

STOCHASTIC RAIL LIFE CYCLE COST MAINTENANCE MODELLING USING MONTE CARLO SIMULATION

RICK VANDOORNE

A dissertation submitted in partial fulfilment of the requirements for the degree of

MASTER OF ENGINEERING (TRANSPORTATION ENGINEERING)

In the

**FACULTY OF ENGINEERING, BUILT ENVIRONMENT & INFORMATION
TECHNOLOGY**

UNIVERSITY OF PRETORIA

February 2017

DISSERTATION SUMMARY

STOCHASTIC RAIL LIFE CYCLE COST MAINTENANCE MODELLING USING MONTE CARLO SIMULATION

RICK VANDOORNE

Supervisor: Professor P.J. Gräbe
Department: Civil Engineering
University: The University of Pretoria
Degree: Master of Engineering (Transportation Engineering)

The objective of this study was to quantify and determine trends in the uncertainty in the life cycle cost (LCC) associated with the maintenance and renewal (M&R) of the rail of a railway track under a fixed set of input parameters and conditions. Rail maintenance models were identified in the literature which use the mean or expected value of the input distributions to determine a corresponding mean or expected LCC. Although these models display important trends with regard to input parameters such as inspection intervals, they provide no means to quantify the uncertainty related to maintenance and renewal decisions. Thus, a numerical model was developed and programmed using MATLAB which allows the quantification of the uncertainty in the LCC estimated for a given set of conditions.

The model uses Monte Carlo simulation to determine the LCC associated with the installation, maintenance and renewal of the rail. The model incorporates imperfect inspections, a hazard function for rail fatigue defects modelled using the Weibull probability distribution and a P-F interval for rail fatigue defects modelled using an exponential probability distribution. The model also allows the modelling of maintenance as either perfect or minimal maintenance as well as the use of either flash butt or alumino-thermic welds to conduct the maintenance. This allowed the development of a method to assess which weld type to use to minimise the minimum attainable mean LCC.

The developed model was validated against a similar stochastic rail maintenance model from the literature. However, the model from the literature considers only the expected LCC and does not show any uncertainty related thereto. The novelty in this study therefore lies in the fact that the LCC uncertainty can be quantified in the form of a probability distribution at any given renewal tonnage for a given set of conditions.

It was found that the distribution of the LCC at a given renewal tonnage followed a lognormal probability distribution. The standard deviation of the lognormal distributions fitted using the method of maximum likelihood was used as a metric to quantify the uncertainty related to the life cycle cost at a given renewal tonnage. The LCC uncertainty was found to increase with an increase in inspection interval length. Furthermore, the uncertainty was also found to increase with a respective increase in renewal tonnage. For varying inspection interval lengths it was found that the uncertainty of combined maintenance costs (planned plus unplanned maintenance costs) tended more strongly towards the uncertainty in the planned maintenance costs for smaller inspection intervals and more strongly towards the uncertainty in unplanned maintenance costs for larger inspection intervals. A critical cost ratio was found of flash butt weld cost to alumino-thermic weld cost at which the minimum attainable mean LCC was equal. It is more economical to use flash butt welding for maintenance if the cost of flash butt welding maintenance produces a cost ratio lower than the critical cost ratio.

The developed model could allow railway operators to assess the risk associated with renewal of the rail at varying renewal tonnages for given conditions such as inspection interval lengths, detectability of rail fatigue defects and the arrival rate of rail fatigue defects.

DECLARATION

I, the undersigned hereby declare that:

I understand what plagiarism is and I am aware of the University's policy in this regard;

The work contained in this dissertation is my own original work;

I did not refer to work of current or previous students, lecture notes, handbooks or any other study material without proper referencing;

Where other people's work has been used this has been properly acknowledged and referenced;

I have not allowed anyone to copy any part of my dissertation;

I have not previously in its entirety or in part submitted this dissertation at any university for a degree.



Rick Vandoorne

11050692

02/02/2017

ACKNOWLEDGEMENTS

I wish to express my appreciation to the following people and organisations who have made this dissertation possible:

