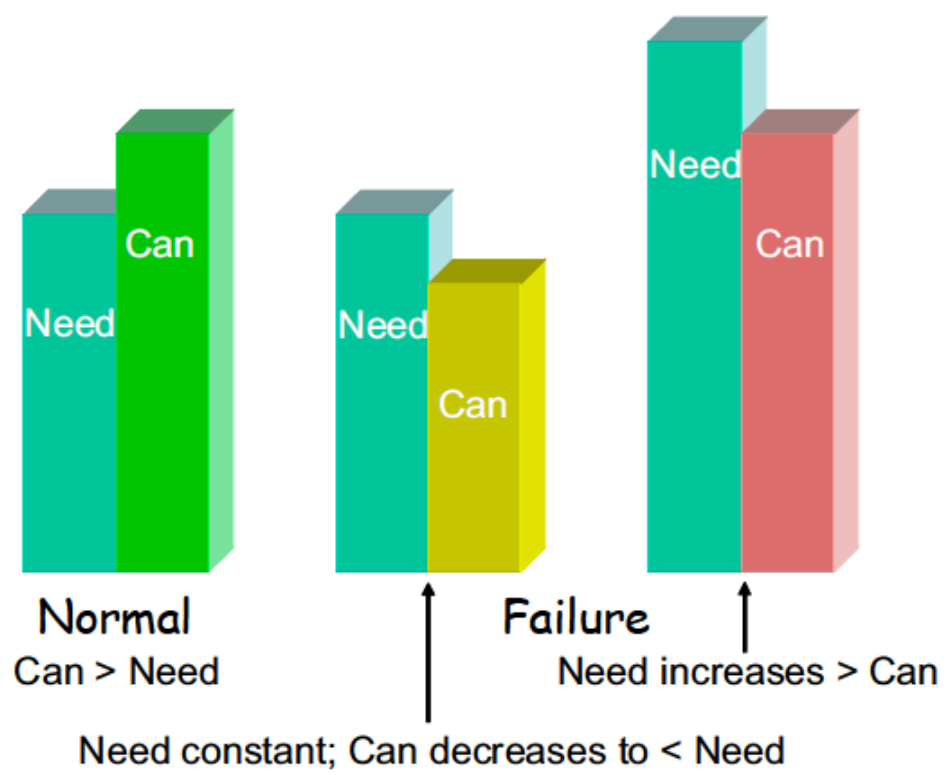


Figure 2.53: Failure modes (Mitchell, 2014)

2.3 Maintenance

The failure analysis process starts by stating the what, where, when and consequences of the problem. Next, the failure modes and causes are identified. The cause and effect path is followed until the root cause is found. Finally, solutions are developed to eliminate the root cause. If the operational or maintenance requirements to eliminate the problem are not cost effective, design changes may be needed.

Components fail as a result of one or more of the following failure agents:

1. Force
2. Reactive environment
3. Time
4. Temperature

Typically, checklists and troubleshooting tables are used to select or determine which failure agents were responsible for a failure.

The cause of machinery failures fall into one or more of the following categories:

1. Faulty design
2. Material defects
3. Fabrication or processing errors
4. Assembly or installation defects
5. Off-design, unintended or especially hostile service conditions
6. Maintenance deficiencies, including neglect
7. Improper operation

2.3.7.1 Failure Modes, Effects and Criticality Analysis

Failure Modes, Effects and Criticality Analysis (FMECA) is a pre-failure systems analysis technique that identifies and prioritises potential failures and formulates avoidance actions. FMECA proactively addresses reliability problems.

The FMECA process determines the following:

- System description

2.3 Maintenance

- System function “what does it do”
- Functional failure “what happens”
- Failure mode “immediate cause”
- Failure cause “root cause”
- Failure effects “consequences, penalty”
- Probability “how likely”
- Criticality “risk”
- Preventive and corrective measures

If the FMECA process identifies a design defect that cannot be corrected by maintenance, it is passed on to Engineering for correction. The basic FMECA process is shown in Figure [2.54](#)

Benefits of the FMECA process include:

- Develops and verifies asset hierarchy
- Focuses resources on highest risk ranked equipment
- Captures maintenance and engineering knowledge
- Identifies and eliminates unnecessary maintenance

2.3.7.2 Root Cause Analysis

Root Cause Analysis (RCA) is a logical post-failure analysis that determines the root cause and effect relationship of a failure. RCA includes actions to minimise or eliminate the reoccurrence of similar failures. RCA addresses the fundamental cause, while recognising that there may be several contributing factors.

RCA's are typically done by multidisciplinary teams composed of Engineering, Operating and Maintenance personnel. Minor failures can be analysed by small teams or individuals.

A RCA identifies:

- What happened

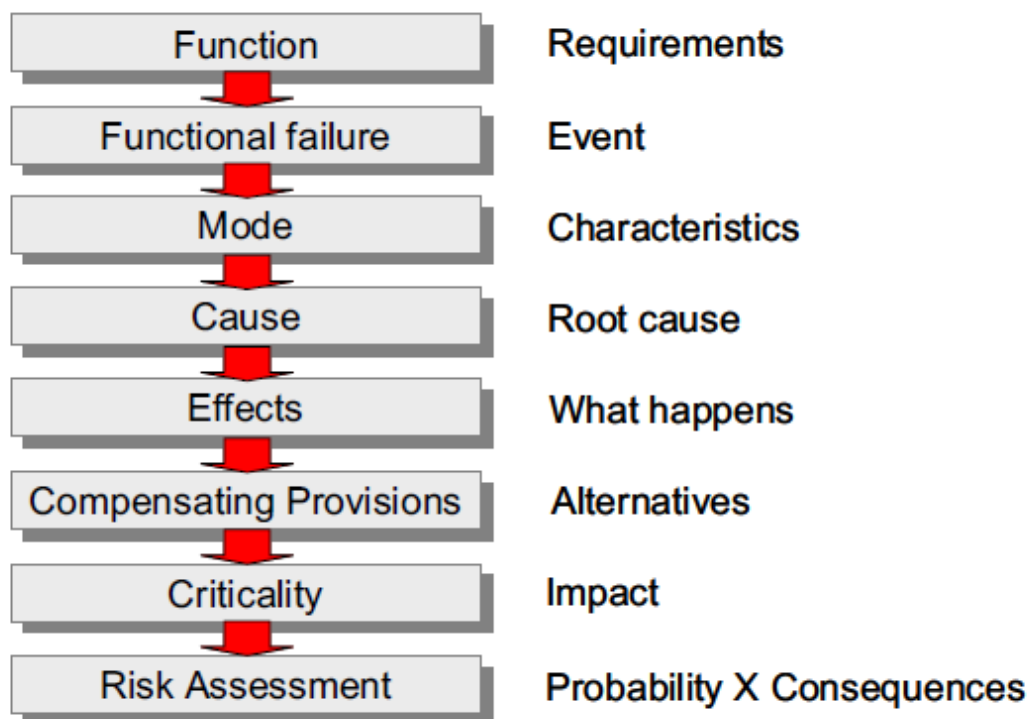


Figure 2.54: The basic FMECA process (Mitchell, 2014)

2.4 Maintenance of Railway Infrastructure

- Why it happened
- Why it was not discovered or prevented
- Corrective action to assure that it does not happen again

Some of the benefits of RCA are:

- Savings from prevented failures
- The developed solutions may address recurring problems that are occurring elsewhere in a plant.

2.4 Maintenance of Railway Infrastructure

This section is based on [Duvel \(2017\)](#). A railway system consists of operations, fixed infrastructure and rolling stock. The three components of a railway system are interdependent, as shown in Figure 2.55. To operate the railway, the infrastructure and rolling stock must be available and functional. To maintain the assets, the infrastructure and rolling stock must be removed from service when required. Operating the assets outside of their specifications increases the deterioration rate of the equipment at the physical interfaces.

Railway asset management is holistic and business centred. It integrates operations and maintenance to increase the overall return on assets, over their total life cycle, and maximise profits. It needs to balance operational performance, financial performance and risk management.

The asset maintainer adds value by ensuring that the assets are reliable, available, affordable and safe. Reliability is achieved by removing infrastructure defects and maintaining the railway to the appropriate standard. Availability is achieved by minimizing maintenance interference with railway operations, reducing the mean time to repair breakdowns and eliminating speed restrictions due to the infrastructure condition. Affordability is achieved by reducing the direct costs of maintenance resources and reducing the indirect costs of infrastructure unavailability. Safety is achieved by complying with maintenance processes, standards, safe working procedures, training and risk management.

2.4 Maintenance of Railway Infrastructure

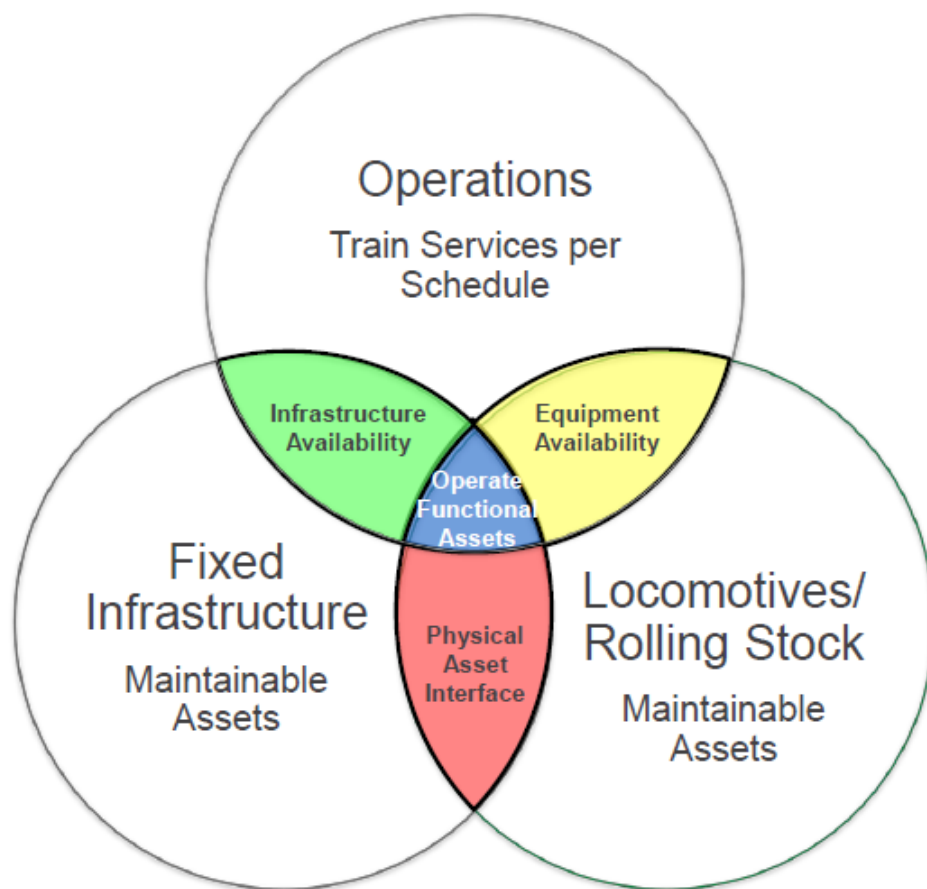


Figure 2.55: Interdependence of railway operations and maintenance (Duvel, 2017)

2.4 Maintenance of Railway Infrastructure

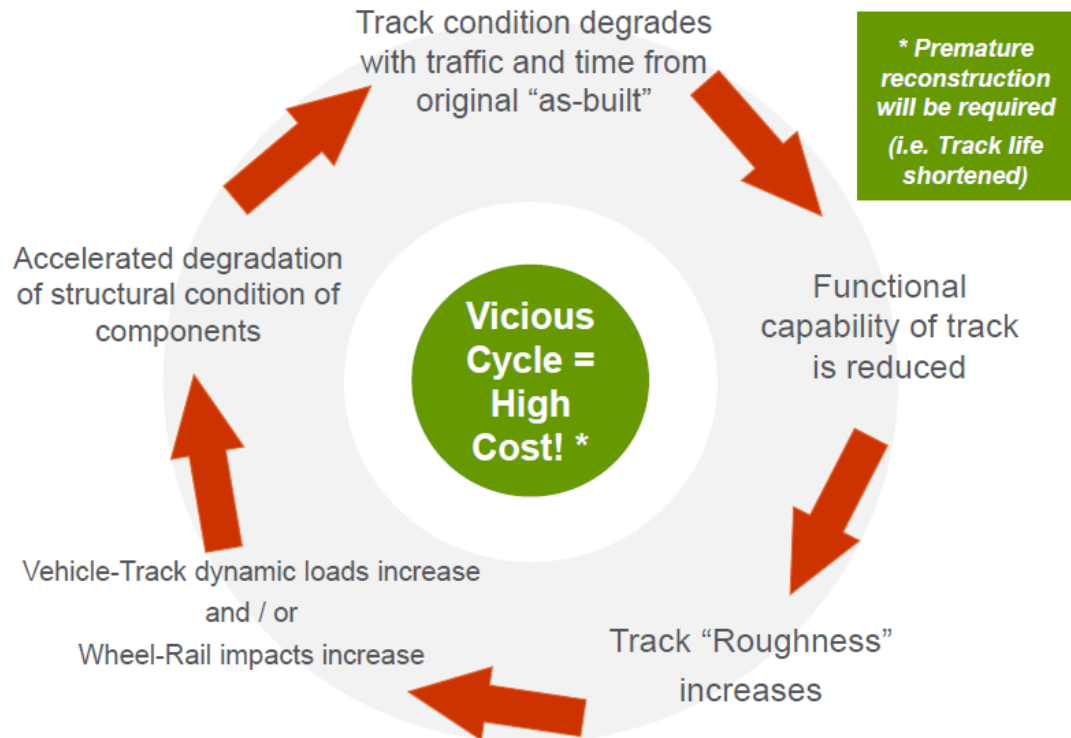


Figure 2.56: The track deterioration cycle (Duvel, 2017)

2.4.1 Track Deterioration

Most of this subsection is based on Duvel (2017). Other references are cited in the text. Railway tracks deteriorate without maintenance. Figure 2.56 shows how the degradation accelerates.

Figure 2.57 shows how the track condition deteriorates as the gross traffic loads, measured in million tonnes, increases. The figure also shows how track rehabilitation and replacement activities improve the track condition. Initially, a new track deteriorates rapidly. After the initial period, the track deteriorates gradually until it reaches the maintenance standard. Once the track condition exceeds the standard, the deterioration accelerates exponentially. Maintenance activities rehabilitate the track before it meets the standard to extend the operational life of the asset. Nevertheless, the overall condition of the track declines after several rehabilitation cycles. This decreases the time between rehabilitation cycles. The increase in maintenance frequency reduces infrastructure availability and interferes with rail operations. Eventually, rehabilitation

2.4 Maintenance of Railway Infrastructure

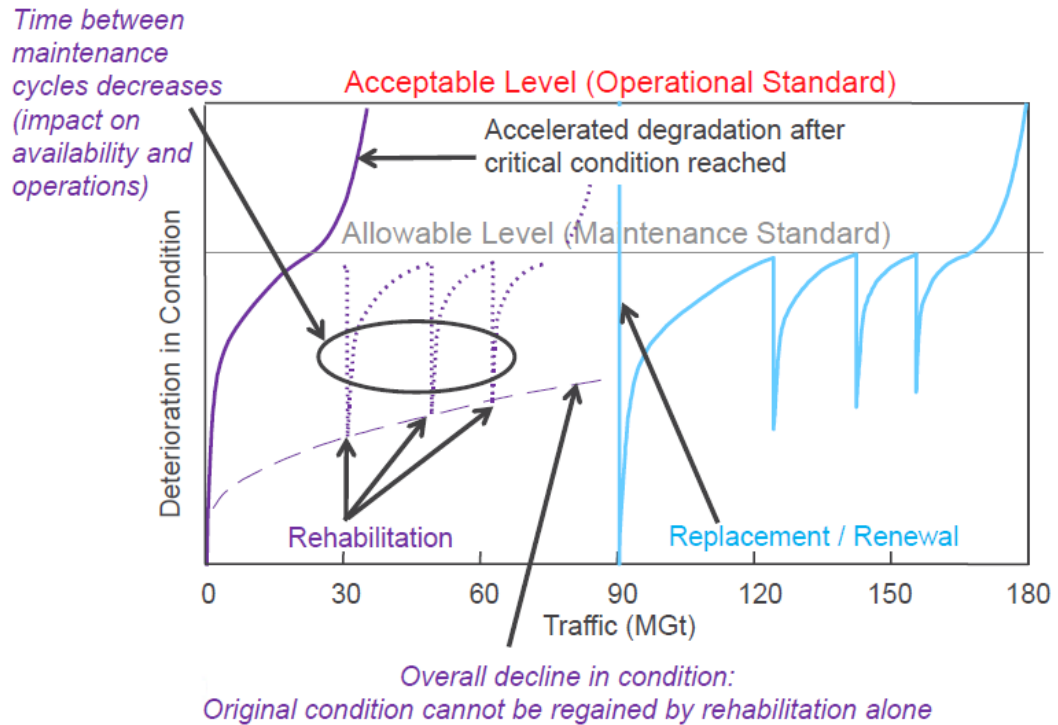


Figure 2.57: Track deterioration, rehabilitation and replacement (Duvel, 2017)

cannot regain a near-to-original condition and the asset must be replaced.

Figure 2.58 shows the deterioration of track condition over time with maintenance interventions. Figure 2.59 shows that constructing the track with higher quality material may extend the asset life. Whereas, Figure 2.60 shows that deferring maintenance may accelerate track deterioration and shorten the asset life (Marutla, 2018).

2.4.2 Asset Lifecycle Process of Railway Infrastructure

Figure 2.61 shows the high-level asset lifecycle process of railway infrastructure at Transnet Freight Rail. Inputs to the asset lifecycle process include the long-term strategic intent, the medium-term strategy and the yearly business plan. The red blocks in the figure show how the asset lifecycle process fits into the four asset lifetime stages introduced in Subsection 2.3.1.1. The maintenance process of railway infrastructure is shown in lifetime stage 3 (van Aardt, 2018).

The lifetime cost of an asset was introduced in Subsection 2.3.1.1. Figure 2.62 shows that maintenance accounts for 70 percent of the lifetime cost of railway infrastructure.

2.4 Maintenance of Railway Infrastructure

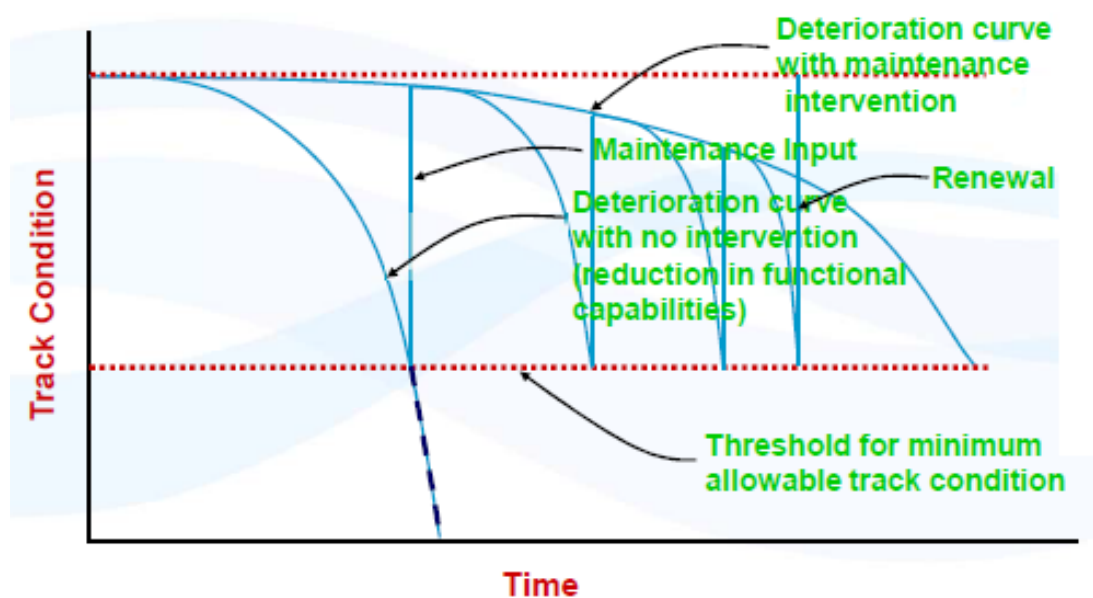


Figure 2.58: Track deterioration with maintenance interventions (Marutla, 2018)

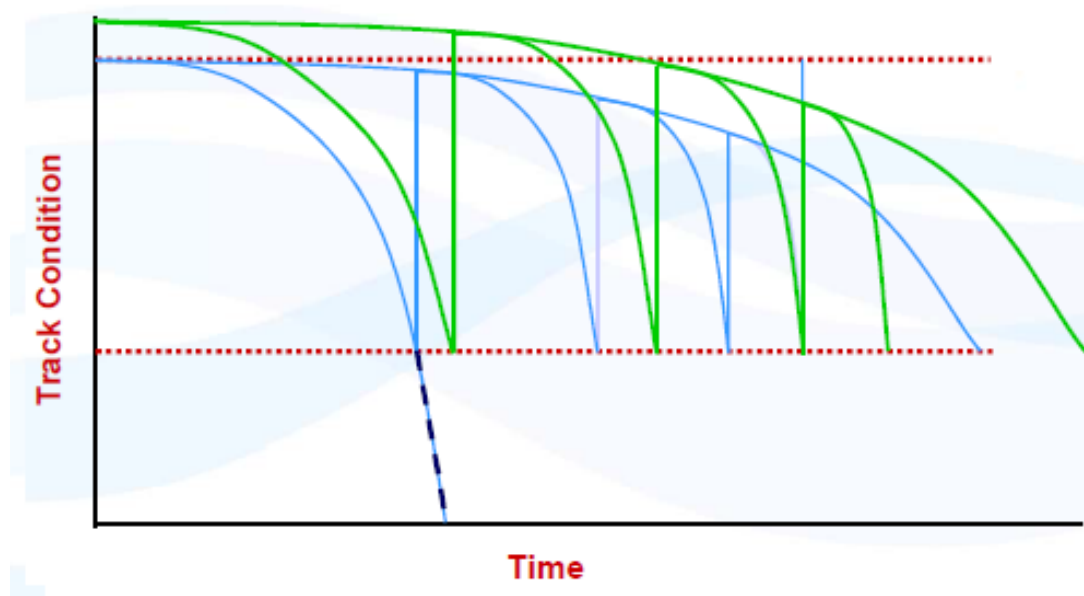


Figure 2.59: Track deterioration with higher initial quality (Marutla, 2018)

2.4 Maintenance of Railway Infrastructure

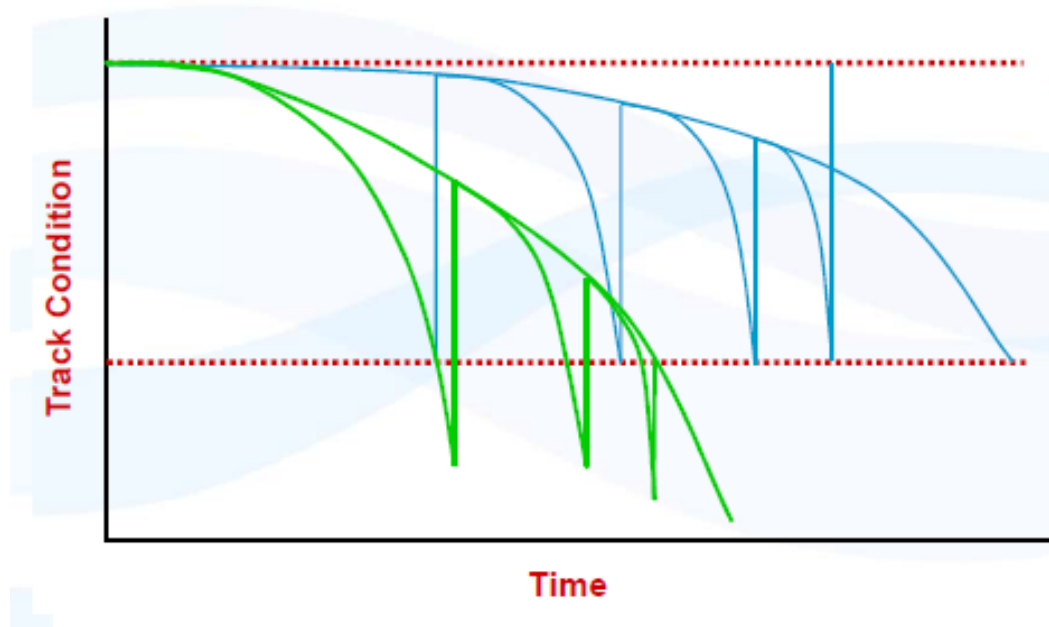


Figure 2.60: Track deterioration when maintenance is deferred (Marutla, 2018)

2.4.3 Maintenance Process of Railway Infrastructure

A general version of the maintenance process was introduced in Subsection 2.3.4. While, the maintenance process of railway infrastructure is shown in Figure 2.61 as part of the operation lifetime stage. An overview of the maintenance process for railway infrastructure is described in this subsection. While, the specific monthly maintenance planning process of TFR is discussed in Chapter 3.