- a) My supervisor and mentor Prof. P.J. Gräbe for his continuous support and guidance throughout the entire research process and for always believing in my capabilities and pushing me further.
- b) Jackie van der Westhuizen for the guidance and feedback that he shared with me during the inception phase of the study.
- c) The Transnet Freight Rail Chair in Railway Engineering for the generous financial support which was provided during the study.
- d) The staff at the Civil Engineering Department at the University of Pretoria for always having an open-door policy which has allowed me to ask numerous questions.
- e) My fellow postgraduate students for the much-needed comic relief during our lunch breaks.
- f) Paul Vorster and Rudolph du Plooy who have kept me motivated and maintained my sanity during the study.
- g) Jaco Vorster for providing the photos of the welding conducted at the University of Pretoria's Test Track and providing advice which aided my choice to pursue my postgraduate studies.
- h) My friends and family for helping me maintain a balanced life style during the course of my postgraduate studies.
- i) My mother in particular for always supporting me in my studies both financially and emotionally.
- j) Rudolph du Plooy for helping proof read my dissertation.
- k) My girlfriend Carmen Weyers (who recently earned the title Dr Carmen Weyers) for helping proof read my dissertation and providing moral support and motivation during the writing of my dissertation. You have always provided a listening ear when I've needed one and have driven me to complete my dissertation in due time. I wish you all the best for your career as a newly qualified doctor!

TABLE OF CONTENTS

	PAGE
1 INTRODUCTION	1-1
1.1 BACKGROUND	1-1
1.2 OBJECTIVES OF THE STUDY	1-2
1.3 SCOPE OF THE STUDY	1-3
1.4 METHODOLOGY	1-3
1.5 ORGANISATION OF THE REPORT	1-4
2 LITERATURE REVIEW	2-1
2.1 TRACK COMPONENTS AND LOADING	2-1
2.1.1 Rail Geometry and Terminology	2-1
2.1.2 Track Forces	2-4
2.1.3 Introduction to Continuous Welded Rail Concepts	2-5
2.2 RAIL DEFECTS AND DETERIORATION	2-7
2.2.1 Rail Damage	2-8
2.2.2 Rail Fatigue Defect Classification	2-12
2.2.3 Rail Fatigue Defect Growth and Deterioration Behaviour	2-13
2.3 RAIL MAINTENANCE, RENEWAL AND INSPECTION	2-17
2.3.1 Rail Inspection Practices	2-18
2.3.2 Rail Grinding	2-18
2.4 MAINTENANCE STRATEGIES AND MAINTENANCE MODELLING	2-19
2.4.1 Maintenance Hierarchy	2-19
2.4.2 Maintenance Optimisation	2-21
2.4.3 Maintenance Modelling	2-22
2.5 STOCHASTIC MODELLING	2-25
2.5.1 The Weibull Distribution and Poisson Process for Rail Defect Modelling	2-25
2.5.2 The Hazard Function and Maintainable versus Non-Maintainable Failure Modes	2-27



2.5.3	The P-F Interval	2-28
2.6	LIFE CYCLE COST ANALYSIS AND UNCERTAINTY	2-30
2.6.1	Uncertainty Estimation	2-32
2.7	REVIEW OF PREVIOUS MAINTENANCE MODELS	2-32
2.7.1	Age-Usage Renewal Model	2-32
2.7.2	Stochastic Rail Defect Model	2-33
2.8	SUMMARY	2-41
3	NUMERICAL MODEL SETUP	3-1
3.1	INTRODUCTION TO THE NUMERICAL MODEL	3-1
3.2	INVERSION AND THE GENERATION OF RANDOM VARIABLES	3-3
3.2.1	Modelling the Arrival of Defects	3-5
3.2.2	Modelling the P-F Interval	3-10
3.3	PROCEDURAL LOGIC OF THE MODEL	3-11
3.3.1	Calculation Procedure for the Last Renewal Tonnage	3-14
3.3.2	Calculation Procedure for the Remaining Renewal Tonnages	3-23
3.3.3	Life Cycle Cost Calculation	3-25
4	ANALYSIS AND DISCUSSION OF RESULTS	4-1
4.1	REFERENCE CASE ANALYSIS	4-1
4.2	TYPICAL RESULTS	4-2
4.3	POSTERIOR DISTRIBUTION ANALYSIS	4-5
4.3.1	Distribution of Defect Inter-Arrival Tonnages	4-7
4.3.2	Distribution of P-F Interval Lengths	4-9
4.3.3	Distribution of Detectability of Defects	4-11
4.3.4	Accuracy, Number of Simulations and Computational Time	4-14
4.4	VALIDATION AGAINST AN EXISTING STOCHASTIC MODEL	4-14
4.5	ASSESSING LIFE CYCLE COST UNCERTAINTY	4-15
4.5.1	Fitting of a Defined Probability Distribution	4-16