The maintenance process for railway infrastructure consists of the following steps (van Aardt, 2018):

1. Understand the asset base
2. Condition assessment and failure analysis
3. Identify the maintenance need
4. Develop the maintenance plan
5. Execute the maintenance plan
6. Feedback and engineering analysis

2.4 Maintenance of Railway Infrastructure

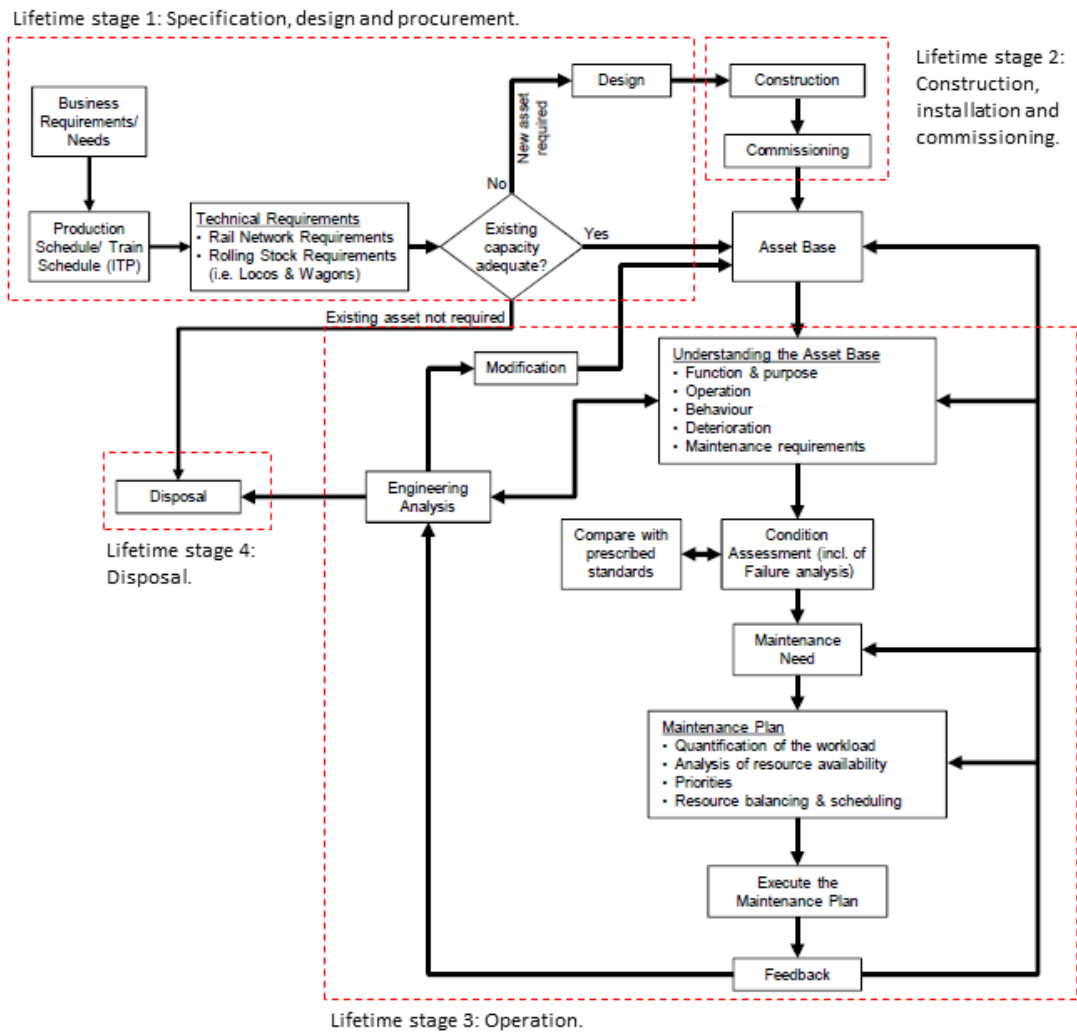


Figure 2.61: High-level asset lifecycle process of railway infrastructure (adapted from van Aardt (2018))

2.4 Maintenance of Railway Infrastructure

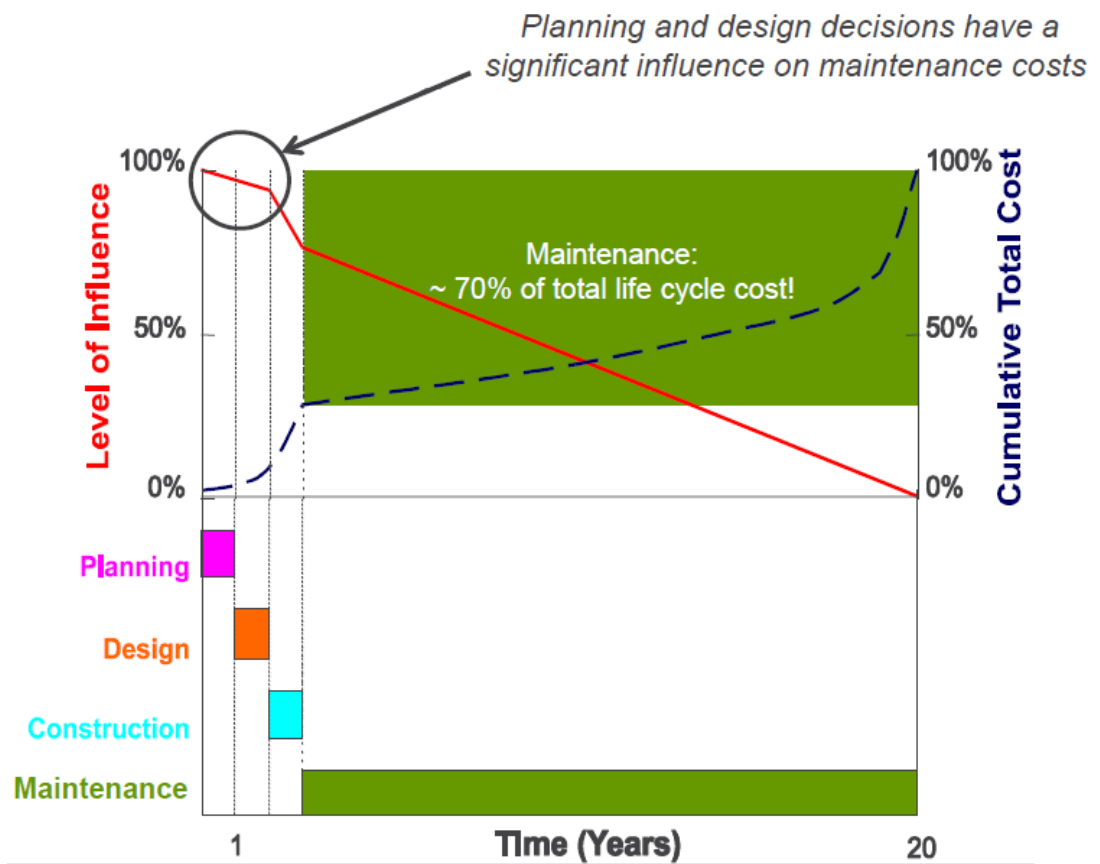


Figure 2.62: The contribution of maintenance to the overall lifecycle cost of a railway (Duvel, 2017)

2.4 Maintenance of Railway Infrastructure

The steps will subsequently be discussed.

Step 1: Understand the Asset Base

Understanding the asset base is the first step of railway infrastructure maintenance. Inputs to this step include the asset inventory list, detailed information about the assets and layouts of the rail network infrastructure. The capacities of the technology subsystems that comprise railway infrastructure are also determined. The subsystems include the railway tracks, bridges, culverts, signalling, electric and telecommunication systems, as discussed in Subsection 2.2.1. Whenever infrastructure projects create additional capacity, new assets are added to the asset base (van Aardt, 2018).

Step 2: Condition Assessment and Failure Analysis

Most of this subsection is based on Duvel (2017). Other references are cited in the text. Condition assessment and failure analysis was introduced in Subsection 2.3.3.3 and Subsection 2.3.7, respectively. Condition assessment is the second step of the maintenance process of railway infrastructure. Condition assessments compare the structural and functional condition of assets to defined standards and limits. Condition assessment also includes failure analysis.

Typically, track components fail progressively. The structural condition of individual components determine the functional condition of the track. The functional parameters of the track can be measured continuously. Standard deviations of these parameters indicate the track roughness. The roughness is an accurate, objective and quantified measure of the overall track condition. Table 2.4 shows examples of functional parameters, while Table 2.5 shows examples of structural parameters. The track components were introduced in Subsection 2.2.1.1 and discussed in Subsection 2.2.1.3, while the basic functional parameters were discussed in Subsection 2.2.1.4. While the examples focus on the railway track, condition assessment is applied to the other subsystems of railway infrastructure technology as well.

Inspections of railway infrastructure can be done manually or with on-track vehicles that are designed for condition assessment. Figure 2.63 and Figure 2.64 show examples of on-track inspection vehicles. The IM2000 measures railway geometry, while the ultrasonic measuring car detects rail defects (van Aardt, 2018).

2.4 Maintenance of Railway Infrastructure

Table 2.4: Functional condition measurement parameters (Duvel, 2017)

Performance Mode	Functional Condition Measurement Parameters
Vertical	Standard deviation of profile
	Standard deviation of twist
	Ride quality
Lateral	Standard deviation of alignment
	Standard deviation of gauge
	Standard deviation of superelevation
Wheel-Rail Contact	Rail transverse profile
	Rail longitudinal profile
Rail Wear	Rail transverse profile
	Crown and side wear
	Remaining rail life

Table 2.5: Structural condition measurement parameters (Duvel, 2017)

Performance Mode	Structural Condition Measurement Parameters
Rail Fatigue	Rail defects
	Rail breaks
Rail Surface	Corrugations
	Rail contact fatigue damage
	Field side flow
Components	Sleeper integrity (cracked, worn, missing)
	Fastener integrity (broken, worn, missing)
Track Support	Ballast rounding or crushing
	Ballast fouling
	Formation condition
	Moisture

2.4 Maintenance of Railway Infrastructure



Figure 2.63: IM2000 railway inspection vehicle (van Aardt, 2018)



Figure 2.64: Ultrasonic measuring car (van Aardt, 2018)

2.4 Maintenance of Railway Infrastructure

The condition assessment data can be used to model the track performance. These models can predict the future performance according to the functional condition measurements or determine the probability of a failure according to the failure history. The condition assessment data can also be used to assess the remaining service life of the track and establish the work requirements to maintain the track within the standards.

Condition assessment includes investigating the causes of track deterioration and failures. Transnet Freight Rail uses the Root Cause Analysis method, presented in Subsection 2.3.7.2, to identify the fundamental cause of failures. While, the Failure Modes, Effects and Criticality Analysis process, presented in Subsection 2.3.7.1, is used to proactively address problem areas.

Cause investigations are prioritised in areas with a poor functional condition. In-depth investigations are done to assess the structural condition, load carrying capability and remaining service life of individual components. Furthermore, cause investigations are done to understand how the deterioration occurred, the failure mode of each component, the deterioration rate and how the deterioration could have been controlled or prevented. An evaluation of the maintenance history and costs may also be done.

Step 3: Identify the Maintenance Need

Condition assessment provides the information necessary to establish the overall maintenance requirements. A needs analysis then prioritises maintenance activities to determine what is financially justifiable given what the business needs. The needs analysis consists of “what if” and “cost-benefit” analyses to determine the optimal mix of maintenance interventions. Refer to Figure 2.43, which shows that the cost of unreliable operations must be balanced with the cost of improving reliability through maintenance interventions. Figure 2.65 shows that the cost of reactive maintenance must be balanced with the cost of proactive maintenance to minimise the total maintenance cost (Duvel, 2017).

The needs analysis includes (Marutla, 2018):

- A list of areas and components that need maintenance.
- The projected condition with and without maintenance interventions.
- A summary of the total costs needed to do the selected maintenance.

2.4 Maintenance of Railway Infrastructure

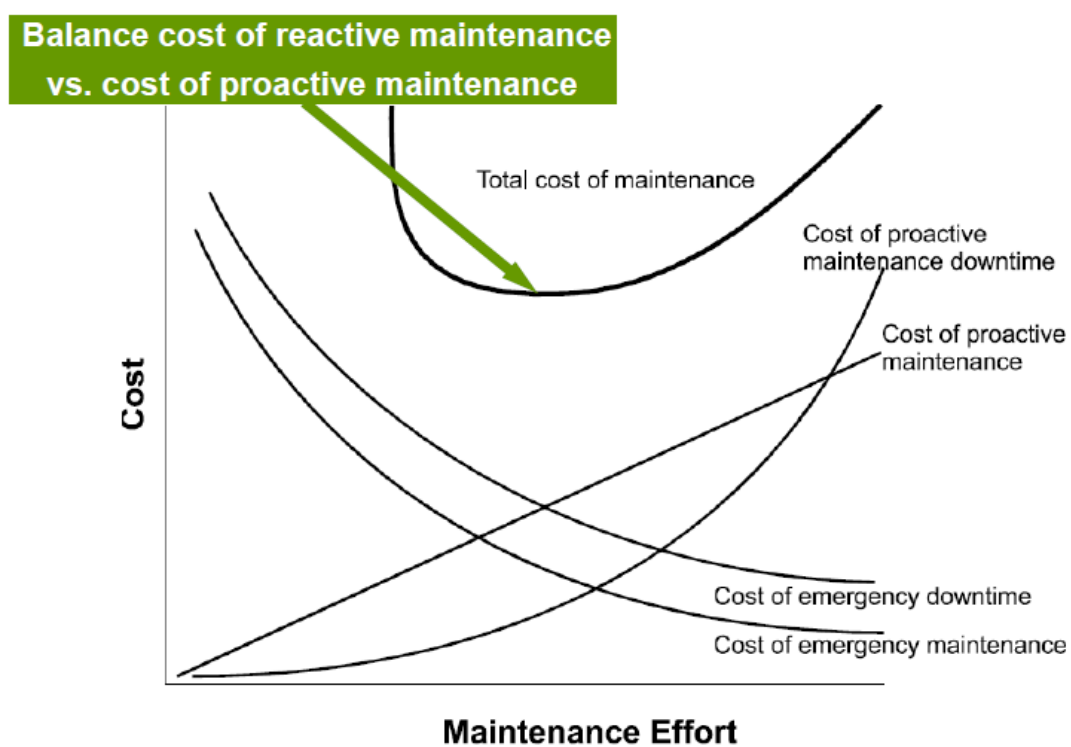


Figure 2.65: The balance of reactive and proactive maintenance costs (Duvel, 2017)

2.4 Maintenance of Railway Infrastructure

- The quantified benefits of the different interventions.
- A prioritised list of maintenance areas and components.

Step 4: Develop the Maintenance Plan

Developing the maintenance plan consists of finalising what maintenance work must be done, quantifying the work requirements, resource planning and scheduling the activities. Maintenance plans are developed for the short-term, as well as for the medium to long-term to manage the asset lifecycles ([van Aardt, 2018](#)). Engineering technicians are responsible for developing the maintenance plan. A maintenance planner requires technical skills and rail infrastructure experience to develop the maintenance plan ([Marutla, 2018](#)).

Resource planning determines what resources are required and then allocates the resources to maintenance jobs. Resources include labour, supervision, contractors, vehicles, materials, components, on-track machines and equipment ([Duvel, 2017](#)).

Maintenance scheduling is based on the availability of resources, the priority of the work and the approval of track possessions ([Marutla, 2018](#)). The maintenance possession scheduling process is presented in Subsection [3.2.1](#). The central Rail Network Planning department schedules most of the on-track maintenance machines. This ensures that the utilisation of the on-track machines are maximised on the national rail network and that the maintenance priorities are addressed. Regional maintenance depots plan and execute most of the other maintenance ([van Aardt, 2018](#)).

Step 5: Execute the Maintenance Plan

Maintenance execution consists of carrying out the maintenance, renewal and upgrade work. It concludes with recording the maintenance activities and condition assessment data, as well as updating the asset inventory ([Duvel, 2017](#)).

The regional maintenance depots are responsible for the planning, execution, monitoring, quality control and sign off of the maintenance work. The central Rail Network Planning department is responsible for electrical control, fault management, possession approvals, material logistics on the rail network, monitoring the Integrated Train Condition Monitoring System and compiling management information ([van Aardt, 2018](#)).

2.4 Maintenance of Railway Infrastructure

Step 6: Feedback and Engineering Analysis

The feedback and engineering analysis step evaluates the quality of maintenance work, adherence to the maintenance standards and quantifies the efficiency of maintenance activities. The performance of the maintenance department is measured by reviewing the maintenance backlog, total breakdowns, breakdowns with an operational impact, train delays due to breakdowns, resource utilisation, productivity, costs and safety. Interventions are developed and implemented to correct problems (Duvel, 2017).

Replacement and renewal cycles are also refined during the engineering analysis. These cycles are based on the railed tonnages, component obsolescence, failure statistics and condition assessments. Non-adherence to the maintenance standards and renewal cycles increases the maintenance backlog and reduce the rail network availability, reliability and safety (van Aardt, 2018).

2.4.4 Maintenance Tactics of Railway Infrastructure

This subsection is based on van Aardt (2018). Subsection 2.3.3 presents a general version of the maintenance tactics. Figure 2.66 shows how the maintenance tactics are organised for railway infrastructure at Transnet Freight Rail. The red text shows how the general maintenance tactics relate to the railway maintenance tactics in black text. The Reliability Centred Maintenance process, described in Subsection 2.3.5, is used to determine which maintenance tactics to apply to railway infrastructure.

Figure 2.67 shows the decision framework for preventive maintenance. As the condition of the rail infrastructure worsens, the likelihood of breakdowns increase. Maintenance tactics are chosen according to the functional condition of the track, as experienced by the user. Maintenance decisions should not be based on the availability of alternative options to sustain train movements, such as alternative routes or authorisation methods .

Figure 2.68 shows the functional condition versus time for the run-to-failure maintenance tactic. Routine preventive maintenance, corrective preventive maintenance and minor breakdown maintenance is not applicable to this tactic. The functional condition of the asset does not deteriorate gradually. Once the functional condition is impaired, a total failure occurs and major breakdown maintenance is needed.

2.4 Maintenance of Railway Infrastructure

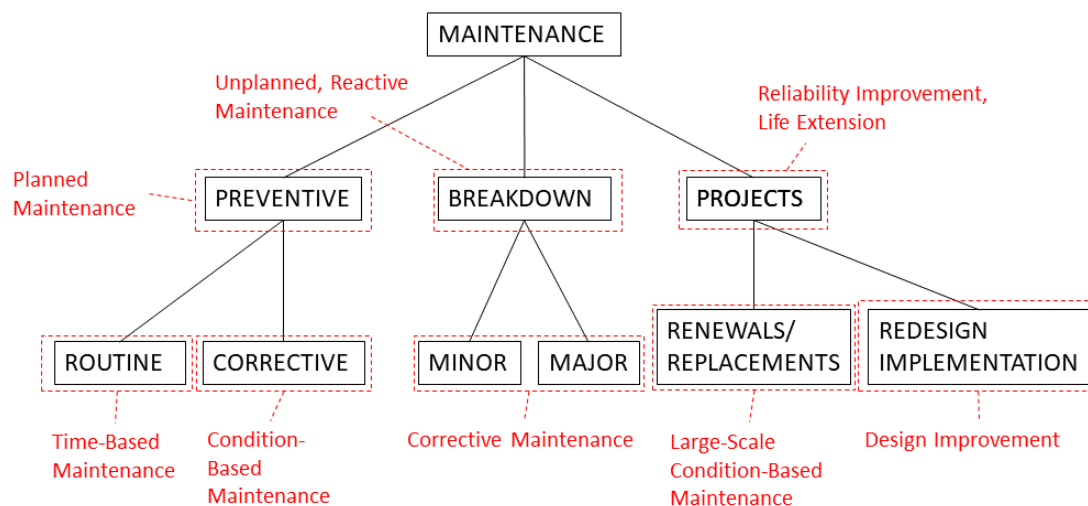


Figure 2.66: Railway maintenance tactics (adapted from van Aardt (2018))

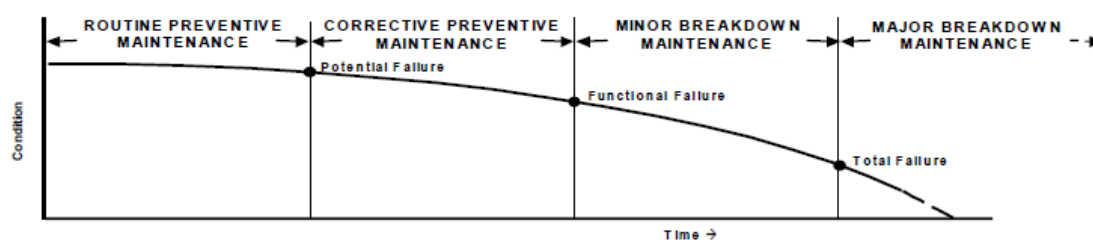


Figure 2.67: Decision framework for preventive maintenance (van Aardt, 2018)

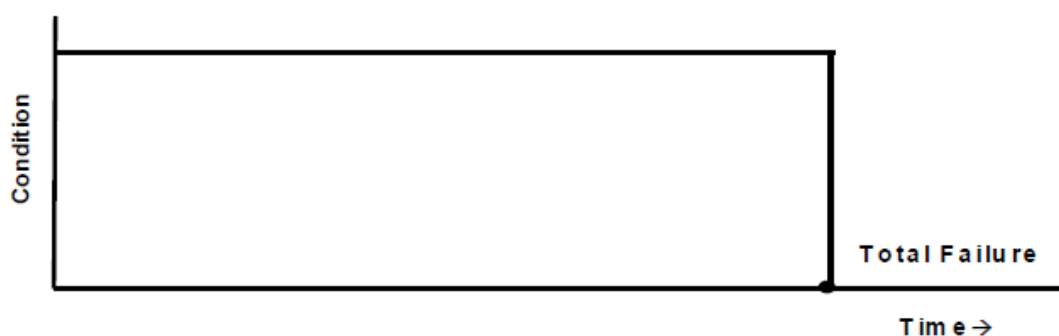


Figure 2.68: The run-to-failure maintenance tactic (van Aardt, 2018)

2.4 Maintenance of Railway Infrastructure

2.4.4.1 Routine Preventive Maintenance

Time-based maintenance, discussed in Subsection 2.3.3.2, is referred to as *routine preventive maintenance* (RPM) in Transnet Freight Rail. RPMs are routine maintenance activities done at regular intervals. General examples include condition monitoring inspections, lubrication, periodic minor overhauls and replacements, cleaning and statutory tasks. Specific examples include lubricating points machines and replacing marshalling yard lights. A points machine is shown in Figure 2.16. Generally, RPMs do not make use of material, only consumables (van Aardt, 2018).

2.4.4.2 Corrective Preventive Maintenance

This subsection is based on van Aardt (2018). Condition-based maintenance, discussed in Subsection 2.3.3.3, is referred to as *corrective preventive maintenance* (CPM) in Transnet Freight Rail. CPM activities correct the asset condition. CPM activities are prioritised and scheduled when a RPM inspection finds that the condition of an asset has deteriorated to the first maintenance standard. Many CPMs need material.

Examples of CPMs include sleeper, rail, fastener and sign replacement, fastening bolts, welding and grinding of rails, repairing fences and cutting trees. Examples of on-track maintenance machines used for CPM work are shown in Figures 2.69–2.73. If the scale of the work exceeds a specified threshold, the CPM is classified as a renewal project. For example, if rails are replaced on more than 1 km of general freight track or more than 240 m on a heavy haul track it is classified as a project. Similar rules apply for all the major rail infrastructure components.

2.4.4.3 Minor Breakdown Maintenance

Corrective maintenance, discussed in Subsection 2.3.3.1, is referred to as *breakdown maintenance* in Transnet Freight Rail. *Minor breakdown maintenance* (MIB) is applicable when the equipment condition has deteriorated to the second maintenance standard. In this case, the equipment is still available, but at a lower performance standard. MIBs may require train speed restrictions or temporary maintenance interventions to reduce the safety risk. Examples of MIBs include repairing faulty points, broken sleepers, broken rails, track slacks, track kick-outs or loose joints (van Aardt, 2018).

2.4 Maintenance of Railway Infrastructure



Figure 2.69: On-track maintenance vehicle for overhead track equipment (van Aardt, 2018)

2.4 Maintenance of Railway Infrastructure



Figure 2.70: Ballast tamper (van Aardt, 2018)



Figure 2.71: Sleeper replacement machine (van Aardt, 2018)

2.4 Maintenance of Railway Infrastructure



Figure 2.72: Rail grinding machine (van Aardt, 2018)



Figure 2.73: Ballast screener (van Aardt, 2018)

2.4 Maintenance of Railway Infrastructure

Table 2.6: Technical challenges of railway infrastructure maintenance (van Aardt, 2018)

Discipline	Symptom	Root Cause
Track	Broken rails	Inadequate rail stress management
	Kick-outs	Thermite weld failures
	Alignment problems	Undetected and unattended rail defects
	Wash-aways	Formation failures
	Missing rail fasteners	Theft of rail fasteners
Electrical	Pantograph and OHTE hook-ups	Sabotage/vandalism/theft
	OHTE equipment failures	Quality assurance after repairs
	Substation equipment failure	Inadequate preventive maintenance
	Distribution power supply failure	Ageing infrastructure
Train Authorisation	Track circuit failures	Theft & vandalism
	Points failures	Obsolete equipment
	CTC equipment failures	Complexity of equipment
	Signal failures	Insufficient backup power supplies
Telecommunication	Transmission failures	Sabotage/vandalism/theft
	Interference	Cable failures
	Train cab system failures	Insufficient backup power supplies
	CTC failures	Software related problems

2.4.4.4 Major Breakdown Maintenance

Corrective maintenance, discussed in Subsection 2.3.3.1, is referred to as breakdown maintenance in Transnet Freight Rail. *Major breakdown maintenance* (MAB) is applicable when the equipment has suffered a total loss of function and is not available for use. The cause of a major breakdown can be internal or external. If the failure results in an operational incident, the disaster recovery process is initiated. If the failure does not result in an operational incident, the maintenance department responds quickly with emergency work. Examples of failures that result in MABs include train derailments, track flooding or wash-aways, veld fires and pantograph hook-ups with the overhead track equipment (van Aardt, 2018).

2.4.5 Technical Challenges of Railway Infrastructure Maintenance

Some of the technical challenges of maintaining rail infrastructure are listed in Table 2.6 (van Aardt, 2018).

2.5 Scheduling Techniques

This section presents scheduling techniques and is based on Talbi (2009). Many scheduling problems are modelled mathematically. Examples of scheduling models include linear programming (LP), integer programming (IP), mixed integer programming (MIP) and constraint programming (CP) models. Solutions for these scheduling models can be found by either *exact* or *approximate* optimisation techniques.

Exact optimisation techniques find optimal solutions for scheduling models. However, it may take a prohibitively long time to calculate optimal solutions for models with a large number of variables or complicated structures. Whereas, approximate techniques find “good” solutions in an acceptable time but do not find optimal solutions. Another option to reduce the solving time is to simplify the model. However, in that case, the model may no longer be an adequate representation of the real-world problem.

Exact optimisation techniques are preferred in instances where it can calculate optimal solutions in an acceptable time. Examples of exact approaches include the simplex algorithm for LPs, branch-and-bound algorithms for IPs and MIPs, and constraint programming for constraint satisfaction models.

Generally, if a scheduling problem is modelled as a LP, the simplex algorithm can find an optimal solution. Although, in instances where the LP model has a large number of variables or complicated structure, an approximate method may be needed to find “good” solutions within an acceptable time.

IP scheduling models make use of only discrete decision variables. Whereas, MIP scheduling models make use of both discrete and continuous decision variables. Small instances of these models may be solved exactly by the branch-and-bound method. Again, IP or MIP models with a large number of variables or complicated structures may be too time consuming to solve exactly for practical use in which case, approximate methods may be needed.

CP is another exact approach used to model scheduling problems. A model consists of a set of variables that are linked by a set of constraints. Every variable takes its value from a finite domain of values. Constraint propagation techniques are used to reduce the variable domains until an optimal solution is found. Here also, the efficiency of the technique depends on the structure of the problem. In general, CP is less suited

2.6 Possession Scheduling

to problems with many feasible solutions and more suited to constrained problems such as scheduling. Talbi (2009) describes constraint programming in more detail.

As mentioned, approximate techniques can be used to find “good” solutions within acceptable times for scheduling models with a large number of variables or complicated structures. Approximate techniques can be divided into two classes: specific heuristics and metaheuristics. Specific heuristics are designed to find solutions to a specific problem. Whereas, metaheuristics are general algorithms that can be used for most problems.

2.6 Possession Scheduling

This section presents an analysis and synthesis of recent work done in the field of possession scheduling. The aim is to identify trends and find commonalities with the work done in this research.

2.6.1 Analysis

This subsection is based on the references in Table 2.7. First, publications were found by using the search phrase “rail infrastructure maintenance possession scheduling”. Then, further publications were found by looking up references cited by the publications within the search results. 57 articles were found and analysed to identify possession scheduling trends based on the following characteristics:

- Tactical or operational applications
- Sequential or integrated planning approaches with regards to possession and train scheduling
- Classification into related work categories:
 1. Maintenance scheduling – traffic impact considered
 2. Train scheduling – with fixed maintenance closures
 3. Combined scheduling of maintenance and trains
- Cyclic, also referred to as periodic, possession and train scheduling
- Infrastructure representation: Macroscopic, mesoscopic or microscopic

2.6 Possession Scheduling

- Allocation of maintenance resources
- Objective functions
- Model types
- Model formulation approaches
- Optimisation techniques
- Application information such as:
 - Rail network descriptions
 - Freight, passenger or mixed railway services
 - Planning horizon
 - Solution precision, *i.e.* the size of the smallest time intervals
 - Number of trains, stations and block sections

These characteristics and their trends will be discussed next.

Railway problems can be classified into strategic, tactical and operational categories. Strategic problems have a long term planning horizon and usually include resource acquisition, construction or modification of existing infrastructure. Examples of strategic problems include the Network Planning Problem and the Line Planning Problem. Tactical problems are concerned with allocating resources to the existing infrastructure. Examples of tactical problems include the Train Timetabling Problem, Train Platforming, Routing or Track Allocation Problems, Rolling Stock Planning Problem and the Crew Planning Problem. Operational problems occur during execution when plans developed at the tactical level need to be adjusted because of disruptions such as late train arrivals, track maintenance, adverse weather conditions or accidents. Operational problems include the rescheduling of trains, crews and rolling stock ([Lusby *et al.*, 2011](#)). A classification of railway problems is shown in Figure 2.74.

When considering maintenance specifically, [Lidén \(2015\)](#) classified railway infrastructure maintenance planning problems into strategic, tactical and operational categories. The strategic problems are:

- Service life and maintenance frequency determination

2.6 Possession Scheduling

Table 2.7: List of references used for the possession scheduling literature analysis

Year	References
2020	Kalinowski <i>et al.</i> (2020), Zhang <i>et al.</i> (2020b), Zhang <i>et al.</i> (2020a), Bešinović <i>et al.</i> (2020), Lidén (2020)
2019	D'Ariano <i>et al.</i> (2019), Bueno <i>et al.</i> (2019), Zhang <i>et al.</i> (2019b), Bababeik <i>et al.</i> (2019), Dao <i>et al.</i> (2019), Zhang <i>et al.</i> (2019a)
2018	Zhang <i>et al.</i> (2018), Arenas <i>et al.</i> (2018), Zhu <i>et al.</i> (2018), Lidén <i>et al.</i> (2018), Lidén (2018b), Lidén (2018a), Kalinowski <i>et al.</i> (2018), Dao <i>et al.</i> (2018), Khalouli <i>et al.</i> (2018)
2017	Van Aken <i>et al.</i> (2017a), Van Aken <i>et al.</i> (2017b), Lidén & Joborn (2017), Luan <i>et al.</i> (2017), Macedo <i>et al.</i> (2017), Looij (2017), Li & Roberti (2017), Dao <i>et al.</i> (2017), Su <i>et al.</i> (2017)
2016	Veelenturf <i>et al.</i> (2016), Lidén & Joborn (2016), Lidén (2016), Famurewa <i>et al.</i> (2016), Vansteenwegen <i>et al.</i> (2016), Pearce & Forbes (2019)
2015	Savelsbergh <i>et al.</i> (2015), Bahramian & Bagheri (2015), Lidén (2015)
2014	Louwerse & Huisman (2014), Boland <i>et al.</i> (2014)
2013	Forsgren <i>et al.</i> (2013), Albrecht <i>et al.</i> (2013), Boland <i>et al.</i> (2013)
2011	Peng <i>et al.</i> (2011)
2010	Pouryousef <i>et al.</i> (2010)
2009	Burdett & Kozan (2009), Budai-Balke (2009)
2007	van Zante-de Fokkert <i>et al.</i> (2007), Caprara <i>et al.</i> (2007)
2006	Budai <i>et al.</i> (2006), Caprara <i>et al.</i> (2006)
2005	den Hertog <i>et al.</i> (2005)
2002	Brucker <i>et al.</i> (2002)
2000	Lake <i>et al.</i> (2000)
1999	Higgins <i>et al.</i> (1999), Cheung <i>et al.</i> (1999)
1998	Higgins (1998)

2.6 Possession Scheduling

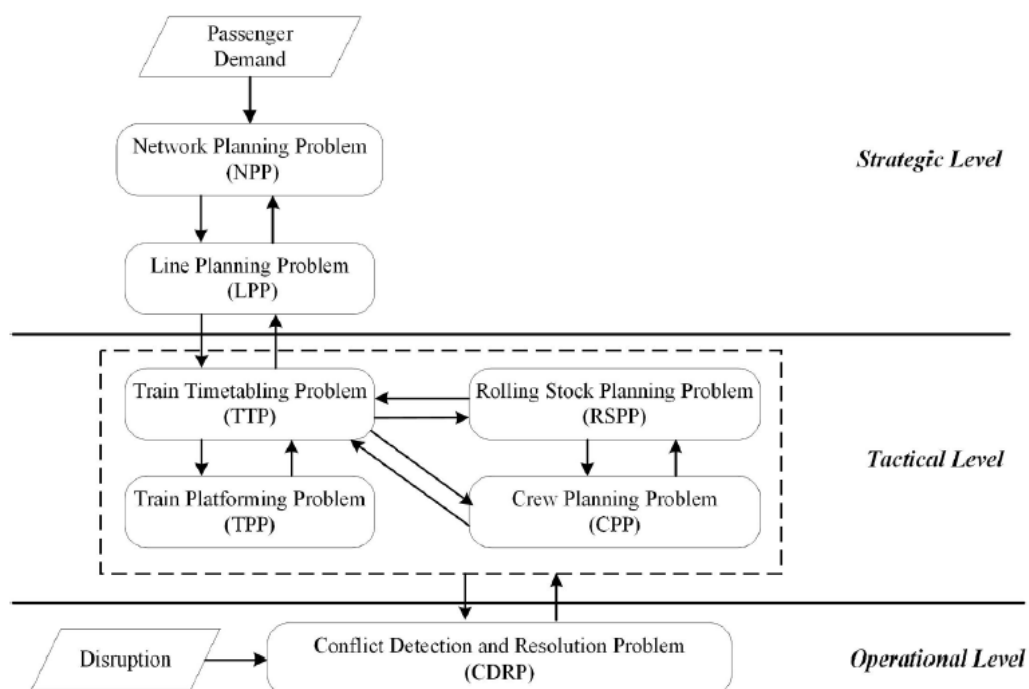


Figure 2.74: Classification of railway problems ([Zhang et al., 2020b](#))

2.6 Possession Scheduling

- Network design considering maintenance
- Renewal scheduling and project planning

The tactical problems in the maintenance domain are:

- Possession scheduling
- Deterioration-based maintenance scheduling
- Maintenance vehicle routing and team scheduling

The operational problem in the maintenance domain is: work timing and resource scheduling.

Possession scheduling plans the allocation of tracks to maintenance activities before the execution of train services. Therefore, most of the publications treat possession scheduling as a tactical planning problem. Some publications consider the operational problem of rescheduling trains when a disruption occurs. With this approach, possessions may be handled as a disruption that blocks trains, *e.g.* [Zhu *et al.* \(2018\)](#).

Two types of possession planning approaches were identified: sequential and integrated. Sequential approaches either:

- Plan the train timetable first, and then schedule the maintenance possessions while reducing the impact on the train service, or
- Plan the possessions first, and then schedule the trains around the possessions.

Whereas, integrated approaches schedule trains and possessions simultaneously. Before 2017, the majority of publications made use of the sequential approach. Recent examples that make use of the sequential approach are [Kalinowski *et al.* \(2020\)](#), [Zhang *et al.* \(2020b\)](#), [Bešinović *et al.* \(2020\)](#) and [Arenas *et al.* \(2018\)](#). From 2017 onwards, the number of integrated approaches has exceeded the number of sequential approaches. Therefore, among the recent literature, there is a clear rise in the use of integrated approaches. The earliest integrated approaches found were [Albrecht *et al.* \(2013\)](#) and [Forsgren *et al.* \(2013\)](#). Recent examples of integrated approaches are [Zhang *et al.* \(2020a\)](#), [Lidén \(2020\)](#), [D'Ariano *et al.* \(2019\)](#), [Bueno *et al.* \(2019\)](#) and [Bababeik *et al.* \(2019\)](#).

[Lidén \(2020\)](#) sorted some of the possession scheduling publications into three related work categories, namely:

2.6 Possession Scheduling

1. Maintenance scheduling with traffic impact considered – *E.g.* Higgins (1998), Budai *et al.* (2006), Boland *et al.* (2013), Savelsbergh *et al.* (2015) and van Zante-de Fokkert *et al.* (2007).
2. Train scheduling with fixed maintenance closures – *E.g.* Caprara *et al.* (2006), Brucker *et al.* (2002), Vansteenwegen *et al.* (2016), Veelenturf *et al.* (2016), Louwerse & Huisman (2014), Van Aken *et al.* (2017a,b), Arenas *et al.* (2018) and Zhu *et al.* (2018).
3. Combined scheduling of maintenance and trains – *E.g.* Albrecht *et al.* (2013), Forsgren *et al.* (2013), Luan *et al.* (2017), D’Ariano *et al.* (2019), Lidén & Joborn (2017) and Lidén *et al.* (2018).

Among the publications before 2016, work related to the first category, “maintenance scheduling with traffic impact considered” appeared most often. However, among the more recent publications, the other two categories dominate. Few articles related to the first category were found from 2016 onwards.

Train timetables can be cyclic or acyclic. A cyclic schedule repeats after a certain period, *e.g.* one hour, one day or one week. Only a few possession scheduling publications consider periodic problems. Examples of these are:

1. Cyclic scheduling of trains and maintenance windows, *e.g.* Lidén (2020).
2. Cyclic train scheduling with fixed maintenance closures, *e.g.* Van Aken *et al.* (2017a,b).
3. Cyclic possession scheduling with traffic impact considered, *e.g.* van Zante-de Fokkert *et al.* (2007).

The rail network can be modelled at different levels of detail, namely: macroscopic, mesoscopic and microscopic. The scope of the timetable determines the detail of the infrastructure representation (Zhang *et al.*, 2020b). At the macroscopic level, stations are modelled as nodes and the track connections between them as arcs. The stations and tracks are given capacity limits (Van Aken *et al.*, 2017a). Only the arrival and departure times are designed and the running time between stations are calculated (Zhang *et al.*, 2019b). Macroscopic representations do not guarantee that the developed timetable will be operationally feasible at the microscopic level. *Block sections*

2.6 Possession Scheduling

are the basic microscopic elements needed to model train movements. A block section is the piece of track between two consecutive train authorisation signals. At most one train may be within a block section at any given time. In microscopic representations, train movements are modelled on block sections. This guarantees that the timetable is operationally feasible, *i.e.* that there are no conflicting train movements (Zhang *et al.*, 2020b). Further benefits of microscopic representations are that running times and minimum train headways are modelled with more accuracy which allows a better assignment of railway capacity to trains and that the timetabling process is more efficient because train conflicts are resolved. The main drawback of the microscopic modelling approach is that the higher level of detail increases the size of the models dramatically. Even so, the recent trend is to model train timetables as much as possible at the microscopic level (Zhang *et al.*, 2019b). Mesoscopic rail network representations combine elements from the macro and micro perspectives. For example, Zhang *et al.* (2020b) use micro representations for the stations and macro representations for the tracks between them. Among the possession scheduling publications in Table 2.7, most authors make use of macroscopic rail network representations. However, from 2017 onwards, a few authors made use of microscopic models (Arenas *et al.*, 2018; D'Ariano *et al.*, 2019; Luan *et al.*, 2017; Zhang *et al.*, 2019b). One example of a mesoscopic model was found Zhang *et al.* (2020b).

Maintenance resources must to be assigned to the possessions to complete the work. These resources, such as crews and machines, may be subject to restrictions such as availability, worktime restrictions and minimum rest times. Most of the reported work do not schedule maintenance resources. Some of the publications that do are Lidén (2020), Kalinowski *et al.* (2020), Lidén *et al.* (2018), Lake *et al.* (2000), Higgins *et al.* (1999) and Higgins (1998).

Four categories of objective functions were identified. Objective functions are related to:

1. Minimising deviations from a reference train timetable
2. Minimising maintenance costs
3. Minimising train travel times
4. Capacity

2.6 Possession Scheduling

Objective functions from the first category appeared most. There are authors that minimise the arrival time deviations of trains (Luan *et al.*, 2017). Bešinović *et al.* (2020) minimise deviations along with turning activities of trains and non-commercial stops. D'Ariano *et al.* (2019) developed a bi-objective goal function that minimises deviations and maximises the number of paired maintenance activities. Other authors from the first category minimise train delays (Arenas *et al.*, 2018; Bueno *et al.*, 2019; Van Aken *et al.*, 2017a,b). Some authors minimise train delays and other features such as train cancellations (Forsgren *et al.*, 2013; Veelenturf *et al.*, 2016; Zhu *et al.*, 2018), maintenance delays (Albrecht *et al.*, 2013) and maintenance completion time (Higgins, 1998; Higgins *et al.*, 1999). Vansteenwegen *et al.* (2016) minimise the number of cancelled trains.

The second category of objective functions minimise maintenance costs (Budai *et al.*, 2006; Budai-Balke, 2009; Famurewa *et al.*, 2016; Lake *et al.*, 2000). Some authors combine maintenance costs with other elements such as:

- Cost of train operations (Lidén, 2020; Lidén & Joborn, 2017; Lidén *et al.*, 2018)
- Impact of maintenance projects on railway operations (Peng *et al.*, 2011)
- Total train travel time (Zhang *et al.*, 2020a, 2019b)

Minimising train travel times is the third objective function category. Within this category there are authors that minimise the total travel time of trains (Zhang *et al.*, 2018, 2019a). While, some authors combine the minimisation of train travel time with other features such as:

- Minimising the number of train cancellations (Zhang *et al.*, 2020b)
- Minimising the maintenance cost (Zhang *et al.*, 2020a)
- Minimising the maintenance tardiness cost (Zhang *et al.*, 2019b)
- Minimising the maintenance activity times (Bababeik *et al.*, 2019)

Only a few authors make use of objective functions related to railway capacity. Kalinowski *et al.* (2020) minimise the overall capacity reduction. Whereas, Boland *et al.* (2013) maximise the total annual throughput while minimising deviations from the original maintenance schedule.

2.6 Possession Scheduling

Section 2.5 discussed scheduling techniques in general. For the possession scheduling problem specifically, examples of integer programming (IP), integer linear programming (ILP), mixed integer programming (MIP), mixed integer linear programming (MILP) and constraint programming models were found. IP scheduling models make use of only discrete decision variables. Whereas, MIP scheduling models make use of both discrete and continuous decision variables. IP and MIP are comprehensive terms that refer to models that may contain linear or quadratic features. While, ILP and MILP refer exclusively to linear models, *i.e.* models without quadratic features in the objective function or constraints. Constraint programming was discussed in Subsection 2.5. Among the reported work, most of the authors make use of MILPs and MIPs. From 2017 onwards, MILPs dominate. Recent examples of MILPs include Bešinović *et al.* (2020), Lidén (2020), D'Ariano *et al.* (2019), Zhang *et al.* (2019b) and Zhang *et al.* (2019a). While recent examples of MIPs include Kalinowski *et al.* (2020), Van Aken *et al.* (2017b) and Famurewa *et al.* (2016). A few authors made use of IPs (Higgins, 1998; Louwerse & Huisman, 2014) and ILPs (Veelenturf *et al.*, 2016; Zhang *et al.*, 2020a). While, one example of a binary IP (Zhang *et al.*, 2020b) and one example of a constraint satisfaction model (Cheung *et al.*, 1999) was identified.

Five modelling approaches were identified: Time-space network, Big-M, Periodic Event Scheduling Problem (PESP) and job shop scheduling formulations, as well as simulation based. Among the identified modelling approaches the time-space network formulations appeared most, examples include Zhang *et al.* (2020a), Zhang *et al.* (2020b), Lidén (2020), Lidén & Joborn (2017) and Luan *et al.* (2017). Examples of Big-M formulations include D'Ariano *et al.* (2019), Zhang *et al.* (2019b) and Arenas *et al.* (2018). While, examples of PESP formulations include Bešinović *et al.* (2020) and Van Aken *et al.* (2017a,b). Burdett & Kozan (2009) is an example of a job shop scheduling formulation. Whereas, Bahramian & Bagheri (2015) presented a simulation based approach.

Techniques for optimising schedules were discussed in Section 2.5. Most of the authors made use of commercial solvers, such as Gurobi and CPLEX, to solve their models. For example, Bešinović *et al.* (2020), Lidén (2020), Lidén (2020), Zhang *et al.* (2019a) and Zhang *et al.* (2018). Metaheuristics was the second most used technique. Even though, only a few applications of metaheuristics were found:

- Tabu search – Brucker *et al.* (2002), Higgins *et al.* (1999) and Higgins (1998)

2.6 Possession Scheduling

- Simulated annealing – Bueno *et al.* (2019), Burdett & Kozan (2009) and Lake *et al.* (2000)
- Problem space search – Albrecht *et al.* (2013)
- Genetic and memetic algorithms – Budai-Balke (2009)
- Ant colony optimisation – Khalouli *et al.* (2018)

Bahramian & Bagheri (2015) used a simulation based method with a genetic algorithm. Heuristic methods were used by Kalinowski *et al.* (2020), Zhang *et al.* (2019b), Arenas *et al.* (2018), Peng *et al.* (2011), and Budai *et al.* (2006). Lagrangian relaxation was used by Zhang *et al.* (2020a), Luan *et al.* (2017), Caprara *et al.* (2006) and Zhang *et al.* (2020b). Zhang *et al.* (2020b) also used an Alternating Direction Method of Multipliers (ADMM) technique.

Possession scheduling problems are usually based on real world railways or data. Examples were found for railways in:

- Australia – Kalinowski *et al.* (2020), Boland *et al.* (2013)
- China – Zhang *et al.* (2020b), Zhang *et al.* (2020a), Zhang *et al.* (2019a), Zhang *et al.* (2018)
- France – Arenas *et al.* (2018)
- Germany – Brucker *et al.* (2002)
- Netherlands – Bešinović *et al.* (2020), Zhu *et al.* (2018), Van Aken *et al.* (2017a,b), Veelenturf *et al.* (2016)
- Sweden – Lidén (2020)

A railway may be used exclusively for passenger or freight trains. Otherwise, a mix of passenger and freight trains operate on a railway. Most of the reported work focused on passenger trains. Examples of these are Zhang *et al.* (2020b), Zhang *et al.* (2020a), Zhang *et al.* (2019a), Zhang *et al.* (2019b) and Bababeik *et al.* (2019). Examples of possession planning on freight railways include Kalinowski *et al.* (2020) and Bueno *et al.* (2019). While, examples of problems based on mixed service railways include Lidén (2020), Bešinović *et al.* (2020) and Arenas *et al.* (2018).

2.6 Possession Scheduling

Planning horizons for most of the schedules range from one hour to seven days. One example was found with a schedule of less than an hour and one example was found with a planning horizon of more than seven days. Examples of planning horizons are:

- 30 minutes – [Van Aken et al. \(2017a\)](#)
- In the order of hours – [Zhang et al. \(2020b\)](#), [Bababeik et al. \(2019\)](#), [Bešinović et al. \(2020\)](#), [D'Ariano et al. \(2019\)](#), [Zhu et al. \(2018\)](#)
- In the order of days – [Lidén \(2020\)](#), [Bueno et al. \(2019\)](#), [Arenas et al. \(2018\)](#), [Forsgren et al. \(2013\)](#), [Albrecht et al. \(2013\)](#)
- Annual – [Kalinowski et al. \(2020\)](#)

The solution precision, *i.e.* the size of the smallest time intervals, range between seconds, minutes and hours. Usually, intervals of one minute are used. Examples of time intervals include:

- Seconds – [Bešinović et al. \(2020\)](#), [Zhang et al. \(2019b\)](#), [Zhu et al. \(2018\)](#)
- Minutes – [Zhang et al. \(2020a\)](#), [D'Ariano et al. \(2019\)](#), [Bueno et al. \(2019\)](#), [Arenas et al. \(2018\)](#), [Luan et al. \(2017\)](#)
- Hours – [Kalinowski et al. \(2020\)](#), [Lidén et al. \(2018\)](#), [Lidén \(2018b\)](#)

A wide range of schedules have been developed in terms of the number of trains, stations and block sections included. The minimum, maximum and averages for these characteristics are:

- Number of trains: min = 10, max = 350, average = 120.
- Number of stations: min = 8, max = 60, average = 30.
- Number of block sections: min = 80, max = 1000, average = 540.

2.6.2 Synthesis

Based on the trends presented in the previous subsection, the following conclusions may be drawn:

2.6 Possession Scheduling

- The number of publications related to possession scheduling have increased significantly since 2016. Nine related works were found from 2011 to 2015, while 35 related works were found from 2016 to 2020.
- Possession scheduling is usually treated as tactical planning problem.
- Both sequential and integrated train and possession planning approaches are used, but recently, there has been a sharp increase in integrated approaches.
- The most active related work categories in recent years has been “combined scheduling of maintenance and trains” and “train scheduling with fixed maintenance closures”.
- Only a few of the related publications make use of periodic train and/or possession scheduling.
- Mostly macroscopic models have been used to model the rail network. However, the recent trend is to use microscopic models as much as possible.
- Most of the publications do not schedule maintenance resources.
- Four objective function categories were identified. Objective functions related to: (1) Minimising deviations from a reference train timetable, (2) Minimising maintenance costs, (3) Minimising train travel times, and (4) Capacity. Objective functions from the first category appeared most often. Whereas, only a few examples from the fourth category were identified.
- The possession scheduling problem is mostly modelled with MILPs and MIPs. From 2017 onwards, mostly MILPs were used. Few examples of other models were found.
- Five modelling approaches were identified: Time-space network, Big-M, Periodic Event Scheduling Problem (PESP), and job shop scheduling formulations, as well simulation based. Of these approaches, time-space network formulations appeared most, while few examples of job shop scheduling formulations and simulation based approaches were identified.

2.6 Possession Scheduling

- Commercial solvers are the most popular optimisation technique for possession scheduling problems. Metaheuristics are the second most popular optimisation technique. Even so, only a few applications of tabu search, simulated annealing, problems space, genetic algorithms and ant colony optimisation were found. A few examples of heuristics and Lagrangian relaxation were also found.
- Most of the possession scheduling problems are based on real world railways or data.
- Most of the case studies consider passenger train services. A few examples of freight and mixed-use railways were also found.
- Most of the planning horizons range between one hour and seven days.
- The solution precision ranges from seconds to hours. Usually, the solution precision is one minute.
- There is a wide range in the number of trains, stations and block sections included in possession schedules. The number of trains range from 10 to 350, the number stations range from 8 to 60 and the number of block sections range from 80 to 1000.

Chapter 3

Case of Application

The case where the maintenance possession scheduling model will be applied is described in this chapter.