4.6	INFLUENCE OF INSPECTION INTERVALS	4-23
4.6.1	Influence of Inspection Intervals on the Mean Life Cycle Cost	4-23
4.6.2	Influence of Inspection Intervals on Life Cycle Cost Uncertainty	4-26
4.7	SELECTING A WELD TYPE TO USE FOR RAIL MAINTENANCE	4-34
4.8	EFFECT OF DERAILMENTS	4-37
5	CONCLUSIONS AND RECOMMENDATIONS	5-1
5.1	CONCLUSIONS	5-1
5.1.1	Modelling Procedure	5-1
5.1.2	Behaviour of the Mean Life Cycle Cost and its Cost Components	5-1
5.1.3	Life Cycle Cost Distribution	5-2
5.1.4	Uncertainty Trends in the Life Cycle Cost Distribution	5-3
5.1.5	Using Flash Butt Welds over Alumino-thermic Welds for Maintenance	5-3
5.1.6	The Effect of Derailments	5-4
5.2	RECOMMENDATIONS	5-4
6	REFERENCES	6-1

LIST OF TABLES

	PAGE
Table 2.1: Rail force components and their sources (Selig & Waters, 1994; Esveld, 2001)	2-5
Table 2.2: Summary of system maintenance quality classification (Pham & Wang, 1996)	2-23
Table 2.3: Summary of Weibull methods used to model rail defects	2-26
Table 2.4: Stochastic parameters used by Zhao et al. (2006)	2-38
Table 2.5: Costs and non-stochastic parameters used by Zhao et al. (2006)	2-38
Table 3.1: Data stored for Category A and Category B defects for each virtual life cycle	3-12
Table 3.2: Description of the shapes used in the flow charts	3-14
Table 4.1: Stochastic input parameters used for the reference case analysis	4-1
Table 4.2: Cost parameters used for the reference case analysis	4-1
Table 4.3: Deterministic input parameters used for the reference case analysis	4-2
Table 4.4: P-value for fitted distributions at varying inspection intervals lengths for $T_R = 800$ MGT	4-17

LIST OF FIGURES

	PAGE
Figure 1.1: Fragmented rail from the Hatfield derailment in October 2000 (a) (Smith, 2003) (b) (Witt, 2007)	1-2
Figure 2.1: Rail profile terminology	2-3
Figure 2.2: Rail cross section indicating the vertical, transverse and longitudinal directions	2-3
Figure 2.3: Rail cross section indicating the vertical, horizontal and transverse planes	2-4
Figure 2.4: Rail force components	2-5
Figure 2.5: (a) Mobile flash butt welding machine (Plasser South Africa, 2015) (b) New FBW (Lekule, 2015)	2-6
Figure 2.6: The alumino-thermic welding process (a) alumino-thermic welding using a crucible (b) the ATW before being cleaned and grinded (c) the ATW being grinded to form a smooth rail profile (photos courtesy of Jaco Vorster)	2-7
Figure 2.7: Rail damage hierarchy (Reinschmidt et al., 2015)	2-8
Figure 2.8: Severe shelling on the gauge corner of a high rail (Kumar, 2006)	2-9
Figure 2.9: Tache ovale defect (Marais & Mistry, 2003)	2-10
Figure 2.10: (a) Head checks on the gauge corner of a rail (Cannon et al., 2003) (b) Macrograph of head check cracks (Olofsson & Nilsson, 2002)	2-11
Figure 2.11: (a) Squat defect as seen on the rail running surface (b) Rail break caused by the propagation of a squat defect (Grassie, 2012)	2-11
Figure 2.12: (a) Spalling caused by wheelburn (Cannon et al., 2003) (b) Spalling caused by head checking (Kumar, 2006)	2-12
Figure 2.13: Cause and effect diagram for rail deterioration (Kumar, 2006)	2-13
Figure 2.14: The effect of temperature variation on the critical defect size (redrawn from Jeong et al., 1997)	2-15
Figure 2.15: Critical defect size versus rail temperature (interpreted from Jeong et al., 1997)	2-15
Figure 2.16: The influence of rail temperature on safe-growth time for transverse fractures in rail with varying initial defect size (redrawn from Jeong et al., 1997)	2-16
Figure 2.17: The relationship between crack length propagation and simulated traffic as determined in a study by Jablonski et al. (1990) which serves to demonstrate the variability in rail defect growth rate under different loading spectra	2-17