3.1 Railway from Bellville to Wellington

The Cape railway corridor stretches from the Western Cape to Gauteng. The maintenance possession scheduler is applied to a segment of this corridor, namely Bellville to Wellington. A map of this segment is shown in Figure 3.1. The segment is maintained by the Bellville infrastructure maintenance depot and has the following characteristics:

- Double track layout between stations with one line dedicated to each direction of travel
- Nine stations where trains may pass one another
- Gauge of 1067 mm
- Track load specification of 20 tonne per axle
- 3 kV DC overhead traction power supply
- Lineside signalling traffic authorisation system with centralised traffic control
- There is a train marshalling yard in Bellville
- The total length of the segment is 53 km

3.1 Railway from Bellville to Wellington

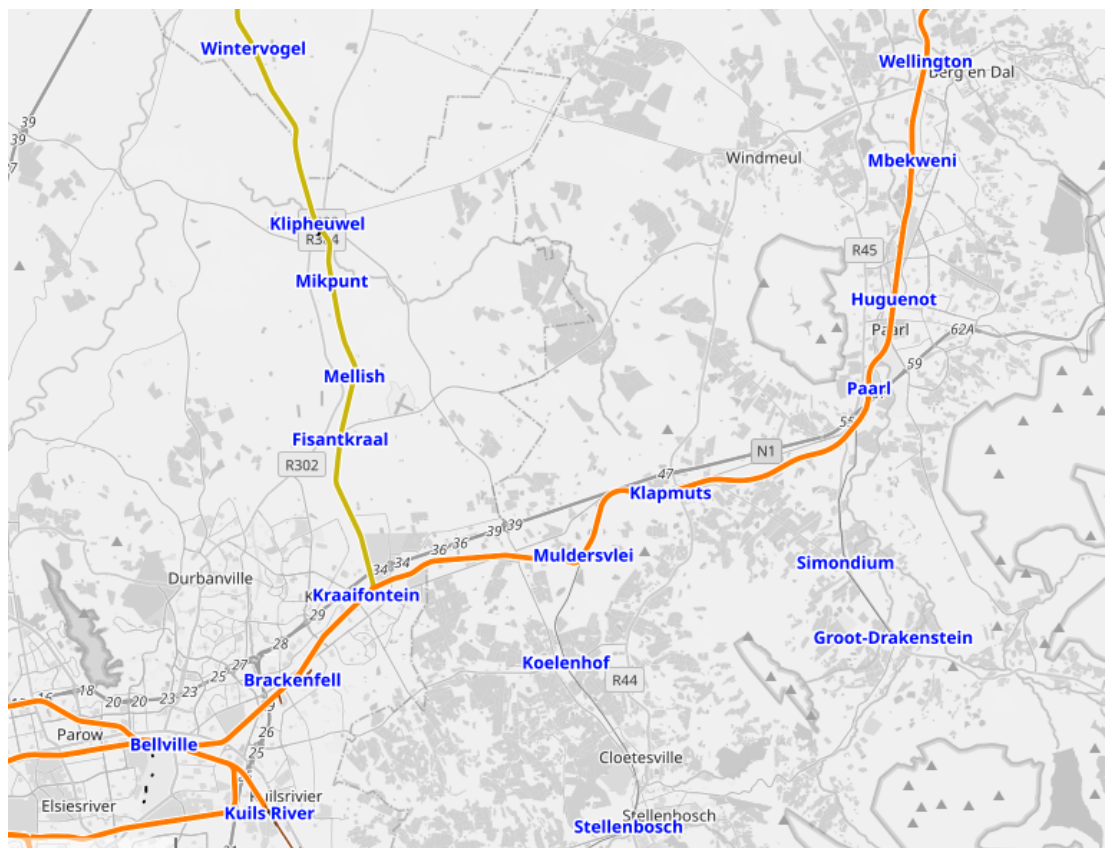


Figure 3.1: Railway from Bellville to Wellington ([OpenRailwayMap website](https://www.openrailwaymap.org/), 2021)

3.1 Railway from Bellville to Wellington



Figure 3.2: Macroscopic representation of the railway from Bellville to Wellington

Passenger and goods trains run on the segment. Passenger trains travel to and from Cape Town, Worcester, East London, Johannesburg and Pretoria. Goods trains travel to and from Cape Town, Bellville, Paarl, Worcester, Kimberley, Johannesburg and Pretoria. Commodities include grain, maize, fertilizer, containers, cement, slag iron, steel, bulk liquid and coal. Goods trains haul between 10 and 50 wagons per train.

A macroscopic representation of the railway line is shown in Figure 3.2. There are nine stations from Bellville to Wellington where trains may pass one another. Between the stations there are double track segments with dedicated directions of travel. One mainline track is dedicated to trains travelling from Bellville to Wellington, referred to as the “down line”, while the other mainline is dedicated to trains travelling from Wellington to Bellville, referred to as the “up line”. If a maintenance possession is taken on one of the mainlines between stations, trains may be authorised to travel on the other mainline against the normal direction of travel. This practice is referred to as “wrong road working” and is subject to specific train authorisation procedures to ensure safe train movements. This practice enables trains to be rerouted around single line maintenance possessions on double track segments.

Based on the process described in Subsections 2.2.2.1 and 2.2.2.8, a master train schedule (MTS) is developed annually for the railway. The MTS defines the planned departure, travel and arrival times for every train. An allocation of space and time on the schedule is referred to as a *train slot*. A train slot specifies the specific times that a train is planned to run from station to station. A week before execution of the train services a selection of trains on the MTS is confirmed to run based on the passenger demand, customer orders, available resources and maintenance possessions. This confirmed weekly train plan is referred to as the integrated train plan (ITP).

The full MTS for the Bellville to Wellington segment has 55 train slots per day:

- 18 train slots for passenger trains – nine in each direction

3.2 The Monthly Maintenance Planning Process of Bellville Depot

- 37 train slots for goods trains – 17 in the direction of Wellington and 20 in the direction of Bellville

On every day of the week, a subset of these train slots are made available to specific train services according to their service designs.

3.2 The Monthly Maintenance Planning Process of Bellville Depot

Transnet Freight Rail has regional maintenance depots that maintain the rail network. The high-level maintenance processes are standardised across these depots. Although, the detailed tasks, timelines and forms of a process may be bespoke at a specific depot. There is a programme in the organisation aimed at standardising and improving the detailed processes. In the mean time, this subsection presents the detailed monthly maintenance planning process of the Bellville depot.

Figure 3.3 shows the monthly maintenance planning process of the Bellville maintenance depot. The monthly maintenance planning meeting is standardised to ensure that the required outcomes are achieved. The meeting and how it fits into the possession scheduling process of Bellville depot is discussed in Subsection 3.2.2.

3.2.1 Maintenance Possession Scheduling at Transnet Freight Rail

Maintenance *possessions* provide maintenance personnel with the authority to occupy a track section for the purpose of maintenance. Maintenance possessions are referred to as *occupations* in Transnet Freight Rail.

There are three types of maintenance possessions: major, minor and unplanned. Major possessions prohibit the movement of trains within the occupied track section. Major possessions are referred to as *total occupations* in TFR. Minor possessions allow trains to operate within the occupied track section. Minor possessions are referred to as *occupations-between-trains* in TFR. Unplanned possessions are authorised when there is an infrastructure breakdown with an operational impact. Unplanned possessions are referred to as *emergency occupations* in TFR.

The maintenance department can negatively affect the train service in three ways:

3.2 The Monthly Maintenance Planning Process of Bellville Depot

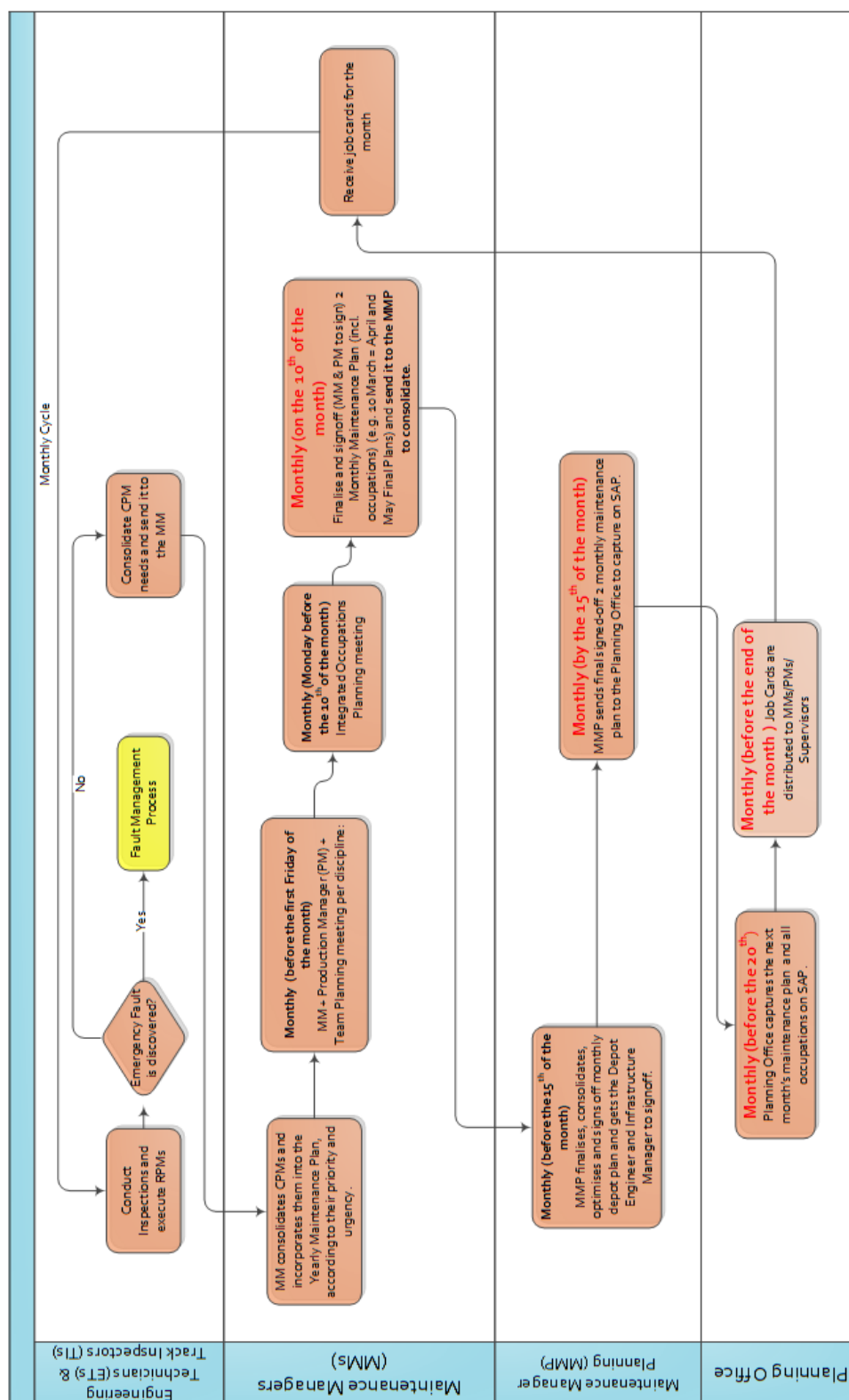


Figure 3.3: The monthly maintenance planning process of Bellville depot

3.2 The Monthly Maintenance Planning Process of Bellville Depot

1. Maintenance possessions remove track sections from service and reduce track availability.
2. Infrastructure breakdowns that have an operational impact lead to train delays or cancellations.
3. Temporary speed restrictions reduce the operational risk of track sections that are in a poor condition, but increase the train cycle times.

Maintenance possessions are needed to do the maintenance that will increase reliability, reduce breakdowns and uplift the temporary speed restrictions.

The operational impact of a major possession depends on the network layout. In the case of a single track, a major possession will restrict all train movements in the track section. In the case of a double track, a major possession may be authorised on one, or both tracks, depending on the work required. If a major possession is authorised on both tracks of a double track, no train movements will be allowed in the track section. If a major possession is authorised on only one track of a double track, train movements may proceed on the unoccupied track. In this case, the operational impact is less than a total closure of the track section. The same principle applies in areas of the network that have multiple tracks in parallel with one another, for example in marshalling yards. Examples of maintenance that require a major possession on both lines of a double track are turnout replacements and substation maintenance. Examples of maintenance that only require a major possession on one line of a double track include maintenance done with on-track maintenance machines, rail welding and maintenance of the electrical overhead track equipment.

Since major track possessions remove track sections from service, the number of major possessions must be minimised to maximise track availability. The number of major possessions can be decreased by coordinating and combining major possessions among the different maintenance disciplines per track section. Major possessions that are combined among different maintenance disciplines are referred to as *block occupations* in TFR. The benefits of coordinating possessions include better track availability and reduced costs from sharing resources. Resources that can be shared include transportation vehicles and traction linesmen that are responsible for electrically isolating

3.2 The Monthly Maintenance Planning Process of Bellville Depot

the occupied track section. A drawback of coordinated possessions is that it may increase complexity for the maintenance planning and execution teams since they need to coordinate with one another.

The possession schedule determines when specific track sections are taken out of service for maintenance work. A conflict exists between scheduling major possessions and trains since major possessions may interfere with the scheduled train service. Rail operations generate revenue in the short term. Whereas, maintenance is an investment in future revenue by providing capacity assurance, improved throughput, quality of operations and reduced operating costs, as mentioned in Subsection 2.3.1.3. As shown in Figure 2.39 and Figure 2.43, the minimum required amount of maintenance must be done to operate assets sustainably and minimise the total cost of reliability. As shown in Figure 2.60, deferred maintenance accelerates track degradation and shortens the track life. Deferred maintenance may also increase the direct costs of restoring the track and the indirect costs from infrastructure breakdowns and unreliable operations. Therefore, the minimum required number of possessions must be scheduled. At the same time, interference between major track possessions and the scheduled train service must be minimised.

Train planners have three options to accommodate major possessions on the train schedule:

1. Depart trains earlier or later.
2. Redirect trains on an alternative path, if such a path exists.
3. Cancel the trains.

Alternatively, if the operational impact of a possession request is unacceptable, the possession may be deferred. In which case, alternative dates and times are proposed to the possession requester. The possession approval process will repeat until a suitable date and time is agreed. The possession approval process is discussed in more detail in the next subsection.

There are possession scheduling considerations from both a train service and track maintenance perspective. Possession scheduling considerations from a train service perspective include:

3.2 The Monthly Maintenance Planning Process of Bellville Depot

- Track availability must be maximised by minimising the number of major possessions.
- Adherence to the train schedule must be maximised by minimising interference between scheduled trains and major possessions.
- Major possessions must be avoided during the peak times of commuter travel, on the applicable track sections.
- Luxury long distance passenger trains may not be delayed or cancelled.
- On double track sections it is preferred that possessions only be taken on one track at a time so that the other track remains available for the train service.

Possession scheduling considerations from a maintenance perspective include:

- Certain on-track maintenance machines are scheduled for maximum utilisation on the national network. Therefore, possessions for those on-track machines are prioritised.
- Certain on-track maintenance machines must to be scheduled in a specific sequence for technical reasons.
- Few possessions of a long duration are preferred over many possessions of a short duration since each possession will incur set up and travel costs.
- Possessions are scheduled and executed as soon as possible. If the majority of the work on the annual maintenance plan is complete before the annual budgeting cycle begins, it supports the motivation for the following year's maintenance budget.

3.2.2 The Maintenance Possession Processes of Transnet Freight Rail

Figure 3.4 shows an overview of the maintenance possession process of TFR. Figure 3.5 shows the monthly possession planning process of the Bellville maintenance depot. Figure 3.6 and Figure 3.7 shows the possession countdown and escalation sub-processes of the Bellville depot, respectively. The maintenance planning technicians are organised into track section planning teams across disciplines. These multidisciplinary teams

3.2 The Monthly Maintenance Planning Process of Bellville Depot

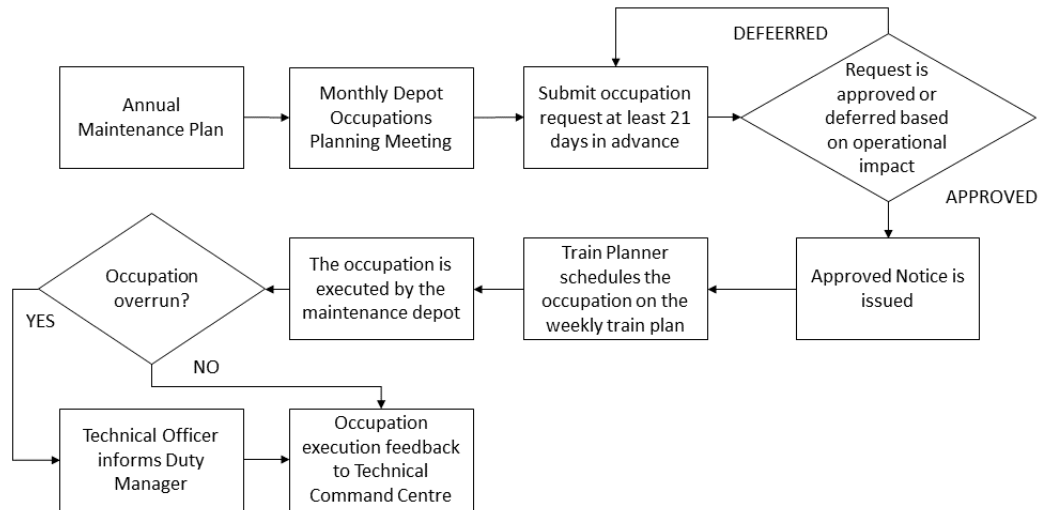


Figure 3.4: Overview of the maintenance possession process at Transnet Freight Rail

integrate the possession requirements per track section to ensure that the required possessions are coordinated and combined. The possession requests are approved by the maintenance managers of every discipline to ensure that there are no possession clashes between disciplines.

The maintenance possession planning processes and meetings are standardised to ensure that the required outcomes are achieved. The section planning meeting is attended by maintenance planning technicians from the track, electrical, signalling, technical support and telecommunications maintenance disciplines. Each technician comes prepared with a list of possessions that must be scheduled for their discipline. The section planning team identifies opportunities to coordinate and combine possessions. A possession schedule is then proposed for the track section.

After the multidisciplinary section planning meetings, a monthly maintenance planning meeting is held per discipline. The meeting is chaired by the maintenance manager while the production manager and maintenance planning technicians attend. Meeting inputs include the monthly depot performance report and preliminary maintenance plans for every track section. The preliminary maintenance plans include the proposed possession schedules as determined in the section planning meetings. The following points are discussed:

3.2 The Monthly Maintenance Planning Process of Bellville Depot

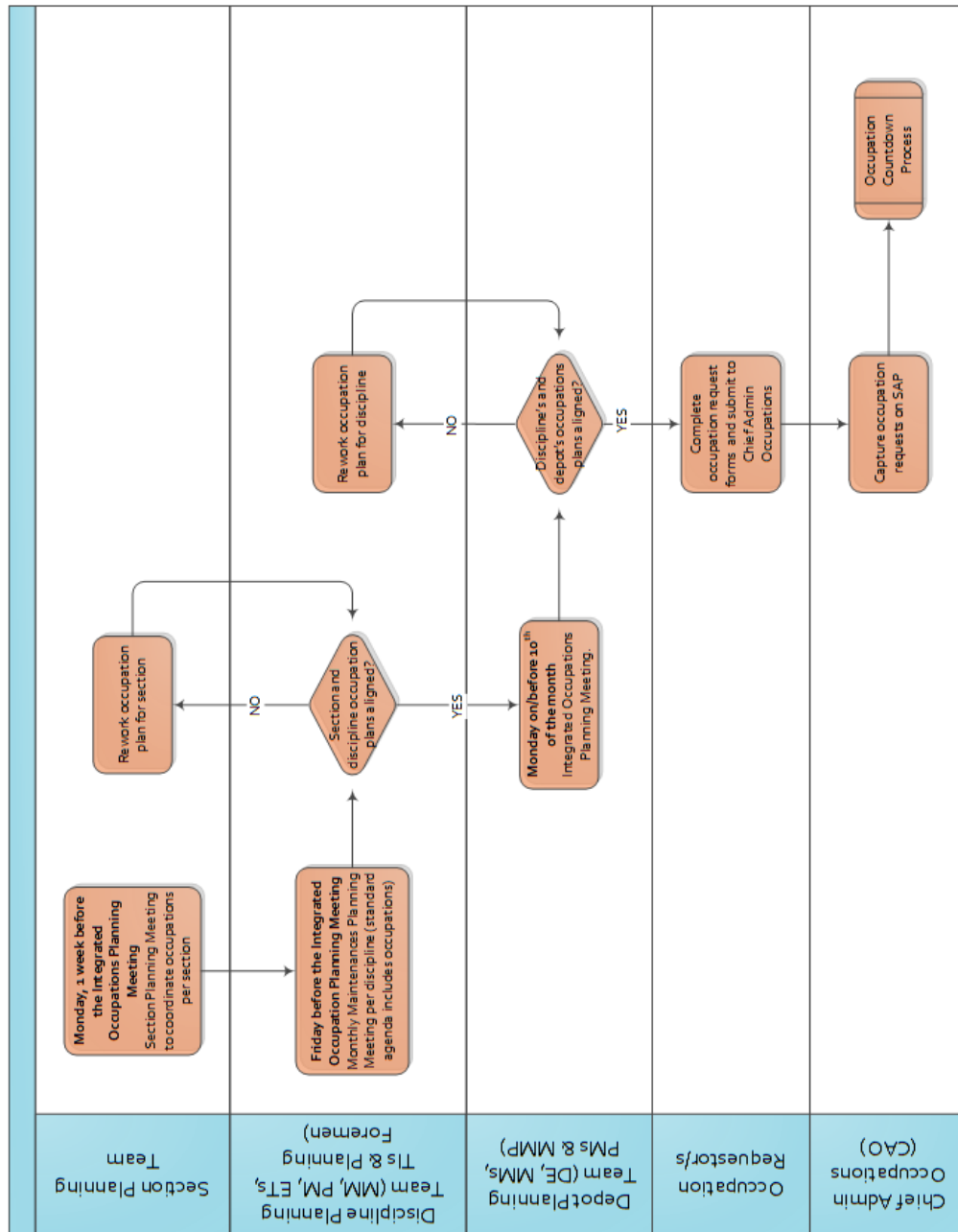


Figure 3.5: The monthly possession planning process of Bellville depot

3.2 The Monthly Maintenance Planning Process of Bellville Depot

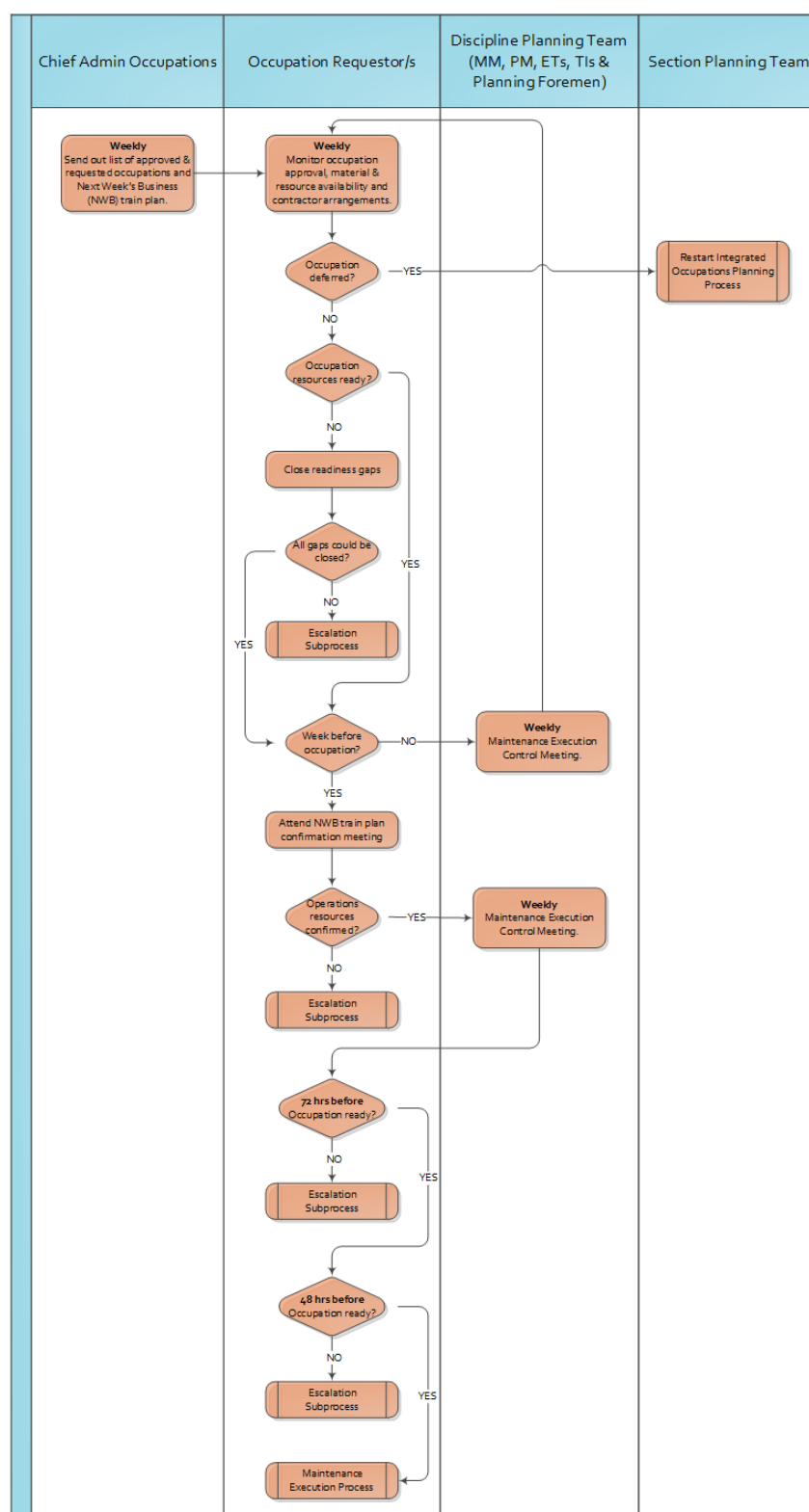


Figure 3.6: The possession countdown process of Bellville depot

3.2 The Monthly Maintenance Planning Process of Bellville Depot

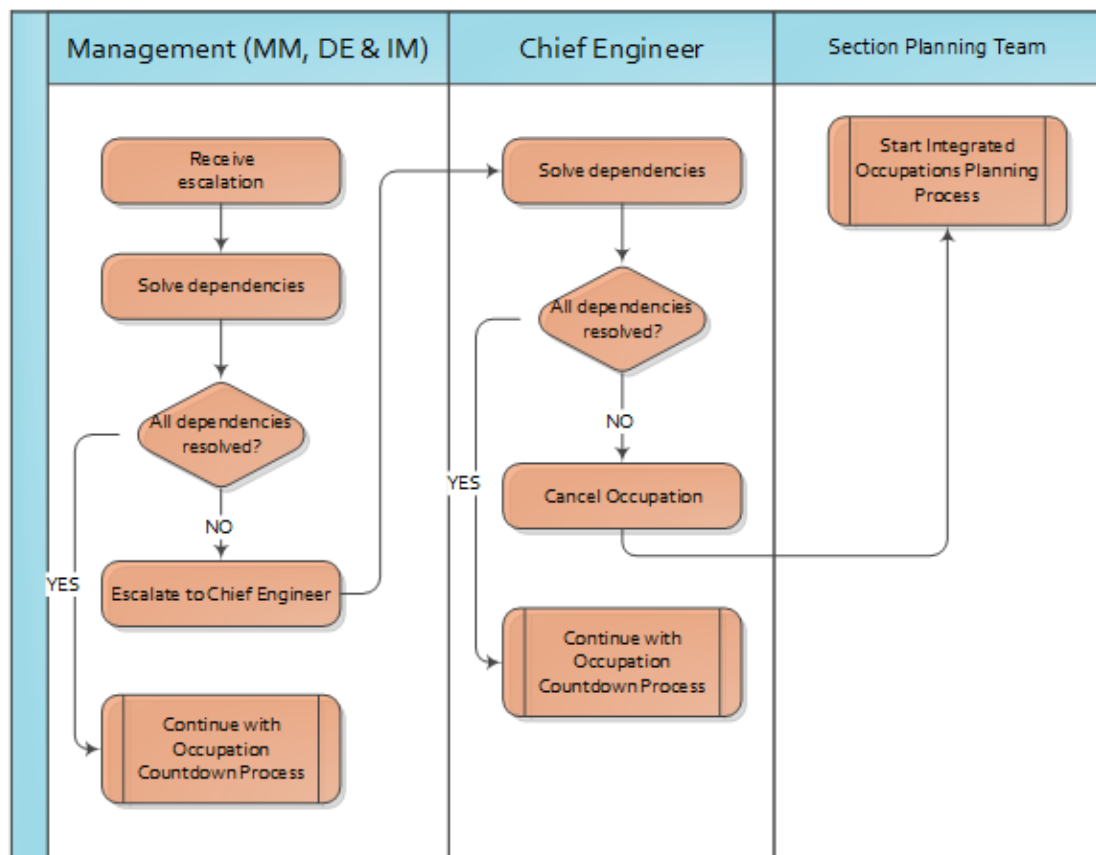


Figure 3.7: The possession countdown escalation process of Bellville depot

3.2 The Monthly Maintenance Planning Process of Bellville Depot

- Monthly maintenance review: open and closed job cards, breakdowns and rescheduling of outstanding work.
- Possessions: location, confirmation of switching points, type of possessions, dates and times.
- Resource allocation: personnel, vehicles, material availability and lead times.
- Finalise the rolling two-month maintenance plan and possession schedule.

After the monthly maintenance meeting per discipline, a depot possession planning meeting is held to confirm the integrated possession plan for the depot. This meeting is referred to as the integrated occupations planning meeting. The meeting is chaired by the Maintenance Manager Planning (MMP) while the depot engineer, maintenance managers, production managers, representatives from local train operations and other relevant stakeholders attend. Inputs to the meeting include the proposed maintenance plan from each discipline. The following possession checklist is confirmed, first for on-track machine possessions and then for all other possessions:

- Block occupations planned where feasible?
- Location confirmed?
- Time and duration confirmed?
- Switching points verified?
- Maintenance teams confirmed?
- Material lead times considered?

Once the details for the possessions are confirmed, the depot's possession plan and the changes required by each discipline to align to the depot's possession plan are finalised. The maintenance planning technicians then complete the possession requests and submit it to the chief administration occupations (CAO). The CAO loads the possession requests on the SAP enterprise resource scheduling system for approval by the National Command Centre (NCC).

3.2 The Monthly Maintenance Planning Process of Bellville Depot

The possessions requests are approved or deferred based on the operational impact. The selected electrical switching points must also be correct. The switching points are devices that electrically isolate the occupied track section.

Maintenance depots that do not follow a standardised possession planning process may experience challenges with regards to possession approvals. Some of these challenges and their root causes are:

1. Possession clashes occur when more than one maintenance discipline apply for a possession in the same track section on the same date and time. In this case, only one possession can be approved and the other will be deferred. Root cause: silo possession planning.
2. Incorrect switching points. Root cause: the possession requestor did not verify the switching points with the electrical maintenance department.
3. Possession requesters not following up on deferred possessions. Root cause: no standard countdown process.
4. Approved possessions cancelled by local operations after approval by the NCC. Root cause: local operations not consulted during the possession planning process.

Figure 3.8 shows the NCC approval process of maintenance possessions. Once the electrical switching points have been validated, the NCC occupations manager does the train service impact analysis. The impact analysis is done by referring to the train diagrams and identifying which train slots will be affected by the possessions. The NCC occupations manager then submits a summary of the impact analysis to the NCC train service managers. Next, the train service managers determine which trains to delay, plan earlier, reroute, or cancel to accommodate the possessions. These train service adjustments are made to minimise the impact on operations as much as possible. The next step is the weekly occupations meeting. At the meeting, key stakeholders determine which possessions to approve and which to defer. The occupations manager chairs the meeting while representatives from the train service, customer care, technical command centre and rolling stock departments as well as the Passenger Rail Association of South Africa attend.

After a possession is approved, a notice is issued. The notice serves as an authority that grants permission to the maintenance team to take possession of the track section

3.2 The Monthly Maintenance Planning Process of Bellville Depot

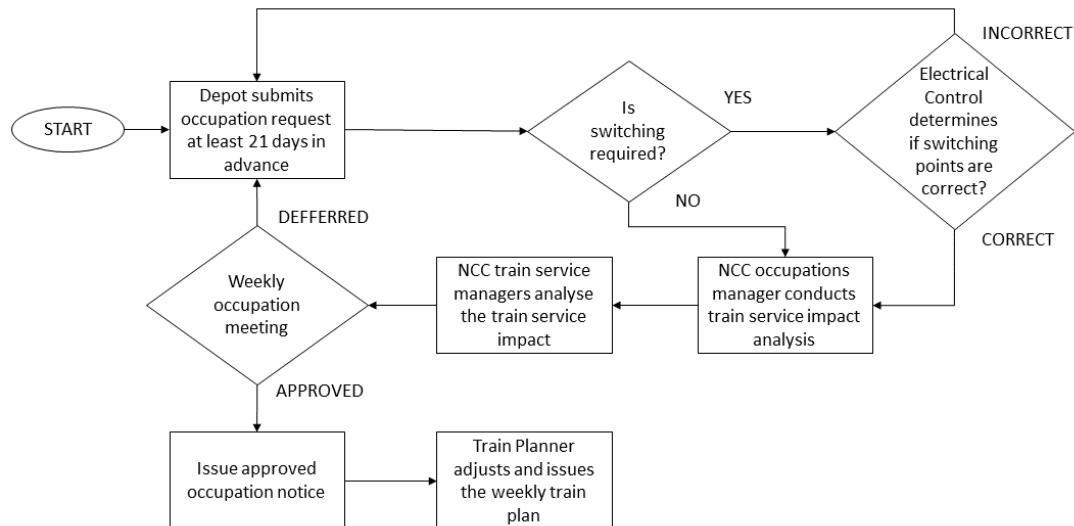


Figure 3.8: The NCC approval process for maintenance possessions

and do the necessary work. The notice is issued to all the stakeholders. The notice includes the full particulars of the maintenance possession. This includes the place, time, date, duration, nature of work, switching points, details of the person who will take the possession and the details of persons whose cooperation is required to complete the work. The notice also indicates the safety and other additional measures that must be taken, *e.g.* the provision of hand signalmen or adjusted train authorisation methods. Once the notice is issued, the train planner adjusts the weekly train plan to accommodate the maintenance possession.

The possession countdown and escalation processes show how the possession requester monitors the readiness of resources to execute the possession. If some of the resources are not ready, the possession may have to be cancelled. If the resources are ready, the maintenance execution teams can proceed with the possession. During the possession, the technical officer in control of the possession monitors the execution. If there are deviations that will delay completion of the possession, the technical officer informs the NCC. The NCC will then make arrangements to minimise disruptions of the train service.

3.2 The Monthly Maintenance Planning Process of Bellville Depot

3.2.3 Maintenance Possession Performance Indicators

Performance indicators of maintenance possessions include:

- Number of trains cancelled due to possessions
- Number of trains delayed due to possessions
- Number of possessions cancelled
- Number of possession delayed to start
- Number of possessions delayed to finish
- Percentage of possessions approved
- Percentage of possession executed
- Rail network availability

Chapter 4

Model

The railway maintenance possession scheduling model that was developed is described in this chapter. Firstly, the components of the microscopic infrastructure representation are described. This is followed by a description of the assumptions, objective function, constraints and validation.

4.1 Components of the Microscopic Infrastructure Representation

The developed model is based on a microscopic representation of the rail network infrastructure. This modelling approach was described in Subsection 2.6. This level of detail allows train movements to be modelled in block sections. In the model, two types of block sections are defined: departure block sections and arrival block sections. A departure block section is defined as a train route from a departure signal at one station to the home signal at the next station. An arrival block section is defined as a train route from a home signal outside a station to a departure signal within the same station. The two types of block sections are illustrated in Figure 4.1. The other elements of the microscopic infrastructure representation are nodes, links and cells. Figure 4.2 shows a microscopic station layout with numbered nodes, links and cells.

4.1 Components of the Microscopic Infrastructure Representation

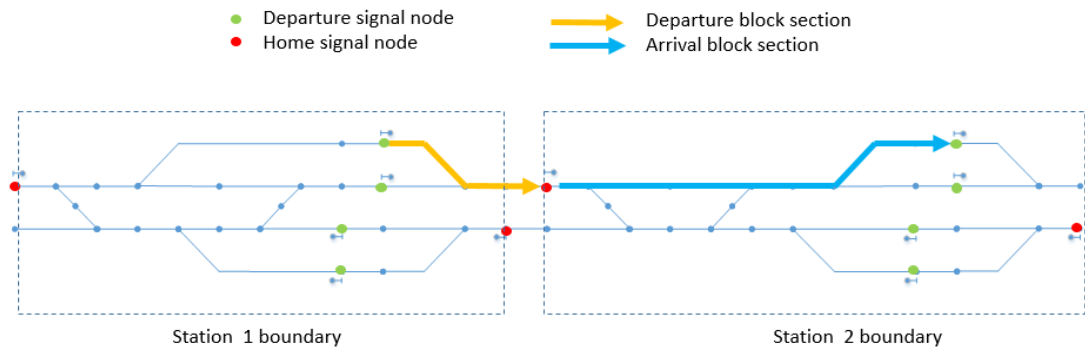


Figure 4.1: Definition of block sections

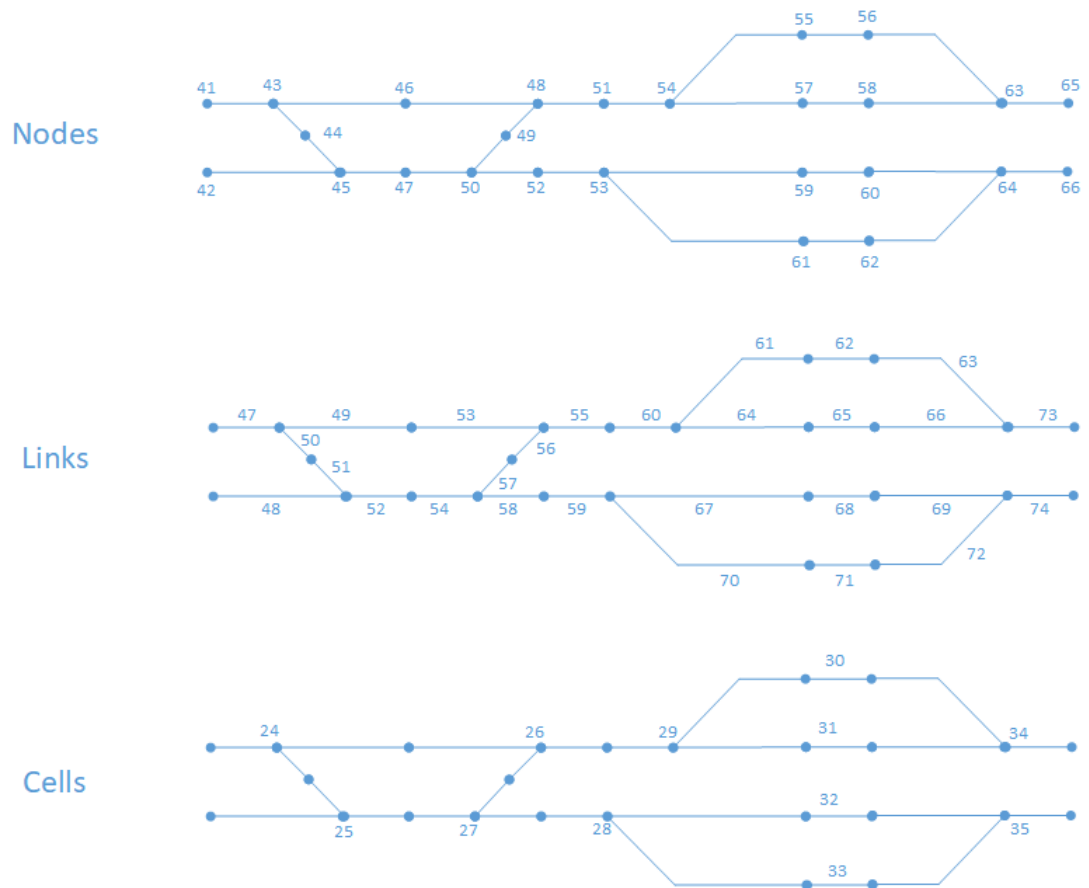


Figure 4.2: Elements of a microscopic station layout

4.2 Assumptions

The following assumptions were made:

- The reservation and release times are the same for all trains on all block sections.
- Time supplements for robust scheduling have already been added to the running times used for train planning.
- At the origin and destination stations, there are dedicated departure and arrival lines for up and down trains, respectively.

4.3 Definition of Sets, Parameters and Decision Variables

The sets, parameters and decision variables that are used in the model are defined in Tables [4.1](#), [4.2](#), and [4.2](#), respectively.

4.3 Definition of Sets, Parameters and Decision Variables

Table 4.1: Set definitions

Set	Definition
R	Set of trains, $r \in R$
S	Set of stations, $s \in S$
B	Set of block sections, $b \in B$
C	Set of cells, $c \in C$
L	Set of links, $l \in L$
N	Set of nodes, $n \in N$
MOT	Set of maintenance possessions, $m \in MOT$
S_r	Set of stations that train r travels through, $S_r \subseteq S$
S'_r	Set of stations that train r travels through, excluding the origin and destination stations $S'_r \subset S$
B_r	Set of block sections that train r may travel through, $B_r \subset B$
B^a	Set of arrival block sections, $B^a \subset B$
$B_{r,n}^+$	Set of block sections that flow out of node n and train r may travel through, $B_{r,n}^+ \subset B$
$B_{r,n}^-$	Set of block sections that flow into node n and train r may travel through, $B_{r,n}^- \subset B$
$B_{r,n}^{a,-}$	Set of arrival block sections that flow into node n and train r may travel through, $B_{r,n}^{a,-} \subset B$
B_c	Set of block sections containing cell c , $B_c \subset B$
C_b	Set of cells in block section b , $C_b \subset C$
C_m	Set of cells included in maintenance possession m , $C_m \subset C$
L_b	Set of links for block section b , $L_b \subset L$
N_r	Set of block section nodes that train r may travel through, excluding the origin and destination nodes, $N_r \subset N$
N_s	Set of departure block section nodes within station s , $N_s \subset N$
$N_{r,s}$	Set of nodes which are the end nodes of the arrival block sections that train r may use to enter station s , $N_{r,s} \subset N$

4.3 Definition of Sets, Parameters and Decision Variables

Table 4.2: Parameter definitions

Parameter	Definition
n_r^o	Origin node index of train r
n_r^d	Destination node index of train r
t_r^s	Earliest departure time of train r from its origin node
t_r^e	Latest departure time of train r from its origin node
mot_m^s	Earliest start time of maintenance possession m
mot_m^e	Earliest start time of maintenance possession m
d_m	Minimum duration of maintenance possession m
s_r^o	Origin station index of train r
s_r^d	Destination station index of train r
l_b^{last}	Last link of arrival block section b
n_b^+	Start node of block section b
n_b^-	End node of block section b
$t_{r,l}$	Running time of train r on link l
$\varepsilon_{b,b'}$	0–1 relationship parameter, equal to one if block sections b and b' have cell in common and these are not arrival block sections with the same last cell, zero otherwise
$t_{r,s}^{dwell}$	Minimum dwell time for train r at station s
t_{res}	Reservation time required by a train before it can enter a block section
t_{rel}	Release time required after a train has exited a block section before the next train may reserve the block section
$t_{r,n_r^d}^{arri}$	The time that train r is normally scheduled to arrive at its destination node
p_r	The probability that train r is confirmed to run on the weekly train plan
M	A large number set for the constraints

Table 4.3: Decision variables definitions

Decision Variable	Definition
$y_{r,b}^{entr}$	Entrance time of train r on block section b
$y_{r,b}^{exit}$	Exit time of train r on block section b
$x_{r,b}$	0–1 route variable, equal to one if train r uses block section b , zero otherwise
$\mu_{r,b,r',b'}$	0–1 sequence variable, equal to 1 if train r is scheduled earlier on block section b than train r' on block section b' that is conflicting with block section b , zero otherwise
$t_{r,s}^{stop}$	Actual dwell time of train r at station s
t_m^{start}	Start time of maintenance possession m
t_m^{end}	End time of maintenance possession m
$\alpha_{r,b,c}$	0–1 variable, equal to one if the entrance time of train r on block section b is larger than or equal to the end time of maintenance possession m on cell c , zero otherwise

4.4 Objective Function

4.4 Objective Function

The objective function minimises the sum of the expected arrival time deviations of all trains at their destination nodes. The probability of each train being confirmed on the train plan is multiplied by the absolute value of the difference between its scheduled arrival time and its normal arrival time. The absolute value is used so that both positive and negative deviations from the normal train schedule is minimised. The scheduled arrival time of train r is equal to the exit time of train r on its final arrival block section plus the running time of train r on the last link of the arrival block section. The objective function is

$$\text{Minimise } Z = \sum_{r \in R} p_r \sum_{b \in B_{r,n_r^d}^-} |y_{r,b}^{exit} + x_{r,b}(t_{r,l_b^{last}} - t_{r,n_r^d}^{arri})|. \quad (4.1)$$

4.5 Constraints

The constraints used for the combined scheduling of trains and maintenance possessions are described in this section.

4.5.1 Train Movement Constraints

Block Section Usage Constraints

A big M formulation is used in constraint (4.2) to ensure that train r is assigned a positive entrance time on block section b only if it makes use of block section b . Constraint (4.3) works in the same way for the exit time of train r on block section b . These constraints are

$$y_{r,b}^{entr} \leq Mx_{r,b}, \quad \forall r \in R, b \in B_r \quad (4.2)$$

and

$$y_{r,b}^{exit} \leq Mx_{r,b}, \quad \forall r \in R, b \in B_r. \quad (4.3)$$

Running Time Constraints

For all the departure block sections, except the departure block section selected from the origin node of train r , constraint (4.4) ensures that train r is scheduled for the

4.5 Constraints

correct duration on block section b . That is, the exit time of train r on block section b must be equal to the sum of the entrance time of train r on block section b plus the sum of the running time of train r on every link in block section b . The constraint is

$$y_{r,b}^{exit} = y_{r,b}^{entr} + x_{r,b} \sum_{l \in L_b} t_{r,l}, \quad \forall r \in R, \quad b \in \{b \mid n_b^+ \neq n_r^o, \quad b \in B_r \setminus B^a\}. \quad (4.4)$$

For the departure block section that is selected as the departure block section of train r from its origin node, constraint (4.5) ensures that the actual dwell time of train r at its origin station is included in the time that train r is scheduled on its original departure block section. Furthermore, since the routing constraint (4.12) ensures that only one departure block section is selected for every train from its origin node, the sum formulation is used. The constraint is

$$\sum_{b \in B_{r,n_r^o}^+} y_{r,b}^{exit} = \sum_{b \in B_{r,n_r^o}^+} y_{r,b}^{entr} + \sum_{b \in B_{r,n_r^o}^+} x_{r,b} \sum_{l \in L_b} t_{r,l} + t_{r,s_r^o}^{stop}, \quad \forall r \in R. \quad (4.5)$$

Constraint (4.6) ensures that the running time of train r on the last link of an arrival block section is excluded from the duration that train r is scheduled on the block section, since trains start to release an arrival block section once they enter the last link of an arrival block section. The constraint is

$$y_{r,b}^{exit} = y_{r,b}^{entr} + x_{r,b} \sum_{l \in L_b \setminus l_b^{last}} t_{r,l}, \quad \forall r \in R, \quad b \in B_r \cap B^a. \quad (4.6)$$

Departure Time Window Constraints

Constraints (4.7) and (4.8) ensure that the scheduled departure times of trains at their origin node are larger than or equal to their earliest departure times and less than or equal to their latest departure times, as follows:

$$\sum_{b \in B_{r,n_r^o}^+} y_{r,b}^{entr} \geq t_r^s, \quad \forall r \in R \quad (4.7)$$

$$\sum_{b \in B_{r,n_r^o}^+} y_{r,b}^{entr} \leq t_r^e, \quad \forall r \in R. \quad (4.8)$$

4.5 Constraints

Exit Time and Entrance Time Transition Between Two Consecutive Block Sections

For the transition of trains from arrival block sections to departure block sections at every station, except the origin and destination stations, constraint (4.9) ensures that trains enter the departure block section only after the actual dwell time of trains on the station and the running time on the last link of the arrival block section have been completed. The constraint is

$$\sum_{n \in N_{r,s}} \sum_{b' \in B_{r,n}^{a,-}} y_{r,b'}^{exit} + t_{r,s}^{stop} + \sum_{n \in N_{r,s}} \sum_{b' \in B_{r,n}^{a,-}} x_{r,b'}(t_{r,b'}^{last}) = \sum_{n \in N_{r,s}} \sum_{b \in B_{r,n}^{+}} y_{r,b}^{entr}, \quad \forall r \in R, s \in S'_r. \quad (4.9)$$

Constraint (4.10) ensures that the entrance time of a train on an arrival block section is equal to the exit time of the train on the preceding departure block section, except at the origin and destination nodes of train r , where there are no block sections that flow into the origin node or out of the destination node that train r can use. The constraints is

$$\sum_{b' \in B_{r,n}^{-}, b' \notin B^a} y_{r,b'}^{exit} = \sum_{b \in B_{r,n}^{+}} y_{r,b}^{entr}, \quad \forall r \in R, n \in \{n \mid n \neq n_r^o, n \neq n_r^d, n \in N_r \setminus N_{r,s}, s \in S'_r\}. \quad (4.10)$$

Minimum Dwelling Time Constraints

Constraint (4.11) ensures that the actual dwell time of train r at station s is larger than or equal to the minimum dwell time specified for train r at station s . The constraint is

$$t_{r,s}^{stop} \geq t_{r,s}^{dwell}, \quad \forall r \in R, s \in S_r. \quad (4.11)$$

4.5.2 Train Routing Constraints

Constraint (4.12) ensures that a single route of connected block sections is selected for every train. That is, only one departure block section from the origin node, only one

4.5 Constraints

arrival block section into the destination node and only one block section into and out of every node that is selected along the route. The constraint is

$$\sum_{b \in B_{r,n}^+} x_{r,b} - \sum_{b \in B_{r,n}^-} x_{r,b} = \begin{cases} 1 & n = n_r^o, \\ -1 & n = n_r^d, \\ 0 & \text{otherwise.} \end{cases} \quad \forall r \in R \quad (4.12)$$

4.5.3 Block Section Occupancy Constraints

Conflicts Between Arrival Block Sections that Share the Same Last Link

Constraints (4.13) and (4.14) determine the sequence of train r and r' if the trains make use of arrival block sections that share the same last link. Furthermore, constraints (4.13) and (4.14) enforce that the entrance time of the second train is greater than or equal to the exit time of the first train plus the sum of the release time, reservation time, actual dwell time of the first train and the running time of the first train on the last link of the arrival block section.