Figure 2.18: Schematic illustration of the shadowing of ultrasound by shallow defects which can cause deeper more severe defects to remain undetected (redrawn from Cannon et al. (2003))	2-18
Figure 2.19: Maintenance strategy hierarchy (adapted from Kumar et al. (2008); Van der Westhuizen & Gräbe (2013))	2-20
Figure 2.20: Effect of perfect and minimal corrective and preventive maintenance on the hazard rate of a system (Lie & Chun, 1989)	2-21
Figure 2.21: Illustration of the concept of a P-F interval for a deteriorating component (Moubray, 1997)	2-29
Figure 2.22: Calculation steps and inputs required to develop a LCC plan (redrawn from Zoeteman (2001))	2-31
Figure 2.23: The development and consequence of rail failures (Zhao et al., 2006)	2-34
Figure 2.24: LCC expressed as the sum of maintenance cost and renewal cost (redrawn from Zhao et al. (2006))	2-39
Figure 2.25: Influence of inspection interval length s_I on the LCC curve (redrawn from Zhao et al. (2006))	2-40
Figure 2.26: Influence of inspection interval length s_I on the minimum attainable LCC (redrawn from Zhao et al. (2006))	2-40
Figure 3.1: The Inverse Transform Method for generating a random variable (redrawn from Rubinstein & Kroese (2008))	3-3
Figure 3.2: Correlation of a 4-dimensional Latin hypercube sample versus the length of the pseudorandom sample	3-5
Figure 3.3: Comparison of the hazard rate for a defect modelled using perfect maintenance, minimal maintenance and perfect maintenance with a grinding reduction factor	3-8
Figure 3.4: CDF for the defect inter-arrival tonnage t_d for a defect modelled using perfect maintenance, minimal maintenance and perfect maintenance with a grinding reduction factor	3-9
Figure 3.5: The effect of maintenance on the total system hazard rate for maintenance modelled using perfect maintenance and minimal maintenance	3-10
Figure 3.6: Flow chart illustrating the process of creating Category B defects in the model	3-15
Figure 3.7: Flow chart illustrating the process of determining the tonnage at which maintenance of Category B defects occurs $\overrightarrow{T_{B_m}}$ as well as the maintenance case number $\overrightarrow{C_B}$	3-16

- Figure 3.8: Flow chart illustrating the procedural logic contained within the Maintenance Function used to determine the tonnage at which maintenance is conducted \vec{T}_M as well as the maintenance case number \vec{C} 3-18
- Figure 3.9: Flow chart illustrating the procedural logic contained within the Detection Function used to determine at which inspection (if at all) the defect is detected 3-19
- Figure 3.10: Flow chart illustrating the creation of the Category A defects for welds present at $T = 0$ MGT 3-20
- Figure 3.11: Flow chart illustrating the process whereby Category A defects are created for welds created by maintaining Category B defects 3-21
- Figure 3.12: Flow chart illustrating the process of modelling the maintenance of all Category A defects and the creation of additional welds as a result thereof 3-22
- Figure 3.13: Flow chart illustrating the procedural logic followed to remove Category B defects which initiated at a tonnage larger than \vec{T}_{Rp} as well as any subsequent welds and Category A defects 3-23
- Figure 3.14: Flow chart illustrating the procedural logic followed to change the maintenance type of any remaining Category B defects with $\vec{T}_{Bm} > \vec{T}_{Rp}$ and removed any subsequent welds and Category A defects 3-24
- Figure 3.15: Flow chart illustrating the procedural logic used to remove any welds installed at tonnages larger than \vec{T}_{Rp} and subsequent defects 3-25
-
- Figure 4.1: Mean LCC versus renewal tonnage T_R for the reference case analysis 4-3
- Figure 4.2: Distribution of LCC at a renewal tonnage of 400 MGT for the reference case analysis presented as (a) a relative frequency plot and (b) a cumulative relative frequency plot 4-4
- Figure 4.3: Distribution of LCC versus renewal tonnage in three-dimensional space 4-5
- Figure 4.4: P-value (a) and number of data points used in the calculation thereof (b) at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of inter-arrival tonnages t_d for ATWs assuming perfect maintenance 4-7
- Figure 4.5: Illustrative behaviour of the CDF used to sample the defect inter-arrival tonnages t_d of Category A defects modelled using minimal maintenance and all Category B defects 4-8
- Figure 4.6: (a) P-value and (b) number of data points used in the calculation thereof at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of P-F interval lengths t_{P-F} for ATWs 4-10
- Figure 4.7: (a) P-value and (b) number of data points used in the calculation thereof at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of P-F interval lengths t_{P-F} for squat defects 4-10