Constraints (4.13) and (4.14) only limit the decision variables if both $x_{r',b'}$ and $x_{r,b}$ are equal to one. If $x_{r',b'}$, $x_{r,b}$ and $\mu_{r,b,r',b'}$ are equal to one, train r is scheduled on block section b before train r' is scheduled on block section b' . In the other case, if $x_{r',b'}$ and $x_{r,b}$ are equal to one and $\mu_{r,b,r',b'}$ is equal to zero, then train r is scheduled on block section b after train r' is scheduled on block section b' . The constraints are

$$M(1 - x_{r',b'}) + M(1 - x_{r,b}) + y_{r',b'}^{entr} - y_{r,b}^{exit} \geq t_{res} + t_{rel} + t_{r,s}^{stop} + t_{r,l_b^{last}} - M(1 - \mu_{r,b,r',b'}), \\ \forall r, r' \in R, n \in N_s, n' \in N_s, b \in B_{r,n}^{a,-}, b' \in B_{r',n'}^{a,-}, r \neq r', s = s^n, l_b^{last} = l_{b'}^{last}, s \in S_r \quad (4.13)$$

and

$$M(1 - x_{r',b'}) + M(1 - x_{r,b}) + y_{r,b}^{entr} - y_{r',b'}^{exit} \geq t_{res} + t_{rel} + t_{r',s}^{stop} + t_{r',l_{b'}^{last}} - M\mu_{r,b,r',b'}, \\ \forall r, r' \in R, n \in N_s, n' \in N_s, b \in B_{r,n}^{a,-}, b' \in B_{r',n'}^{a,-}, r \neq r', s = s^n, l_b^{last} = l_{b'}^{last}, s \in S_r. \quad (4.14)$$

Conflicts Between Other Types of Block Sections

If train r and r' are selected to run on block sections that conflict with one another, but the block sections are not arrival block sections that share the same last link, constraints

4.5 Constraints

(4.15) and (4.16) ensure that the entrance time of the second train is larger than or equal to the exit time of the first train plus the sum of the release and reservation times. Constraints (4.15) and (4.16) only limit the decision variables if $x_{r',b'}$, $x_{r,b}$ and $\varepsilon_{b,b'}$ are equal to one. If $x_{r',b'}$, $x_{r,b}$, $\varepsilon_{b,b'}$ and $\mu_{r,b,r',b'}$ are equal to one, train r is scheduled on block section b before train r' is scheduled on block section b' . In the other case, if $x_{r',b'}$, $x_{r,b}$ and $\varepsilon_{b,b'}$ are equal to one and $\mu_{r,b,r',b'}$ is equal to zero, then train r is scheduled on block section b after train r' is scheduled on block section b' . The constraints are

$$\begin{aligned} M(1 - x_{r',b'}) + M(1 - x_{r,b}) + y_{r',b'}^{entr} - y_{r,b}^{exit} &\geq t_{res} + t_{rel} - M(1 - \mu_{r,b,r',b'}), \\ \forall r, r' \in R, b \in B_r, b' \in B_{r'}, r \neq r', \varepsilon_{b,b'} &= 1 \end{aligned} \quad (4.15)$$

and

$$\begin{aligned} M(1 - x_{r',b'}) + M(1 - x_{r,b}) + y_{r,b}^{entr} - y_{r',b'}^{exit} &\geq t_{res} + t_{rel} - M\mu_{r,b,r',b'}, \\ \forall r, r' \in R, b \in B_r, b' \in B_{r'}, r \neq r', \varepsilon_{b,b'} &= 1. \end{aligned} \quad (4.16)$$

4.5.4 Maintenance Task Scheduling Constraints

Maintenance Task Time Constraints

Constraint (4.17) ensures that maintenance possession m is scheduled on or after its earliest start time. While, constraint (4.18) ensures that maintenance possession m is scheduled on or before its latest start time. Constraint (4.19) ensures that maintenance possession m is scheduled for at least the minimum duration. The constraints are

$$t_m^{start} \geq mot_m^s, \quad \forall m \in MOT, \quad (4.17)$$

$$t_m^{start} \leq mot_m^e, \quad \forall m \in MOT \text{ and} \quad (4.18)$$

$$t_m^{end} - t_m^{start} \geq d_m, \quad \forall m \in MOT. \quad (4.19)$$

4.5 Constraints

Maintenance Task Entrance Constraints

Constraints (4.20) and (4.21) sequence trains and maintenance possessions that make use of the same cells. In the case where $x_{r,b}$ is equal to one and $\alpha_{r,b,c}$ is equal to zero, constraint (4.20) ensures that maintenance possession m is only scheduled to start after the exit of train r on block section b plus the release time. In the case where both $x_{r,b}$ and $\alpha_{r,b,c}$ are equal to one, constraint (4.21) ensures that the entrance time of train r on block section b is greater than or equal to the end time of maintenance possession m plus the reservation time. The constraints are

$$y_{r,b}^{exit} + t_{rel} \leq t_m^{start} + M(1 - x_{r,b}) + M\alpha_{r,b,c}, \quad \forall r \in R, m \in MOT, c \in C_m, b \in B_r \cap B_c \quad (4.20)$$

and

$$M(1 - x_{r,b}) + y_{r,b}^{entr} \geq t_m^{end} + t_{res} - M(1 - \alpha_{r,b,c}), \quad \forall r \in R, m \in MOT, c \in C_m, b \in B_r \cap B_c. \quad (4.21)$$

4.5.5 The Domain of Variables

The domain of the variables is defined by constraints (4.22) to (4.27). The entrance and exit times of trains on block sections, start and end time of maintenance possessions and the actual dwell time of trains are defined as integer variables. The remainder of the variables are defined as binary variables. The constraints are

$$y_{r,b}^{entr}, y_{r,b}^{exit} \in N, \quad \forall r \in R, b \in B_r, \quad (4.22)$$

$$x_{r,b} \in \{0, 1\}, \quad \forall r \in R, b \in B_r, \quad (4.23)$$

$$\mu_{r,b,r',b'} \in \{0, 1\}, \quad \forall r, r' \in R, b \in B_r, b' \in B_{r'}, r \neq r', \quad (4.24)$$

$$t_{r,s}^{stop} \in N, \quad \forall r \in R, s \in S_r, \quad (4.25)$$

4.6 Validation

$$t_m^{start}, t_m^{end} \in N, \quad \forall m \in MOT \text{ and} \quad (4.26)$$

$$\alpha_{r,b,c} \in \{0, 1\}, \quad \forall r \in R, \quad m \in MOT, \quad c \in C_m, \quad b \in (B_r \cap B_c). \quad (4.27)$$

4.6 Validation

The model validation consisted of the following steps:

1. Develop and implement the model on a small test case study. The test case study consisted of a single mainline with four stations where trains can cross and a mixed train service of seven trains scheduled over a five hour period.
2. Solve the model for different scenarios.
3. Observe and inspect the solution of every scenario for unexpected behaviour.
4. Evaluate the correct functioning of each constraint by comparing the expected output values with the output produced by the model.

It was found that the model conformed to the desired output requirements.

Chapter 5

Experiments and Results

A set of experiments was formulated to test the capability of the model. This chapter describes these experiments and their results.

5.1 Experiments

A 12 month sample of maintenance possession data was taken for the railway segment between Bellville and Wellington to develop realistic instances for the experiments. Only major possessions, that remove tracks from service, were considered. From the sample, the following observations were made:

- The majority of possessions are taken on a single line, have a duration of five hours, start at 09:00 and finish at 14:00.
- Almost all of the possessions are on mainline track segments, with very few on station loop lines.
- Between zero and three possessions are scheduled per day.
- Most of the possessions are scheduled during the day time working hours. Although, some of the possessions are scheduled during the night, *i.e.* between 18:00 and 06:00.

The set of experiments is listed in Table 5.1. The following settings are common to all the experiments:

5.1 Experiments

- All experiments were done with the Friday train schedule, which is the busiest day of the week.
- The possession durations were set to five hours.
- The reservation and release times were set to five minutes each to enforce a 10 minute headway between trains on conflicting block sections.
- “Wrong road working” described in Subsection 3.1 was enabled on the mainline track next to each possession.
- Except for Experiment 1, all experiments use the actual probabilities of the trains.

Firstly, two baseline experiments were done without maintenance. In the first baseline experiment, the probability of all trains was set to 100 percent. In the second, the actual probabilities of the trains were used.

Experiments were set up for scenarios with one, two and three single track mainline possessions, as well as a scenario with a double track mainline possession. Firstly, the model was allowed to select any starting time so that the possessions are completed before the 24 hour scheduling period is finished. Secondly, similar scenarios were run where the start times of the possessions were limited to between 07:00 and 12:00 so that day time working hours are enforced for the maintenance teams.

5.1 Experiments

Table 5.1: List of experiments

Experiment	Description
1	No maintenance, 100 percent probability for all trains.
2	No maintenance, actual probabilities for all trains.
3	One mainline possession on the down line, cell 36, between Kraaifontein and Muldersvlei.
4	One mainline possession on the down line, cell 36, between Kraaifontein and Muldersvlei. Possession start time limited to between 07:00 and 12:00.
5	Two mainline possessions on the down line: the first possession is on cell 36 between Kraaifontein and Muldersvlei and the second possession is on cell 46 between Muldersvlei and Klapmuts.
6	Two mainline possessions on the down line: the first possession is on cell 36 between Kraaifontein and Muldersvlei and the second possession is on cell 46 between Muldersvlei and Klapmuts. Possession start times limited to between 07:00 and 12:00.
7	Three mainline possessions: the first two possessions are the same as described in Experiment 5. The third possession is on cell 80 between Paarl and Huguenot on the up line.
8	Three mainline possessions: the first two possessions are the same as described in Experiment 5. The third possession is on cell 80 between Paarl and Huguenot on the up line. Possession start times limited to between 07:00 and 12:00.
9	One double track mainline possession on cells 22 and 33 between Brackenfell and Kraaifontein. Solving time limited to 30 minutes.
10	One double track mainline possession on cells 22 and 33 between Brackenfell and Kraaifontein. Solving time limited to 30 minutes. Possession start time limited to between 07:00 and 12:00.

5.2 Results

The results of the experiments are presented in this subsection. The solutions are presented as train diagrams. The trains are numbered and the maintenance possessions are named m1, m2 and m3. The trains are represented by the diagonal lines, while the possessions are represented by the rectangles.

The result of each experiment is:

1. The value of the objective function is 1.93 minutes, the solving time was 3 minutes and 13 seconds and the solution is optimal. The solution is presented in [Figure 5.1](#).
2. The value of the objective function is 1.88 minutes, the solving time was 3 minutes and 9 seconds and the solution is optimal. The solution is presented in [Figure 5.2](#).
3. The value of the objective function is 1.88 minutes, the solving time was 3 minutes and 36 seconds and the solution is optimal. The solution is presented in [Figure 5.3](#).
4. The value of the objective function is 1.88 minutes, the solving time was 5 minutes and 51 seconds and the solution is optimal. The solution is presented in [Figure 5.4](#).
5. The value of the objective function is 1.88 minutes, the solving time was 6 minutes and 58 seconds and the solution is optimal. The solution is presented in [Figure 5.5](#).
6. The value of the objective function is 1.88 minutes, the solving time was 4 minutes and 9 seconds and the solution is optimal. The solution is presented in [Figure 5.6](#).
7. The value of the objective function is 1.88 minutes, the solving time was 3 minutes and 29 seconds and the solution is optimal. The solution is presented in [Figure 5.7](#).

5.3 Evaluation

8. The value of the objective function is 1.88 minutes, the solving time was 8 minutes and 26 seconds and the solution is optimal. The solution is presented in Figure [5.8](#).
9. The value of the objective function is 58.17 minutes and the solving time was 28 minutes and 46 seconds. The solution is presented in Figure [5.9](#).
10. The value of the objective function is 210.66 minutes and the solving time was 11 minutes and 44 seconds. The solution is presented in Figure [5.10](#).

5.3 Evaluation

The experiments have shown that the model is capable of producing solutions to typical possession scheduling instances within an acceptable time. Instances where one, two or three single line possessions were scheduled on the mainline tracks were solved to optimality within the longest run time of 8 minutes and 36 seconds. In the instances where a double line possession was scheduled on the mainlines, the model was able to produce good solutions within a 30 minute solving time limit. These results indicate that the model will be useful as a decision support tool since it can produce optimal and good solutions at short notice.

5.3 Evaluation

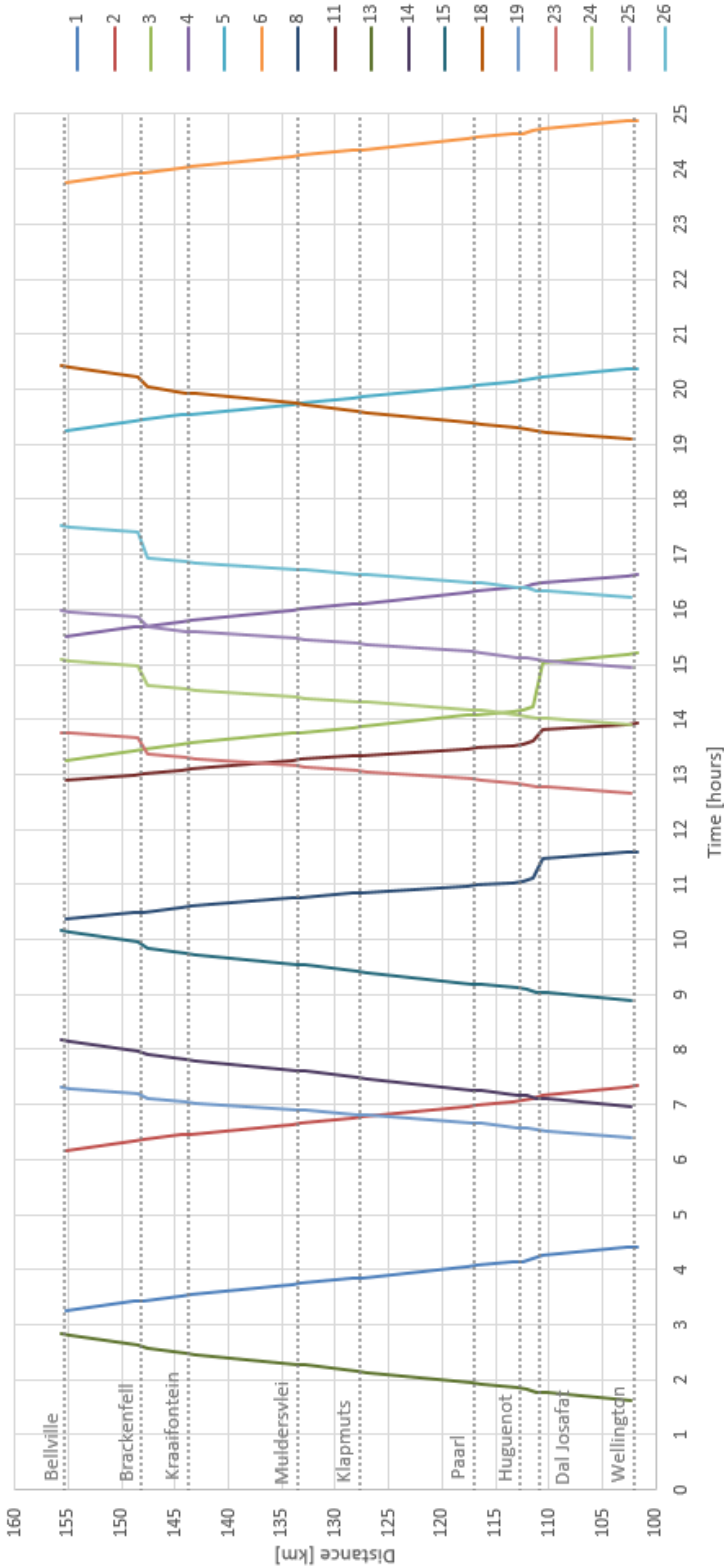


Figure 5.1: The solution of Experiment 1

5.3 Evaluation

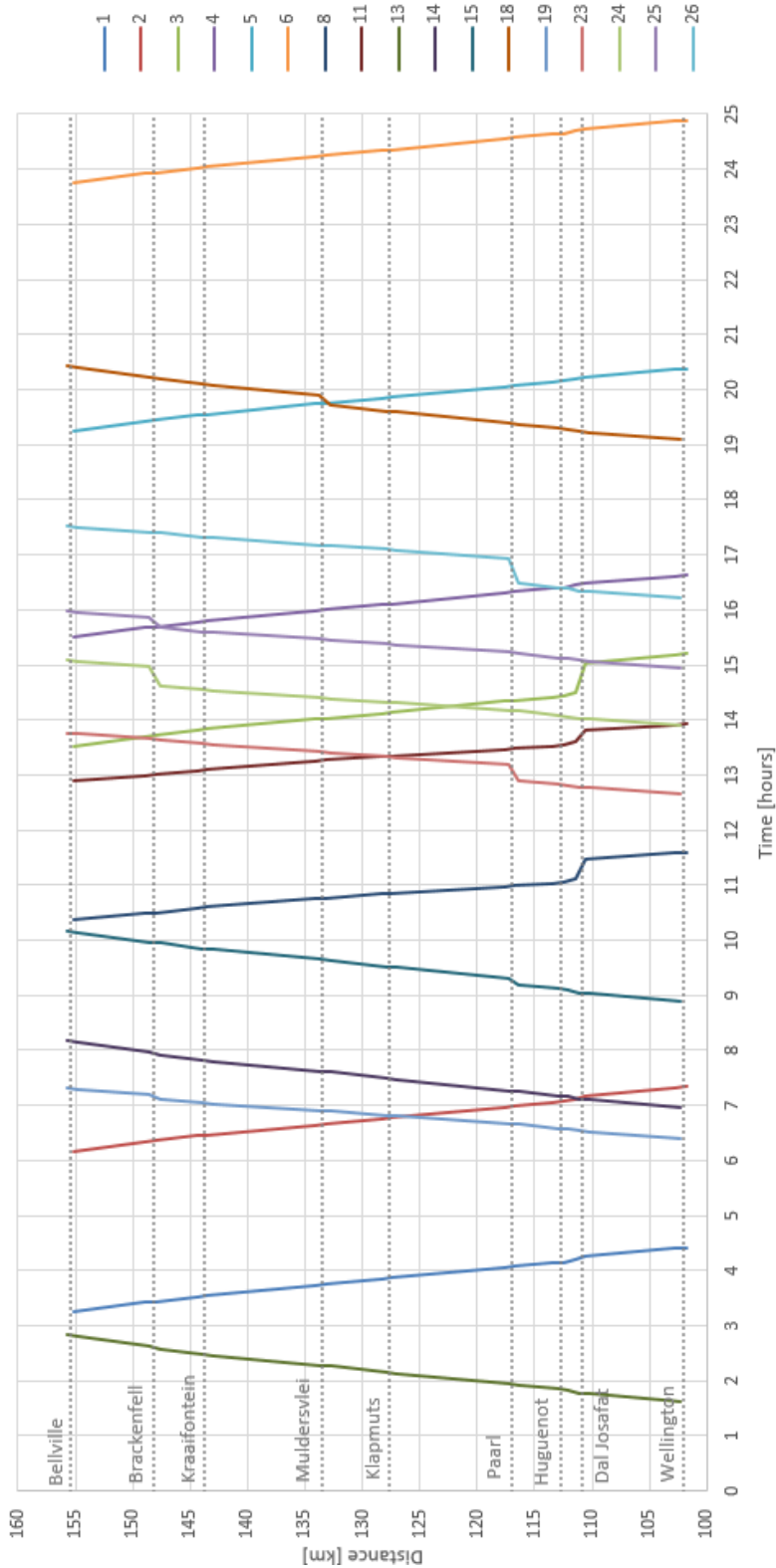


Figure 5.2: The solution of Experiment 2

5.3 Evaluation

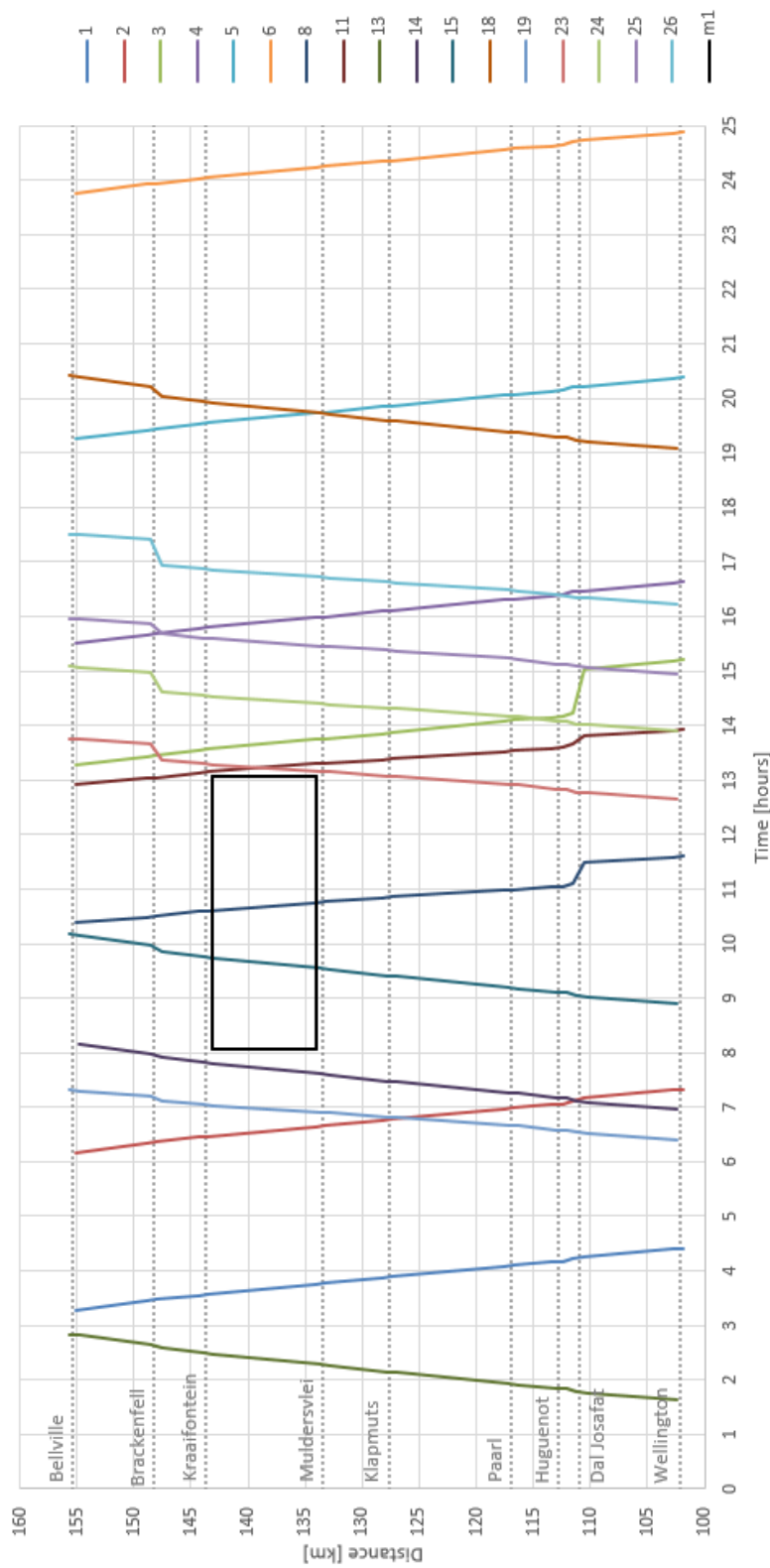


Figure 5.3: The solution of Experiment 3

5.3 Evaluation

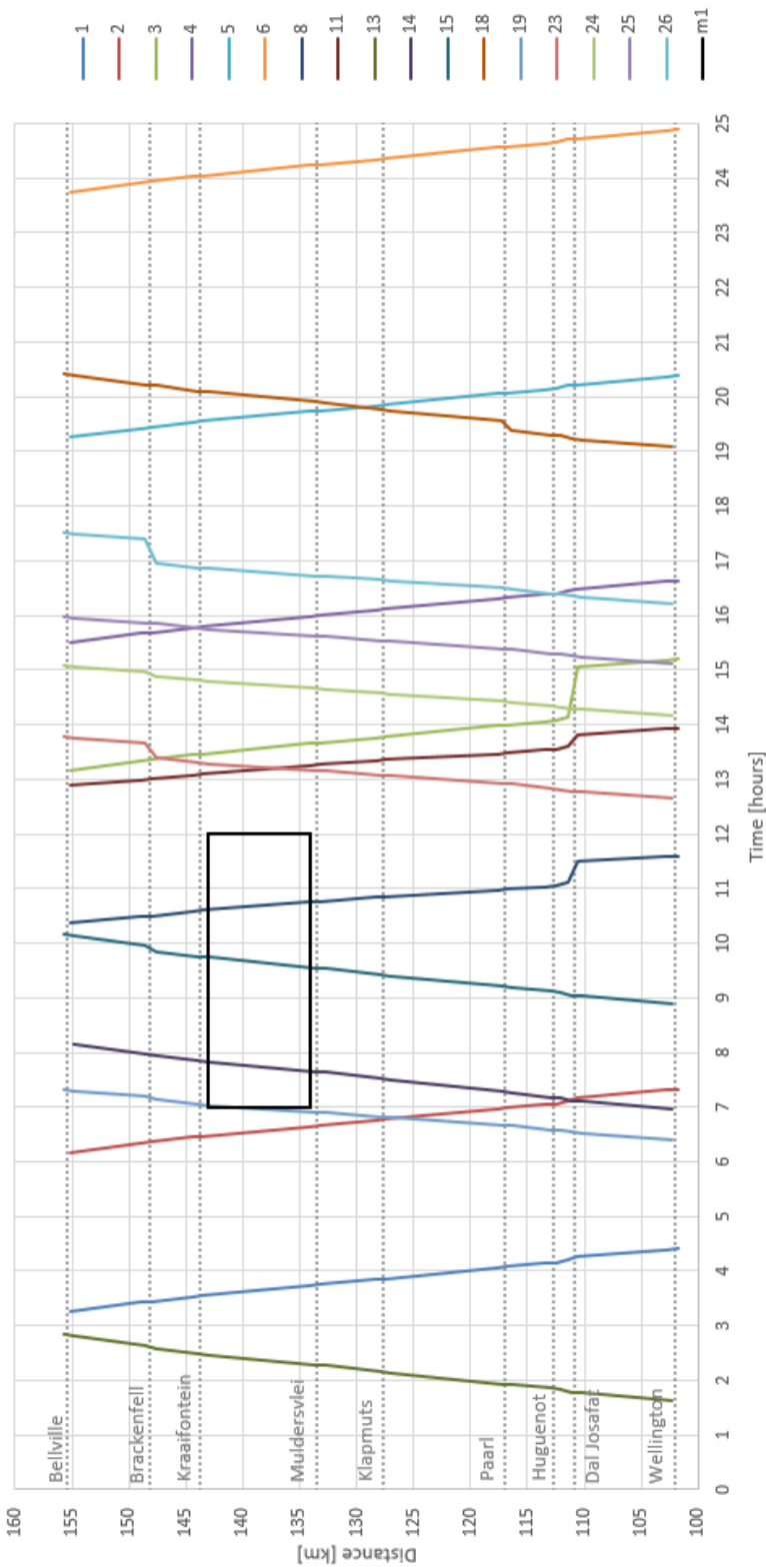


Figure 5.4: The solution of Experiment 4

5.3 Evaluation

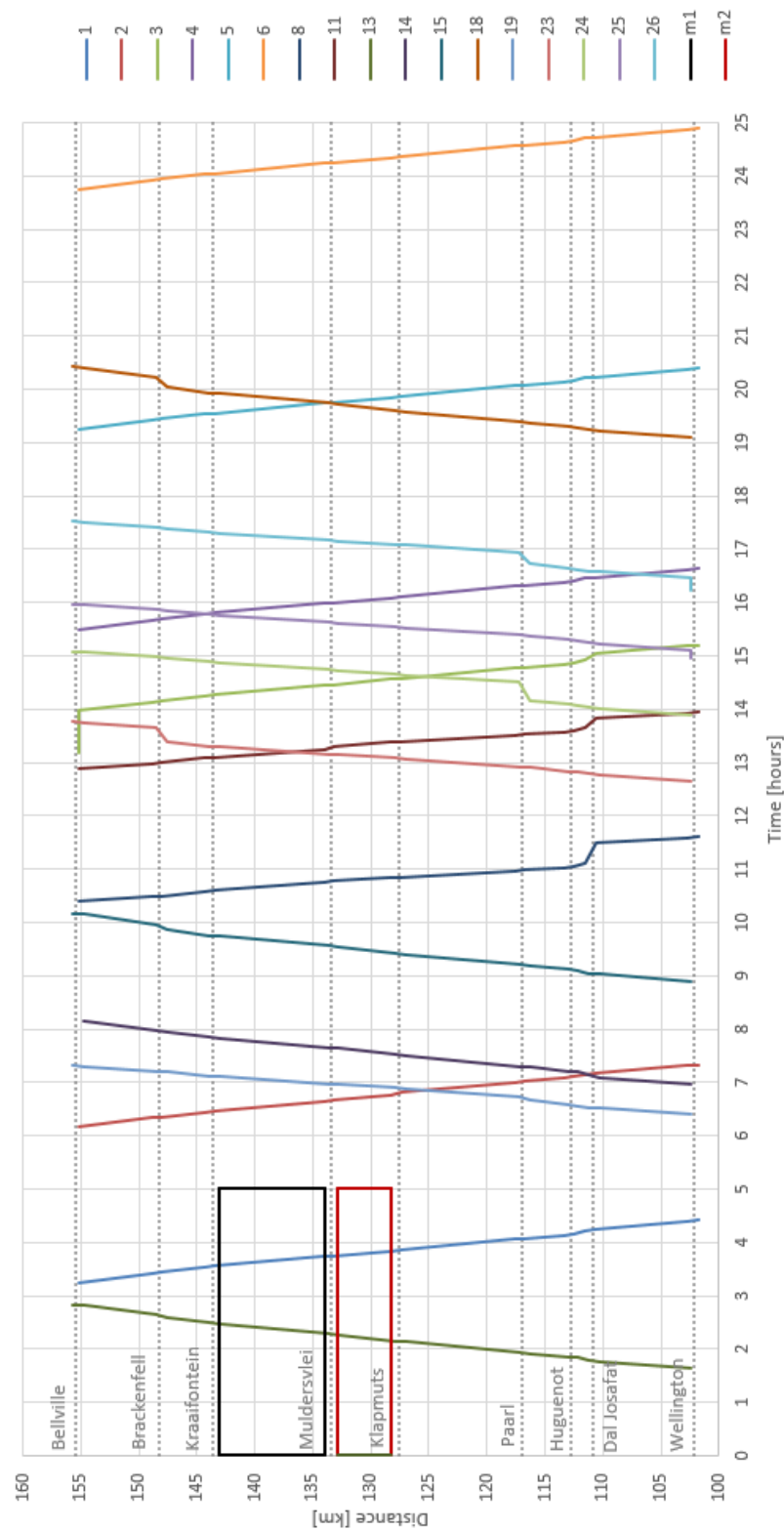


Figure 5.5: The solution of Experiment 5

5.3 Evaluation

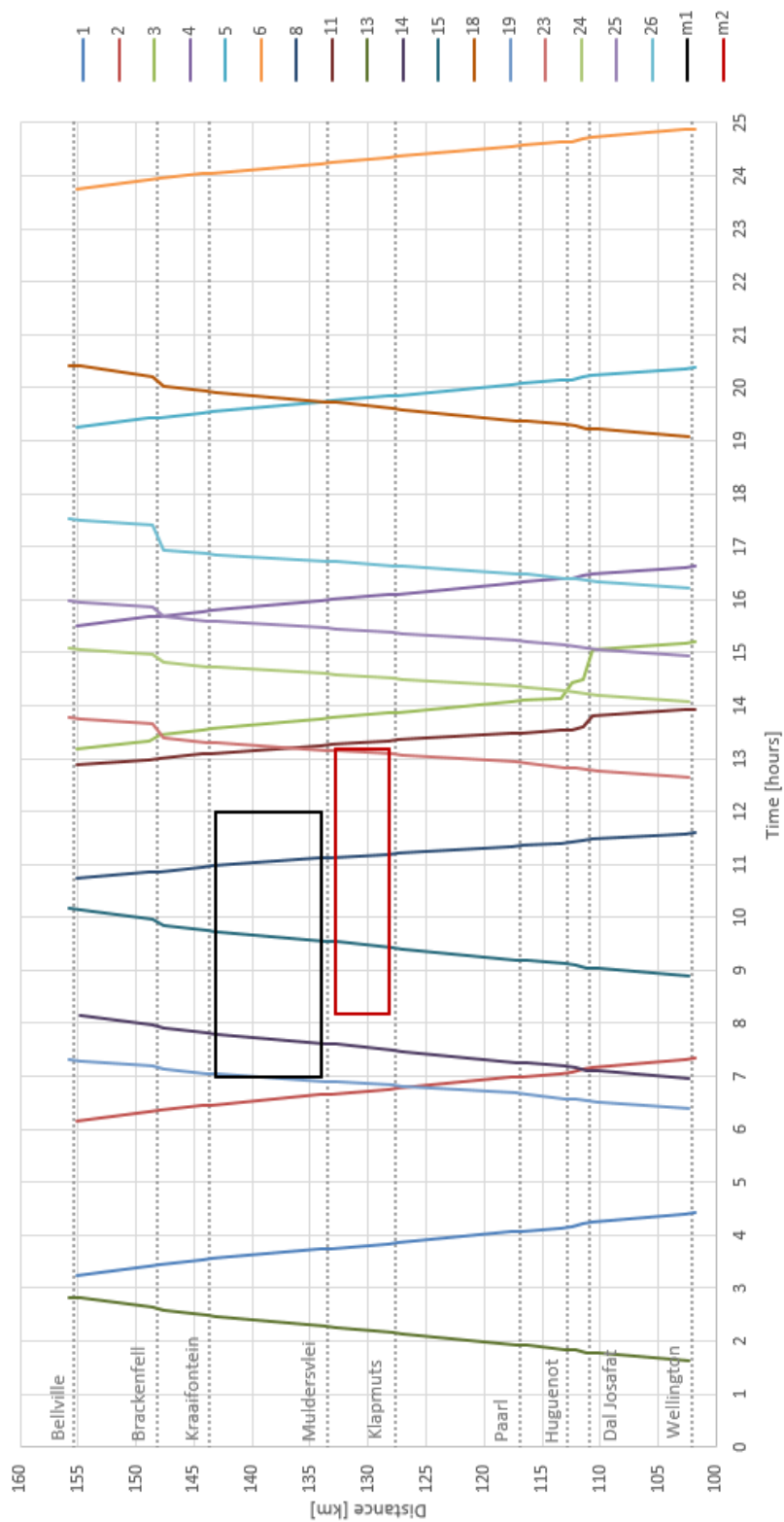


Figure 5.6: The solution of Experiment 6

5.3 Evaluation

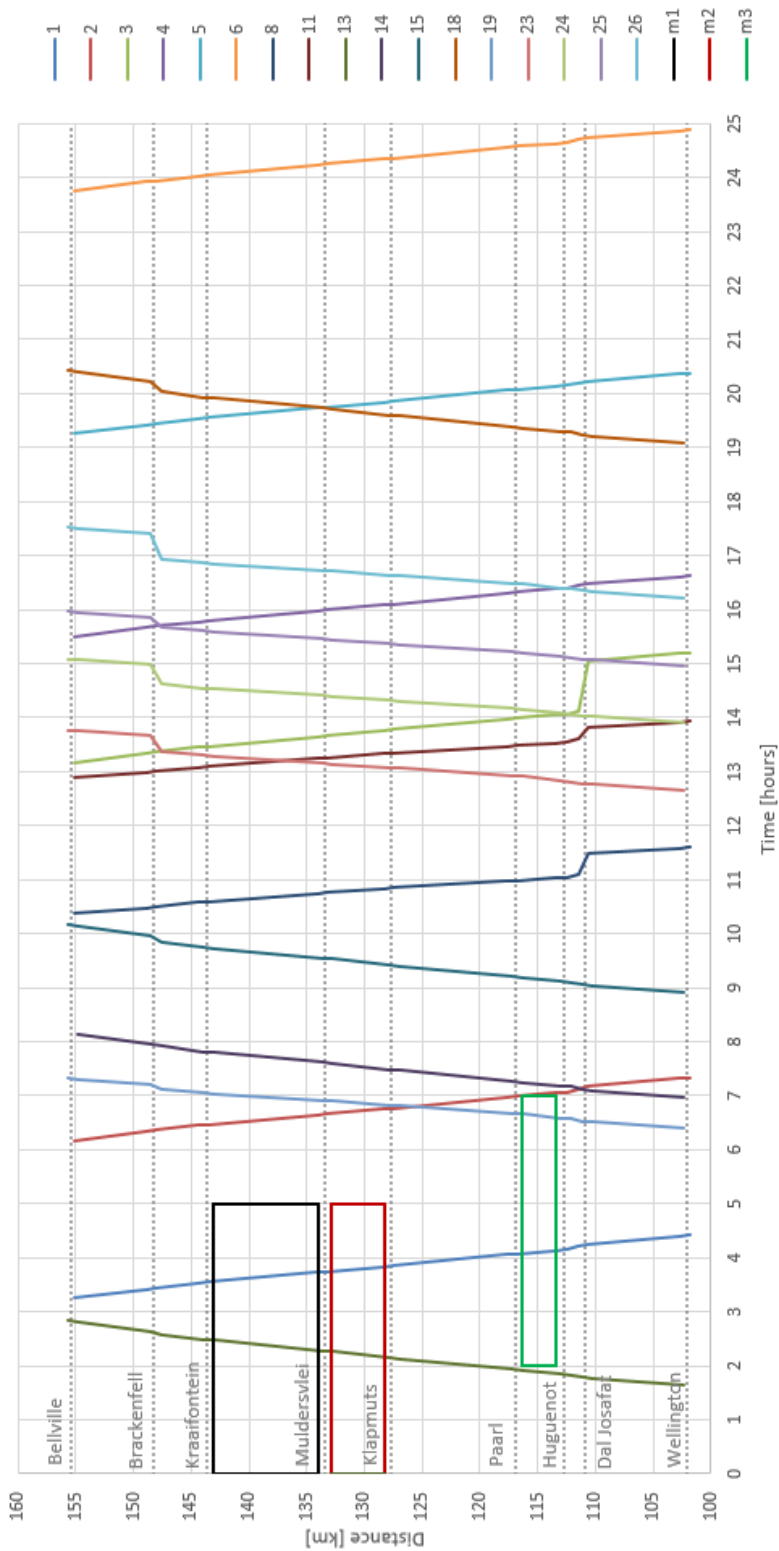


Figure 5.7: The solution of Experiment 7

5.3 Evaluation

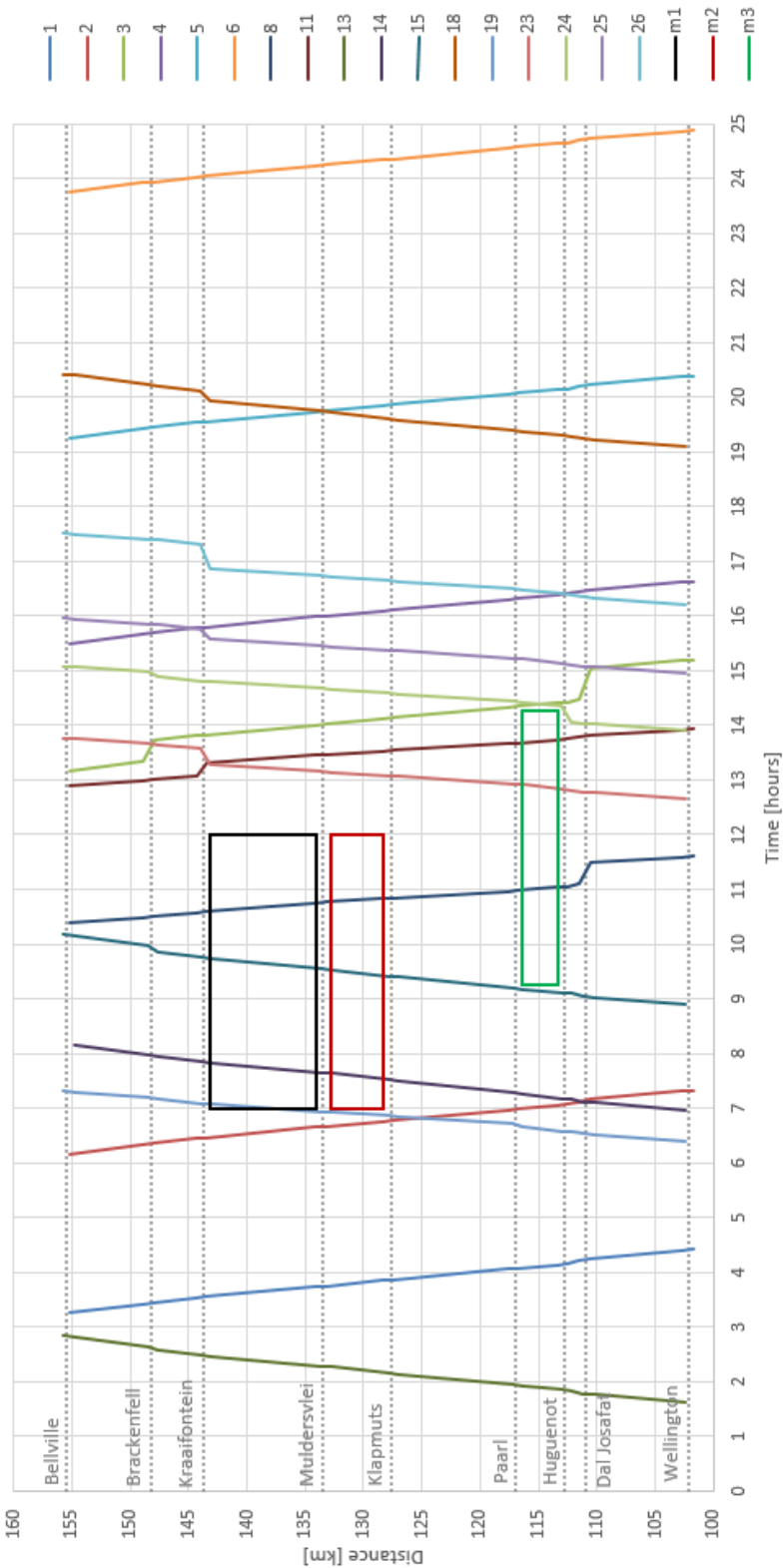


Figure 5.8: The solution of Experiment 8

5.3 Evaluation

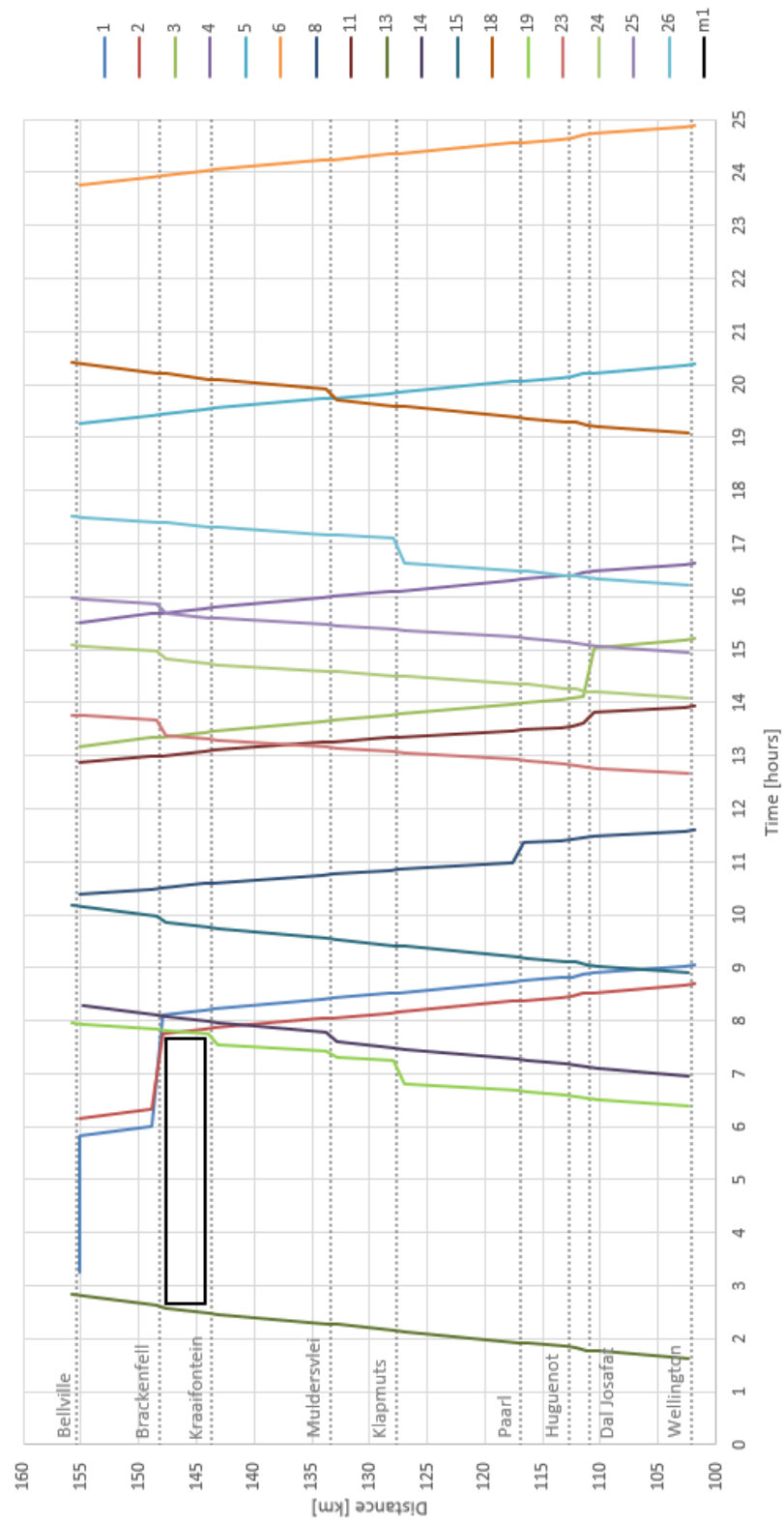


Figure 5.9: The solution of Experiment 9

5.3 Evaluation

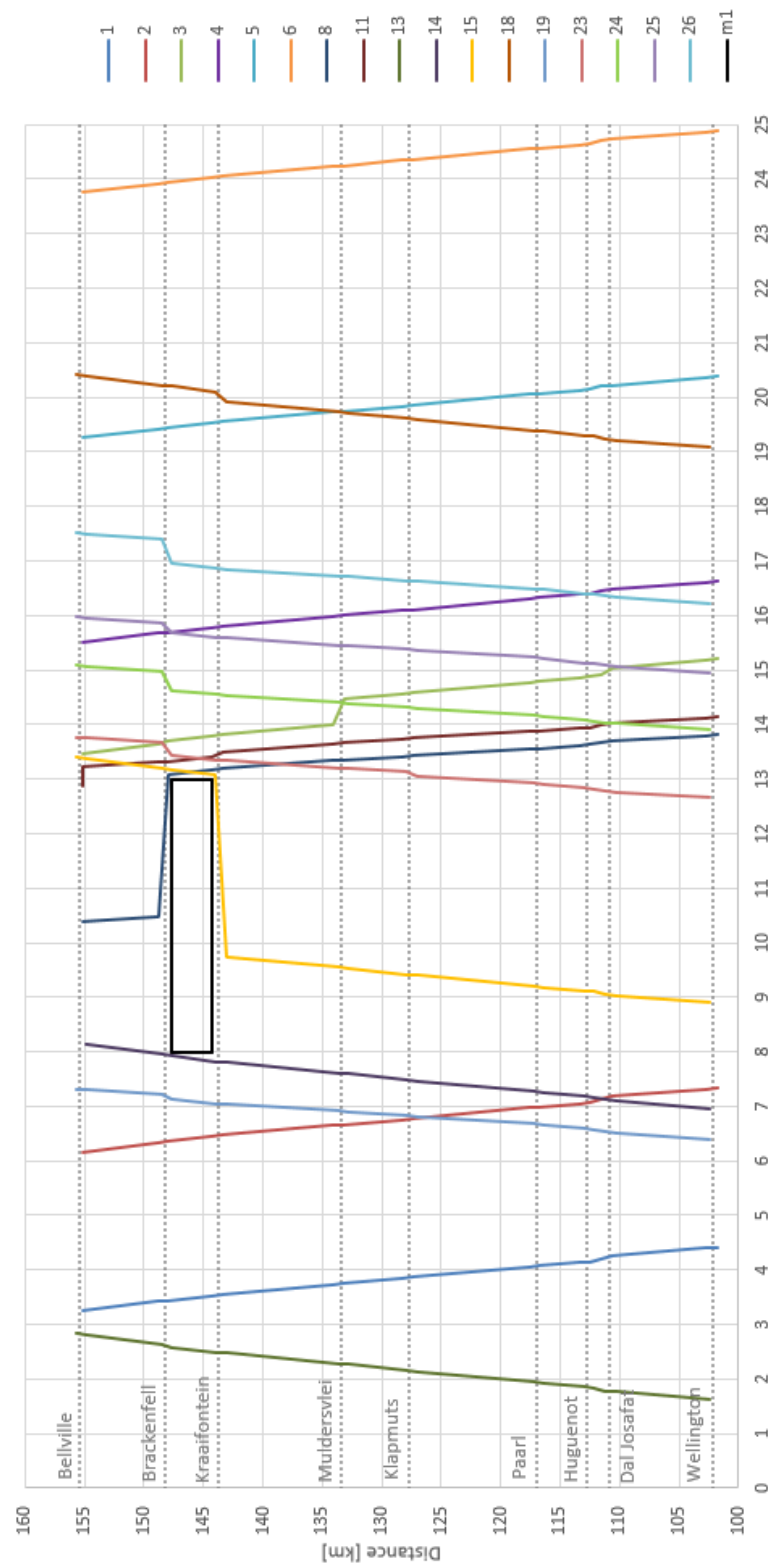


Figure 5.10: The solution of Experiment 10

5.4 Transnet Feedback

This section presents feedback from senior technical experts of train scheduling, rail network maintenance and capital projects within Transnet Freight Rail (TFR) regarding the usability of the maintenance possession scheduler (MPS). The feedback is summarised in Figure 5.11.

The MPS was well accepted by the TFR representatives. The experts stated that it can work in practice. Their feedback is presented next.

- Firstly, the experts validated the model. They stated that the MPS does realistic scheduling of single and double track maintenance possessions, as well as trains. Also, importantly, the train headways are correct.
- Next, the experts provided feedback on the usability of the MPS. They said that the train diagram helps greatly since it shows the impact of possessions and how trains are adjusted. They continued by saying that the train diagram provides insight to the decision makers and helps train planners to choose better departure times for trains. The experts also noted several use cases for the MPS. Namely, it can be used for train and possession scheduling at various timeframes. Firstly, it can be used during the possession planning process that occurs a few weeks before execution of the train service. Secondly, it can be used during short term train scheduling a few days before the execution of the train service. Finally, it can be used to reschedule the train service after an operational disruption occurs a few hours before the execution of the train service.
- When the experts considered whether the possession planning process would be better with the MPS, they noted the following. In the current process, there is no tool that shows the impact of possessions. Furthermore, trains are mostly cancelled to accommodate possessions. Otherwise, trains are delayed by hours and departed after the possession in another train slot. Also, the manual rescheduling of the train service to accommodate possessions takes three to four hours. The experts then noted that the current process would be better with the MPS since the impact of possessions can be seen on the train diagram, the MPS assists decision making, the MPS is faster than the manual process and there are benefits expected from using the MPS tool.