- Figure 4.8: (a) P-value and (b) number of data points used in the calculation thereof at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of P-F interval lengths t_{p-F} for tache ovale defects 4-11
- Figure 4.9: (a) Proportion of successful inspections as well as (b) the total number of times defects were inspected at the respective renewal tonnages for varying number of simulated life cycle N_{sim} for ATW defects 4-12
- Figure 4.10: (a) Proportion of successful inspections as well as (b) the total number of times defects were inspected (b) at the respective renewal tonnages for varying number of simulated life cycle N_{sim} for squat defects 4-13
- Figure 4.11: Proportion of successful inspections (a) as well as the total number of times defects were inspected (b) at the respective renewal tonnages for varying number of simulated life cycle N_{sim} for tache ovale defects 4-13
- Figure 4.12: Comparison of LCC and maintenance costs calculated from the model developed by Zhao et al. (2006) and the model developed for this study 4-15
- Figure 4.13: Relative frequency plot of observed total LCC c_{LCC} at $T_R = 400, 600$ and 800 MGT respectively 4-16
- Figure 4.14: Illustration of the goodness-of-fit of different probability distributions to the ECDF of the LCC c_{LCC} for the reference case analysis with $s_I = 1$ MGT and $T_R = 800$ MGT 4-18
- Figure 4.15: Fitted CDF and the ECDF for the univariate distribution of total LCC c_{LCC} for the reference case analysis with $s_I = 1$ MGT at $T_R =$ (a) 200 MGT, (b) 400 MGT, (c) 600 MGT and (d) 800 MGT 4-19
- Figure 4.16: Lognormal probability distributions at fixed renewal tonnages fitted over the bivariate histogram of the original data from which they were fitted 4-21
- Figure 4.17: Lognormal PDFs at varying renewal tonnages for the reference case analysis with $s_I = 1.0$ MGT 4-22
- Figure 4.18: Normalised lognormal PDFs at varying renewal tonnages for the reference case analysis with $s_I = 1.0$ MGT 4-23
- Figure 4.19: The effect of the inspection interval s_I on the resulting mean LCC curve 4-24
- Figure 4.20: Planned, unplanned and combined maintenance costs for inspection interval lengths $s_I = 0.5$ and 10.0 MGT 4-25
- Figure 4.21: Proportion of combined maintenance cost originating from planned and unplanned maintenance activities for inspection interval lengths $s_I = 0.5$ and 10.0 MGT respectively 4-26
- Figure 4.22: Mean LCC curve showing the 1st and 99th percentiles for $s_I = 0.1$ and 5.0 MGT 4-27

Figure 4.23: Uncertainty in the form of standard deviation and its relationship with varying inspection interval lengths s_I as measured using a standard lognormal PDF normalised against the mean LCC at $T_R = 800$ MGT	4-27
Figure 4.24: Influence of renewal tonnage T_R on total LCC uncertainty σ_{LCC} at varying inspection intervals s_I	4-28
Figure 4.25: Influence of inspection interval length s_I on LCC uncertainty σ_{LCC} at varying renewal tonnages T_R	4-29
Figure 4.26: Influence of renewal tonnage T_R on planned maintenance cost uncertainty at varying inspection intervals s_I	4-30
Figure 4.27: Influence of renewal tonnage T_R on planned maintenance cost uncertainty at varying inspection intervals s_I on the renewal tonnage interval (0,200) to illustrate the curve behaviour	4-31
Figure 4.28: Influence of renewal tonnage T_R on unplanned maintenance cost uncertainty at varying inspection intervals s_I	4-32
Figure 4.29: Planned, unplanned and combined maintenance cost standard deviation for (a) $s_I = 0.5$ MGT and (b) $s_I = 10$ MGT	4-33
Figure 4.30: Influence of renewal tonnage T_R on total maintenance cost uncertainty at varying inspection intervals s_I	4-34
Figure 4.31: (a) PDF and (b) CDF for ATWs and FBWs	4-35
Figure 4.32: Average LCC c_{LCC} versus renewal tonnage T_R for varying c_{pFBW}/c_{pATW} cost ratios	4-36
Figure 4.33: Minimum attainable LCC $c_{LCC_{min}}$ versus cost ratio c_{pFBW}/c_{pATW}	4-37
Figure 4.34: Influence of derailments on the mean LCC c_{LCC} versus renewal tonnage T_R curves	4-39
Figure 4.35: The influence of derailments on (a) the minimum attainable mean LCC $c_{LCC_{min}}$ and (b) the renewal tonnage at which it occurs $T_{R_{cr}}$	4-40

LIST OF SYMBOLS

c_d	cost of a derailment
$c_{d_{norm}}$	cost of all derailments during a life cycle normalised against tonnage borne
c_f	cost of a unplanned maintenance action in the Zhao et al. (2006) model
$c_{f_{ATW}}$	cost of a unplanned maintenance action of a failed defect using alumino-thermic welds
$c_{f_{FBW}}$	cost of a unplanned maintenance action of a failed defect using flash butt welds
c_g	cost of rail grinding
c_I	cost of ultrasonic inspection
c_{LCC}	total life cycle cost
c_{LCC_f}	total life cycle cost of unplanned maintenance activities
c_{LCC_p}	total life cycle cost of planned maintenance activities
$\overline{c_{LCC}}$	mean total life cycle cost
$\overline{c_{LCC_{min}}}$	minimum attainable mean total life cycle cost
c_p	cost of a planned maintenance action in the Zhao et al. (2006) model
$c_{p_{ATW}}$	cost of a planned maintenance action of a detected defect using alumino-thermic welds
$c_{p_{FBW}}$	cost of a planned maintenance action of a detected defect using flash butt welds
c_R	cost of rail renewal
h	the total number of defects modelled using the Zhao et al. (2006) model
h_j	the number of Type B defects modelled in the Zhao et al. (2006) model
i	indexing variable
j	indexing variable
k	indexing variable
m	number of ultrasonic inspections conducted during the interval $(0, T)$
n_0	the number of alumino-thermic welds at $t = 0$ in the Zhao et al (2006) model
$n_{0_{ATW}}$	initial number of alumino-thermic welds
$n_{0_{FBW}}$	initial number of flash butt welds
n_{bins}	the number of equiprobable bins to use in the Chi-Square goodness-of-fit test
n_d	the number of derailments which occurred during a life cycle
n_{data}	the number of data points in the sample to be tested using the Chi-square test
$n_{0_{FBW}}$	initial number of FBWs
p	indexing variable
q_g	annual grinding rate



s_g	rail grinding interval length
s_I	inspection interval length
t	relative tonnage or time
t_{Ad}	tonnage-to-defect initiation relative to the defect or weld for a Category A defect
t_{AP-F}	P-F interval length for a Category A defect
t_{BP-F}	P-F interval length for a Category B defect
t_{Bd}	tonnage-to-defect initiation relative to the defect or weld for a Category B defect
t_d	defect inter-arrival tonnage
t_{P-F}	P-F interval length
C_A	maintenance case number for a Category A defect
C_B	maintenance case number for a Category B defect
LCC	the life cycle cost as calculated using the Zhao et al. (2006) model
N_d	expected number of defects detected by inspections using the Zhao et al. (2006) model
N_f	the number of rail failures as modelled using the Zhao et al. (2006) model
N_{sim}	number of virtual life cycles simulated
T	cumulative tonnage
T_{Ad}	virtual life cycle tonnage at defect initiation for a Category A defect
T_{Af}	virtual life cycle tonnage at failure for a Category A defect
T_{AI}	virtual life cycle tonnage at weld installation for a Category A defect
T_{Am}	virtual life cycle tonnage at which a Category A defect was maintained
T_{Bd}	virtual life cycle tonnage at defect initiation for a Category B defect
T_{Bf}	virtual life cycle tonnage at failure for a Category B defect
T_{Bm}	virtual life cycle tonnage at which a Category B defect was maintained
T_d	cumulative defect initiation tonnage
T_R	renewal tonnage
T_{Rcr}	renewal tonnage at which the mean life cycle cost is a minimum
U	a uniformly distributed random variable
α	Weibull probability distribution shape parameter
β	Weibull probability distribution scale parameter
γ	hazard rate reduction factor for grinding
η_j	the probability that a defect of type j will be detected by ultrasonic inspection



λ	expected rate of occurrence of an event
μ_j	the expected P-F interval length for defect type j
μ_{LCC}	the mean life cycle cost at a fixed renewal tonnage
$\lambda_a(t)$	hazard function of a single Type A defect at tonnage t in the Zhao et al. (2006) model
$\lambda_b(t)$	hazard function of a single Type B defect at tonnage t in the Zhao et al. (2006) model
$\nu_a(t)$	hazard function of all Type A defects for the Zhao et al. (2006) model
$\nu_b(t)$	hazard function of all Type B defects for the Zhao et al. (2006) model
ξ	probability that a rail failure will cause a derailment for the Zhao et al. (2006) model
ρ	Pearson's correlation coefficient
σ_{LCC}	standard deviation of the total life cycle cost at a fixed renewal tonnage
$\sigma_{LCC_{com}}$	standard deviation of the total combined maintenance cost at a fixed renewal tonnage
σ_{LCC_f}	standard deviation of the total unplanned maintenance cost at a fixed renewal tonnage
σ_{LCC_p}	standard deviation of the total planned maintenance cost at a fixed renewal tonnage