5.4 Transnet Feedback

Validated	Use cases	Expected benefits
<ul style="list-style-type: none"> Realistic scheduling of single and double line possessions and trains Headways correct 	<ul style="list-style-type: none"> Possession and train scheduling at various time frames before execution: <ul style="list-style-type: none"> Possession scheduling (weeks) Train scheduling (days) Train rescheduling after disruption (hours) 	<ul style="list-style-type: none"> Better scheduling of crews – cost savings, reduced security risks Better scheduling of trains – reduced security risk Higher possession approval rate – Standard way of determining the train schedule impact – better motivations Better decisions – impact can be visualised System optimization and increasing throughput Minimise human errors, subjectivity, emotions, inexperience Optimal time slots for possessions
Train diagram	Better than current process	
<ul style="list-style-type: none"> Helps greatly Shows impact of occupations Shows train adjustments Provides insight Helps to choose better departure times 	<ul style="list-style-type: none"> Current process <ul style="list-style-type: none"> can't see impact on the train plan trains mostly cancelled to accommodate possessions, otherwise delayed by hours and departed after the occupation in another train slot Manual scheduling takes 3-4 hours 	<ul style="list-style-type: none"> With the MPS tool <ul style="list-style-type: none"> Impact of occupations can be seen on train diagram. Assists decision making Faster than the manual process Refer to expected benefits
Likes	Shortcomings	Recommended for further development, testing and implementation
<ul style="list-style-type: none"> Visual solutions – train diagram Scheduling flexibility for occupations and trains Applicable in different use cases Practical application that reflects the current operations and train movements Expected benefits 	<ul style="list-style-type: none"> Occupations Between Trains also have impact on the train service: it takes time to make sure that the line is safe before and after maintenance crews occupy a track in between trains Real world train running times can vary greatly as a result of incidents such as theft, vandalism, natural disasters and protests 	<ul style="list-style-type: none"> Definitely – Currently, no tools show impact of occupations

Figure 5.11: Feedback from senior technical experts of train scheduling, rail network maintenance and capital projects

5.4 Transnet Feedback

- The following benefits are expected by the experts. Firstly, better scheduling of train crews: delaying trains at major hubs can reduce crewing costs, since crews sign-on later, as well as reduce security risks, since crews are waiting at safer stations. Secondly, better scheduling of trains: delaying trains at major hubs reduces the risks of theft and vandalism of train equipment. Thirdly, a higher possession request approval rate: the maintenance department can motivate better for time slots if there is a standard way of determining the impact on the train schedule. The TFR representatives then continued by stating further expected benefits of the MPS such as: better scheduling decisions can be made because the impact of possessions can be visualised, it can assist with system optimisation and increasing throughput, it can reduce human errors, subjectivity, emotions and inexperience during scheduling and it can select optimal time slots for possessions.
- The experts liked that the solutions are visual, that the MPS provides scheduling flexibility for both possessions and trains, that it is applicable in various use cases and that it is applied on the TFR system and reflects the current operations and train movements.
- Some shortcomings were also noted by the experts. The shortcomings were related to the variability that is inherent in rail operations that is not reflected in the model. For example, minor possessions may also have an impact on the train service since it takes time to declare the track safe before and after maintenance crews occupy the track. Furthermore, real world train running times can vary greatly as a result of incidents such as theft, vandalism, natural disasters and protests.
- Finally, the experts stated they will definitely recommend Transnet to further develop, test and implement the MPS since there are currently no tools that show the impact of possessions on the train schedule.

This concludes the discussion of the expert feedback. The next chapter will present the project summary, appraisal and recommended future works.

Chapter 6

Conclusions

This chapter summarises the work done, presents a project appraisal and recommends future work.

6.1 Project Summary

Maintenance of rail infrastructure is an important element in rail operations in order to keep traffic moving. However, maintenance causes infrastructure to be taken out of service, which impacts traffic flow. In this study, the requirements of a maintenance possession scheduler for a South African application was investigated, and a proposed solution was subsequently developed. The main objective of the scheduler was to minimise the deviation of the train service on a subset of rail infrastructure while ensuring that the required maintenance is done.

A literature study, presented in Chapter 2, was done on a number of themes. It includes an overview of the local railway operator at which the research was done, with a look at the role of industrial engineering as a function in the railway operator business. Themes including railway infrastructure and operations, planning of railway operations, and maintenance in the context of rail operations were also discussed to enable the researcher to understand the complete picture. The topic of possession scheduling, which was critical for this research, was then studied and presented. Past and recent works were studied and research areas and trends were synthesised during this process, including time span of possession scheduling in optimisation models, and whether it was done on microscopic, mesoscopic or macroscopic level. The various

6.2 Project Appraisal and Future Works

optimisation objectives formulated by researchers were also noted, among other sub themes.

In Chapter 3, an application case was identified to apply a possession scheduling model to. The case was identified as the railway infrastructure between Bellville and Wellington in the Western Cape province of South Africa. The nature of this case was described and all characteristics were noted with the future possession scheduling model in mind.

A novel mixed-integer linear programming model was formulated for this case, which was presented and discussed in Chapter 4. The model formulation was implemented in Cplex and validated using specific requirement statements. Also presented were the objective function and the many constraints, which are the heart of the implemented model.

Demonstrative experiments were executed and discussed in Chapter 5. It was shown that the model can do possession scheduling for 24 operational hours on a microscopic level. The experiments proved that the implemented model performed well and it delivered optimal results in less than eight minutes, which makes it a feasible maintenance possession scheduler for day-to-day work in the immediate planning horizon.

The study showed that an operations research technique like mixed-integer linear programming, which is learned as a core skill by industrial engineers, can be applied in South African railway possession scheduling.

6.2 Project Appraisal and Future Works

The research objectives set out in Section 1.5 have been achieved as described in the project summary. The following future work is proposed:

- Apply the model to larger real world instances.
- Develop and implement solution techniques that can solve the large instances.
- Develop and solve the model with two objectives, for example, minimising the overtime cost of maintenance teams and the train service deviations.
- Test and implement the model with subject matter experts in industry.

This chapter summarised the work done, provided an appraisal of the project and recommended future work.

References

- ALBRECHT, A.R., PANTON, D.M. & LEE, D.H. (2013). Rescheduling rail networks with maintenance disruptions using problem space search. *Computers & Operations Research*, **40**, 703–712. [121](#), [123](#), [124](#), [126](#), [128](#), [129](#)
- ARENAS, D., PELLEGRINI, P., HANAFI, S. & RODRIGUEZ, J. (2018). Timetable rearrangement to cope with railway maintenance activities. *Computers & Operations Research*, **95**, 123–138. [121](#), [123](#), [124](#), [125](#), [126](#), [127](#), [128](#), [129](#)
- BABABEIK, M., ZERGUINI, S., FARJAD-AMIN, M., KHADEMI, N. & BAGHERI, M. (2019). Developing a train timetable according to track maintenance plans: A stochastic optimization of buffer time schedules. *Transportation research procedia*, **37**, 27–34. [121](#), [123](#), [126](#), [128](#), [129](#)
- BAHRAMIAN, Z. & BAGHERI, M. (2015). A simulation-based optimization approach for passenger train timetabling with periodic track maintenance and stops for praying. *Journal of Modern Transportation*, **23**, 148–157. [121](#), [127](#), [128](#)
- BEŠINOVIĆ, N., WIDARNO, B. & GOVERDE, R.M. (2020). Adjusting freight train paths to infrastructure possessions. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1–6, IEEE. [121](#), [123](#), [126](#), [127](#), [128](#), [129](#)
- BOLAND, N., KALINOWSKI, T., WATERER, H. & ZHENG, L. (2013). Mixed integer programming based maintenance scheduling for the hunter valley coal chain. *Journal of Scheduling*, **16**, 649–659. [121](#), [124](#), [126](#), [128](#)
- BOLAND, N., KALINOWSKI, T., WATERER, H. & ZHENG, L. (2014). Scheduling arc maintenance jobs in a network to maximize total flow over time. *Discrete Applied Mathematics*, **163**, 34–52. [121](#)

REFERENCES

-
- BOTHA, D., NUTTALL, M. & MATSEKE, A. (2016). Rail Bound Telecommunications, Lecture Notes of Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [55](#)
- BRUCKER, P., HEITMANN, S. & KNUST, S. (2002). Scheduling railway traffic at a construction site. *OR Spectrum*, **24**, 19–30. [121](#), [124](#), [127](#), [128](#)
- BUDAI, G., HUISMAN, D. & DEKKER, R. (2006). Scheduling preventive railway maintenance activities. *The Journal of the Operational Research Society*, **57**, 1035–1044. [121](#), [124](#), [126](#), [128](#)
- BUDAI-BALKE, G. (2009). *Operations research models for scheduling railway infrastructure maintenance*. 456, Rozenberg Publishers. [121](#), [126](#), [128](#)
- BUENO, P., VILELA, P.R., CHRISTOFOLETTI, L.M. & VIEIRA, A.P. (2019). Optimizing railway track maintenance scheduling to minimize circulation impacts. In *2019 Joint Rail Conference*, American Society of Mechanical Engineers Digital Collection. [121](#), [123](#), [126](#), [128](#), [129](#)
- BURDETT, R. & KOZAN, E. (2009). Techniques for inserting additional trains into existing timetables. *Transportation Research Part B: Methodological*, **43**, 821–836. [121](#), [127](#), [128](#)
- CAPRARA, A., MONACI, M., TOTH, P. & GUIDA, P.L. (2006). A lagrangian heuristic algorithm for a real-world train timetabling problem. *Discrete Applied Mathematics*, **154**, 738–753. [121](#), [124](#), [128](#)
- CAPRARA, A., KROON, L., MONACI, M., PEETERS, M. & TOTH, P. (2007). Passenger railway optimization. *Handbooks in operations research and management science*, **14**, 129–187. [121](#)
- CHEUNG, B.S., CHOW, K., HUI, L.C. & YONG, A.M. (1999). Railway track possession assignment using constraint satisfaction. *Engineering applications of artificial intelligence*, **12**, 599–611. [121](#), [127](#)
- CORREIA, C., FLYNN, D., ULIANA, E. & WORMALD, M. (2013). *Financial Management*. Juta, 7th edn. [73](#)

REFERENCES

-
- DAO, C., BASTEN, R. & HARTMANN, A. (2018). Maintenance scheduling for railway tracks under limited possession time. *Journal of Transportation Engineering, Part A: Systems*, **144**. [121](#)
- DAO, C.D., BASTEN, R.J. & HARTMANN, A. (2017). Optimal maintenance clustering for rail track components subject to possession capacity constraints. *International Journal of Architectural and Environmental Engineering*, **11**, 752–756. [121](#)
- DAO, C.D., HARTMANN, A., LAMPER, A. & HERBERT, P. (2019). Scheduling infrastructure renewal for railway networks. *Journal of Infrastructure Systems*, **25**. [121](#)
- D’ARIANO, A., MENG, L., CENTULIO, G. & CORMAN, F. (2019). Integrated stochastic optimization approaches for tactical scheduling of trains and railway infrastructure maintenance. *Computers & Industrial Engineering*, **127**, 1315–1335. [121](#), [123](#), [124](#), [125](#), [126](#), [127](#), [129](#)
- DARWISH, H. & VAN DYK, L. (2016). The industrial engineering identity: from historic skills to modern values, duties, and roles. *South African Journal of Industrial Engineering*, **27**, 50–63. [17](#), [18](#), [19](#), [20](#), [25](#)
- DEN HERTOOG, D., VAN ZANTE-DE FOKKERT, J., SJAMAAR, S. & BEUSMANS, R. (2005). Optimal working zone division for safe track maintenance in the netherlands. *Accident Analysis and Prevention*, **37**, 890–893. [121](#)
- DoT WEBSITE (2015). National Rail Policy Draft Green Paper. Accessed: 2018-08-18. [10](#), [11](#), [12](#), [13](#), [14](#), [15](#)
- DoT WEBSITE (2017). National Rail Policy Draft White Paper. Accessed: 2018-08-18. [8](#), [9](#), [13](#)
- DPE WEBSITE (2018). DPE Overview. Accessed: 2018-08-19. [13](#)
- DU PLESSIS, J. (2017a). Overhead Track Equipment Systems, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [34](#), [36](#), [38](#), [39](#)
- DU PLESSIS, J. (2017b). Traction Power Supply, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [34](#), [38](#), [40](#)

REFERENCES

- DUVEL, J. (2017). Track Maintenance Management, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [3](#), [77](#), [78](#), [97](#), [98](#), [99](#), [100](#), [104](#), [105](#), [106](#), [108](#), [109](#), [110](#), [111](#)
- EBERSÖHN, W. & GRÄBE, P. (2008a). Rail Transport, Lecture Notes of Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [1](#), [2](#)
- EBERSÖHN, W. & GRÄBE, P. (2008b). Track Terminology, Performance and Components, Lecture Notes of Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [26](#), [28](#), [29](#), [30](#), [31](#), [32](#), [33](#), [35](#)
- FAMUREWA, S.M., NISSEN, A. & KUMAR, U. (2016). Scheduling of railway infrastructure maintenance tasks using train free windows. In *Current Trends in Reliability, Availability, Maintainability and Safety*, 425–436, Springer. [121](#), [126](#), [127](#)
- FORSGREN, M., ARONSSON, M. & GESTRELIUS, S. (2013). Maintaining tracks and traffic flow at the same time. *Journal of Rail Transport Planning & Management*, **3**, 111–123. [121](#), [123](#), [124](#), [126](#), [129](#)
- GRÄBE, P. (2017). Transport in Context, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [2](#), [26](#)
- HIGGINS, A. (1998). Scheduling of railway track maintenance activities and crews. *The Journal of the Operational Research Society*, **49**, 1026–1033. [121](#), [124](#), [125](#), [126](#), [127](#)
- HIGGINS, A., FERREIRA, L. & LAKE, M. (1999). Scheduling rail track maintenance to minimise overall delays. In *The 14th International Symposium on Transportation and Traffic Theory*. [121](#), [125](#), [126](#), [127](#)
- INSTITUTE OF INDUSTRIAL AND SYSTEMS ENGINEERS (2019). Industrial Engineering Body of Knowledge. Accessed: 2019-03-23. [19](#)
- KALINOWSKI, T., LIDÉN, T. & WATERER, H. (2018). Tight MIP formulations for bounded length cyclic sequences. *Preprint*. [121](#)
- KALINOWSKI, T., MATTHEWS, J. & WATERER, H. (2020). Scheduling of maintenance windows in a mining supply chain rail network. *Computers & Operations Research*, **115**. [121](#), [123](#), [125](#), [126](#), [127](#), [128](#), [129](#)

REFERENCES

-
- KHALOULI, S., BENMANSOUR, R. & HANAFI, S. (2018). Ant colony optimisation combined with variable neighbourhood search for scheduling preventive railway maintenance activities. *International Journal of Intelligent Engineering Informatics*, **6**, 78–98. [121](#), [128](#)
- KRUGER, D. (2012). The Return Circuit: Electric Traction, Lecture Notes of Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [34](#)
- LAKE, M., FERREIRA, L. & MURRAY, M. (2000). Minimising costs in scheduling railway track maintenance. *WIT Transactions on The Built Environment*, **50**. [121](#), [125](#), [126](#), [128](#)
- LI, R. & ROBERTI, R. (2017). Optimal scheduling of railway track possessions in large-scale projects with multiple construction works. *Journal of Construction Engineering and Management*, **143**. [121](#)
- LIDÉN, T. (2015). Railway infrastructure maintenance – a survey of planning problems and conducted research. *Transportation Research Procedia*, **10**, 574–583. [3](#), [120](#), [121](#)
- LIDÉN, T. (2016). *Towards concurrent planning of railway maintenance and train services*, vol. 1746. Linköping University Electronic Press. [121](#)
- LIDÉN, T. (2018a). *Concurrent planning of railway maintenance windows and train services*, vol. 1957. Linköping University Electronic Press. [121](#)
- LIDÉN, T. (2018b). Reformulations for integrated planning of railway traffic and network maintenance. In *18th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2018)*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. [121](#), [129](#)
- LIDÉN, T. (2020). Coordinating maintenance windows and train traffic: a case study. *Public Transport*, **12**, 261–298. [121](#), [123](#), [124](#), [125](#), [126](#), [127](#), [128](#), [129](#)
- LIDÉN, T. & JOBORN, M. (2016). Dimensioning windows for railway infrastructure maintenance: Cost efficiency versus traffic impact. *Journal of Rail Transport Planning & Management*, **6**, 32–47. [121](#)

REFERENCES

-
- LIDÉN, T. & JOBORN, M. (2017). An optimization model for integrated planning of railway traffic and network maintenance. *Transportation Research Part C: Emerging Technologies*, **74**, 327–347. [121](#), [124](#), [126](#), [127](#)
- LIDÉN, T., KALINOWSKI, T. & WATERER, H. (2018). Resource considerations for integrated planning of railway traffic and maintenance windows. *Journal of rail transport planning & management*, **8**, 1–15. [121](#), [124](#), [125](#), [126](#), [129](#)
- LOMBARD, P.J. (2010). Train Service Planning, Lecture Notes of Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [55](#), [57](#), [58](#), [62](#), [66](#)
- LOMBARD, P.J. (2017). Train Service Planning, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [59](#), [61](#)
- LOOIJ, P. (2017). Adjusting train routing in case of planned infrastructure maintenance. [121](#)
- LOUWERSE, I. & HUISMAN, D. (2014). Adjusting a railway timetable in case of partial or complete blockades. *European journal of operational research*, **235**, 583–593. [121](#), [124](#), [127](#)
- LUAN, X., MIAO, J., MENG, L., CORMAN, F. & LODewIJKS, G. (2017). Integrated optimization on train scheduling and preventive maintenance time slots planning. *Transportation Research Part C: Emerging Technologies*, **80**, 329–359. [121](#), [124](#), [125](#), [126](#), [127](#), [128](#), [129](#)
- LUSBY, R.M., LARSEN, J., EHRGOTT, M. & RYAN, D. (2011). Railway track allocation: models and methods. *OR Spectrum*, **33**, 843–883. [120](#)
- MACEDO, R., BENMANSOUR, R., ARTIBA, A., MLADENović, N. & UROŠEVIĆ, D. (2017). Scheduling preventive railway maintenance activities with resource constraints. *Electronic Notes in Discrete Mathematics*, **58**, 215–222. [121](#)
- MARTIN, P. (2004). *Railways in South Africa: Historic Highlights, Strategic Evolution and Transformation 1845 - 2002*. Spoornet. [9](#), [10](#)

REFERENCES

- MARUTLA, F. (2017). Track Terminology, Performance and Components, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [26](#), [27](#), [30](#), [31](#), [32](#), [33](#), [34](#), [39](#)
- MARUTLA, F. (2018). Infrastructure Management, Presentation at Railway Infrastructure Maintenance Management Course, University of Pretoria. [100](#), [101](#), [102](#), [108](#), [110](#)
- MATSEKE, A. (2016). Rail Bound Telecommunications, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [55](#)
- MITCHELL, J.S. (2014). *Physical Asset Management Handbook*. Reliabilityweb.com, 4th edn. [69](#), [71](#), [72](#), [74](#), [76](#), [77](#), [79](#), [80](#), [81](#), [85](#), [86](#), [87](#), [88](#), [92](#), [93](#), [96](#)
- OPENRAILWAYMAP WEBSITE (2021). Open Railway Map. Accessed: 2021-11-15. [133](#)
- OSTROFSKY, H. (2014). Train Authorisation Technology, Lecture Notes of Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [39](#), [42](#), [43](#), [44](#), [45](#), [49](#), [50](#), [51](#)
- OSTROFSKY, H. (2017). Train Authorisation Technology, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [41](#), [46](#), [47](#), [48](#), [51](#), [52](#), [53](#)
- PEARCE, R.H. & FORBES, M. (2019). Disaggregated benders decomposition for solving a network maintenance scheduling problem. *Journal of the Operational Research Society*, **70**, 941–953. [121](#)
- PENG, F., KANG, S., LI, X., OUYANG, Y., SOMANI, K. & ACHARYA, D. (2011). A heuristic approach to the railroad track maintenance scheduling problem. *Computer-Aided Civil and Infrastructure Engineering*, **26**, 129–145. [121](#), [126](#), [128](#)
- POURYOUSEF, H., TEIXEIRA, P. & SUSSMAN, J. (2010). Track maintenance scheduling and its interactions with operations: Dedicated and mixed high-speed rail (hsr) scenarios. In *Joint Rail Conference*, vol. 49071, 317–326. [121](#)

REFERENCES

- SAVELSBERGH, M., WATERER, H., DALL, M. & MOFFIET, C. (2015). Possession assessment and capacity evaluation of the central queensland coal network. *EURO Journal on Transportation and Logistics*, **4**, 139–173. [121](#), [124](#)
- SCHUTTE, C., KENNON, D. & BAM, W. (2016). The status and challenges of industrial engineering in South Africa. *The South African Journal of Industrial Engineering*, **27**. [18](#), [20](#), [25](#)
- SPEROTTO, F. (2015). The development of the industrial engineering profession in South Africa. *The South African Journal of Industrial Engineering*, **26**, 1–9. [17](#), [18](#), [25](#)
- SPOORNET (n.d.). *Spoornet Western Cape Region*. Spoornet. [9](#), [10](#), [13](#)
- SU, Z., JAMSHIDI, A., NÚÑEZ, A., BALDI, S. & DE SCHUTTER, B. (2017). Multi-level condition-based maintenance planning for railway infrastructures—a scenario-based chance-constrained approach. *Transportation Research Part C: Emerging Technologies*, **84**, 92–123. [121](#)
- TALBI, E. (2009). *Metaheuristics: From Design to Implementation*. Wiley, 1st edn. [118](#), [119](#)
- TFR (2012). *Traction substation electricians handbook revision*. Transnet. [34](#)
- TFR WEBSITE (2018a). Company Overview. Accessed: 2018-08-18. [13](#), [17](#)
- TFR WEBSITE (2018b). Railway Country. Accessed: 2018-08-18. [8](#), [9](#), [10](#), [11](#)
- TRANSNET WEBSITE (2017). Freight Rail 2017. Accessed: 2018-08-20. [16](#)
- TRANSNET WEBSITE (2018a). Financial Results Presentation. Accessed: 2018-08-20. [2](#), [14](#), [15](#)
- TRANSNET WEBSITE (2018b). Our Divisions. Accessed: 2018-08-19. [13](#)
- TRANSNET WEBSITE (2018c). Overview. Accessed: 2018-08-19. [13](#), [14](#)
- TRANSNET WEBSITE (2018d). Transnet Engineering. Accessed: 2018-08-19. [14](#)
- TRANSNET WEBSITE (2018e). Transnet Freight Rail. Accessed: 2018-08-19. [15](#)

REFERENCES

- TRANSNET WEBSITE (2018f). Transnet National Ports Authority. Accessed: 2018-08-19. [14](#)
- TRANSNET WEBSITE (2018g). Transnet Pipelines. Accessed: 2018-08-19. [14](#)
- TRANSNET WEBSITE (2018h). Transnet Port Terminals. Accessed: 2018-08-19. [14](#)
- UNIVERSITY OF STELLENBOSCH WEBSITE (2018). Department of Industrial Engineering. Accessed: 2018-12-02. [18](#)
- VAN AARDT, J. (2018). Railway Infrastructure Maintenance Management, Presentation at Railway Infrastructure Maintenance Management Course, University of Pretoria. [100](#), [102](#), [103](#), [105](#), [107](#), [110](#), [111](#), [112](#), [113](#), [114](#), [115](#), [116](#), [117](#)
- VAN AKEN, S., BEŠINOVIĆ, N. & GOVERDE, R.M. (2017a). Designing alternative railway timetables under infrastructure maintenance possessions. *Transportation Research Part B: Methodological*, **98**, 224–238. [121](#), [124](#), [126](#), [127](#), [128](#), [129](#)
- VAN AKEN, S., BEŠINOVIĆ, N. & GOVERDE, R.M. (2017b). Solving large-scale train timetable adjustment problems under infrastructure maintenance possessions. *Journal of rail transport planning & management*, **7**, 141–156. [121](#), [124](#), [126](#), [127](#), [128](#)
- VAN NIEKERK, J. (2012). Condition Monitoring Systems, Lecture Notes of Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [54](#)
- VAN NIEKERK, J. (2017). Integrated Train Condition Monitoring, Presentation at Introduction to Multidisciplinary Concepts in Railway Engineering Course, University of Pretoria. [54](#)
- VAN ZANTE-DE FOKKERT, J., DEN HERTOOG, D., VAN DEN BERG, F. & VERHOEVEN, J. (2007). The netherlands schedules track maintenance to improve track workers safety. *Interfaces*, **37**, 133–142. [121](#), [124](#)
- VANSTEENWEGEN, P., DEWILDE, T., BURGGRAEVE, S. & CATTRYSSSE, D. (2016). An iterative approach for reducing the impact of infrastructure maintenance on the performance of railway systems. *European Journal of Operational Research*, **252**, 39–53. [121](#), [124](#), [126](#)

REFERENCES

-
- VEELENTURF, L.P., KIDD, M.P., CACCHIANI, V., KROON, L.G. & TOTH, P. (2016). A railway timetable rescheduling approach for handling large-scale disruptions. *Transportation Science*, **50**, 841–862. [121](#), [124](#), [126](#), [127](#), [128](#)
- VLOK, P.J. (2014). Introduction to Elementary Statistical Analysis of Failure Time Data: Long Term Cost Optimization and Residual Life Estimation, Lecture Notes of Quality Management 444, University of Stellenbosch. [69](#), [70](#), [77](#), [79](#)
- ZHANG, C., GAO, Y., LI, W., YANG, L. & GAO, Z. (2018). Robust train scheduling problem with optimized maintenance planning on high-speed railway corridors: the china case. *Journal of Advanced Transportation*, **2018**. [121](#), [126](#), [127](#), [128](#)
- ZHANG, C., GAO, Y., YANG, L., KUMAR, U. & GAO, Z. (2019a). Integrated optimization of train scheduling and maintenance planning on high-speed railway corridors. *Omega*, **87**, 86–104. [121](#), [126](#), [127](#), [128](#)
- ZHANG, C., GAO, Y., YANG, L., GAO, Z. & QI, J. (2020a). Joint optimization of train scheduling and maintenance planning in a railway network: A heuristic algorithm using lagrangian relaxation. *Transportation Research Part B: Methodological*, **134**, 64–92. [121](#), [123](#), [126](#), [127](#), [128](#), [129](#)
- ZHANG, Q., LUSBY, R.M., SHANG, P. & ZHU, X. (2020b). Simultaneously re-optimizing timetables and platform schedules under planned track maintenance for a high-speed railway network. *Transportation Research Part C: Emerging Technologies*, **121**, 102823. [121](#), [122](#), [123](#), [124](#), [125](#), [126](#), [127](#), [128](#), [129](#)
- ZHANG, Y., D’ARIANO, A., HE, B. & PENG, Q. (2019b). Microscopic optimization model and algorithm for integrating train timetabling and track maintenance task scheduling. *Transportation Research Part B: Methodological*, **127**, 237–278. [121](#), [124](#), [125](#), [126](#), [127](#), [128](#), [129](#)
- ZHU, Y., GOVERDE, R.M. & QUAGLIETTA, E. (2018). Railway timetable rescheduling for multiple simultaneous disruptions. In *CASPT 2018: 14th conference on advanced systems in public transport and Transit-Data*. [121](#), [123](#), [124](#), [126](#), [128](#), [129](#)