LIST OF ABBREVIATIONS

%HA	percentage head area
ATW	alumino-thermic weld
CDF	cumulative distribution function
CEN	European Committee for Standardization
CM	corrective maintenance
CWR	continuous welded rail
DSS	decision support system
ECDF	empirical cumulative distribution function
FBW	flash butt weld
IHHA	International Heavy Haul Association
IM	infrastructure manager
LCC	life cycle cost
LCCA	life cycle cost analysis
MGT	million gross tonnes
M&R	maintenance and renewal
NHPP	non-homogeneous Poisson process
PDF	probability density function
PEM	point estimation method
P-F Interval	potential failure – failure interval
PM	preventive maintenance
PMF	probability mass function
RAMS	reliability, availability, maintainability and safety
RCF	rolling contact fatigue
RV	random variable
SF	survivor function
UIC	International Union of Railways
UK	United Kingdom
WILMA	Wayside Intelligent Longstress Management System

1 INTRODUCTION

Demands being placed on railway transport infrastructure assets by operators and users alike are increasing (Khouy et al., 2014; Morant et al., 2012; Zoeteman, 2001). These demands include increasing speeds, axle loads and number of trains per day and must be satisfied whilst maintaining strict requirements on reliability, availability and safety. Furthermore, operational and maintenance costs are rising which means that these increasing demands must be met under tightening budget limits (Shafiee et al., 2016). These challenging operational conditions may conflict with the scheduling and execution of maintenance works (Zoeteman, 2001). Therefore, a well-structured, efficient and effective maintenance strategy is required to ensure that the required level of reliability, availability, maintainability and safety (RAMS) is achieved at a minimum life cycle cost (LCC).

1.1 BACKGROUND

One of the most important assets of the entire railway operation remains the rail itself (Esveld, 2001). The performance of the rail is largely dictated by the maintenance and renewal (M&R) decisions made during its life cycle. Current practices in M&R suggest that decisions are made based largely on past experience and expert opinion. Thus, there is a need for a decision support system (DSS) in order to quantitatively justify M&R decisions based on minimising the estimated LCC. A LCC approach will take into account short term budget requirements as well as long term capital expenditure costs.

Kumar et al. (2008) identify that the railway industry owns high risk in terms of potential loss of life and damage or destruction to freight and assets if the railway infrastructure is not operated and maintained effectively. As a component, the rail itself is particularly susceptible to high risk as it contains little redundancy (Cannon et al., 2003). There are numerous examples of the potential repercussions of poor infrastructure maintenance conditions in the railway industry. The Hatfield derailment in the United Kingdom (UK) in October 2000 killed 4 and injured 34 people and had serious financial repercussions for both the infrastructure manager (IM) and the maintenance company (The Guardian, 2006; Kumar et al., 2008). The derailment was caused by a fatigued rail which subsequently fragmented into more than 200 pieces as illustrated in Figure 1.1 (Smith, 2003). The overall consequential cost of the derailment to Network Rail was estimated at £733 million (Kumar et al., 2008). Another maintenance related incident saw the derailment of the German ICE at Eschede on 3 June 1998 which claimed the lives of more than 100 people (Smith, 2003). This incident was caused by a fatigue fracture of a wheel and was thus not rail-related but still serves to illustrate the importance of maintenance.



Figure 1.1: Fragmented rail from the Hatfield derailment in October 2000 (a) (Smith, 2003) (b) (Witt, 2007)

Infrastructure failures in the railway industry are more likely to manifest themselves as operational delays rather than derailments. However, these operational delays occur with much higher frequency than derailments. Furthermore, an exponential relationship between daily traffic volume and delay costs resulting from in-service failures has been suggested by Schlake et al. (2011) which provides a strong business case for investigating and understanding the relationship between costs associated with the implementation of different maintenance policies and the costs associated with train delay and derailments. A careful balance must be achieved between the costs associated with maintenance practices and the risk costs associated with the lack thereof in order to ensure the continued profitable operation of railways (Endrenyi et al., 2001).

1.2 OBJECTIVES OF THE STUDY

A model was developed using Monte Carlo simulation to determine the probabilistic distribution of LCC as a function of renewal tonnage. The objectives achieved through the model were to:

- Illustrate the bivariate probability distribution of LCC versus renewal tonnage.
- Determine whether an existing family of probability distributions can be fitted to the distribution of the LCC for a fixed renewal tonnage and given set of input parameters.
- Identify and discuss trends with regard to the uncertainty in the LCC distribution using an appropriate metric.
- Determine an appropriate method to determine whether flash butt welding or alumino-thermic welding should be used to minimise the minimum attainable mean LCC.

The model had to incorporate the following:

- The modelling of the arrival of rail fatigue defects using an appropriate probability distribution.
- Imperfect inspections.
- The influence of preventive rail grinding on the arrival of rail fatigue defects.
- The influence of derailments on the mean LCC.
- The use of either alumino-thermic welds (ATWs) or flash butt welds (FBWs) for maintenance activities.
- An increase in hazard rate due to the increase in weld quantity associated with maintenance activities.

1.3 SCOPE OF THE STUDY

The following assumptions apply with regard to the life cycle costing procedure:

- No discount rate was taken into account. Therefore, implicitly the discount rate and inflation rate are equal.
- Yearly maintenance costs are fixed.
- The disposal cost of any assets at the end of the life time are not included.
- All capital costs are assumed to occur at the beginning of the life cycle.
- All costs specified include labour, plant and delay costs.

Only rail fatigue defects were modelled in this study. The following defects were not considered:

- Rail corrugations.
- Rail wear.
- Rail buckling or fracture due to thermally induced stresses.

1.4 METHODOLOGY

A literature study was conducted to understand the current state-of-the-art in maintenance modelling. Previous rail maintenance models published in the literature were studied. A mathematical background to the statistical techniques used in the model was also obtained during the literature study. Thereafter, a model was developed and programmed using MATLAB. The model took into account the knowledge acquired through the literature study and was aimed at modelling input parameters in a flexible manner such that assumptions of previous models studied in the literature review could be overcome. Thereafter, numerous analyses were run with varying input parameters. Data from the analyses were subsequently

interpreted, analysed and discussed with regard to the objectives of the study outlined in Section 1.2. Conclusions and recommendations for future research in the field were then made.

1.5 ORGANISATION OF THE REPORT

This dissertation is structured as follows:

- Section 1 provides an introduction and background to the study. A summary is provided of the objectives, scope and methodology of the study.
- Section 2 serves as a review of the applicable literature in the field of concern.
- Section 3 describes the operation of the numerical model developed.
- Section 4 presents the results of the model and provides an interpretation and discussion thereof.
- Section 5 summarises the results and concludes the study providing recommendations for future research in the field.

2 LITERATURE REVIEW

This section provides a review of the literature studied to supplement the study. Initially, a brief introduction to basic railway concepts and components is provided as considered relevant to this study. The focus is then shifted to a discussion of the physics of deterioration and failure with regard to rail defects and rail failure. Thereafter, incorporation of these concepts into stochastic analysis principles is discussed and basic life cycle cost (LCC) principles are reviewed. This section is concluded with a summary.

2.1 TRACK COMPONENTS AND LOADING

Conventional or ballasted railway track is the predominant track structure type used throughout the world. Although slab track structures are becoming progressively more popular for high speed railway lines, their use in freight railway lines is limited to sections of track passing through tunnels. Therefore, it should be noted that the focus of this study is on conventional track in a heavy haul setting.

The components of a conventional permanent way structure may be broadly categorised into two groups namely the superstructure and the substructure. The superstructure consists of the rails, fastening system, pads and sleepers. The substructure consists of the ballast, subballast, placed soil and natural ground (Selig & Waters, 1994).

Each track component has specific functions to fulfil in order to ensure efficient and safe operation of the railway track system as a whole. The rails specifically serve to (Esveld, 2001):

- Guide the train wheels evenly and continuously.
- Transfer the concentrated wheel loads to the sleepers without excessive deflection.
- Act as conductors for the signal circuit.
- Provide a traction return leg for the traction circuit.

The remaining permanent way track components serve numerous functions too. However, these are not listed in order to maintain brevity. For a comprehensive description of the functions of the remaining track components refer to Esveld (2001).

2.1.1 Rail Geometry and Terminology

Numerous rail profiles are in existence across the globe. A list of approximately 40 different rail profiles is provided by Esveld (2001). Different national standards exist for the classification of rail profiles. On

South African heavy haul lines the 60 E1 profile (previously classified as the UIC 60 profile) is predominantly used as classified by the European Committee for Standardization (CEN). General terminology used with regard to a typical rail profile is illustrated in Figure 2.1. The rails are inclined at an angle of 1:20 from the vertical towards the gauge side of the track. This rail cant is provided by the sleeper shape. Figure 2.2 and Figure 2.3 show the terminology used to indicate directions and planes respectively with respect to the rail.

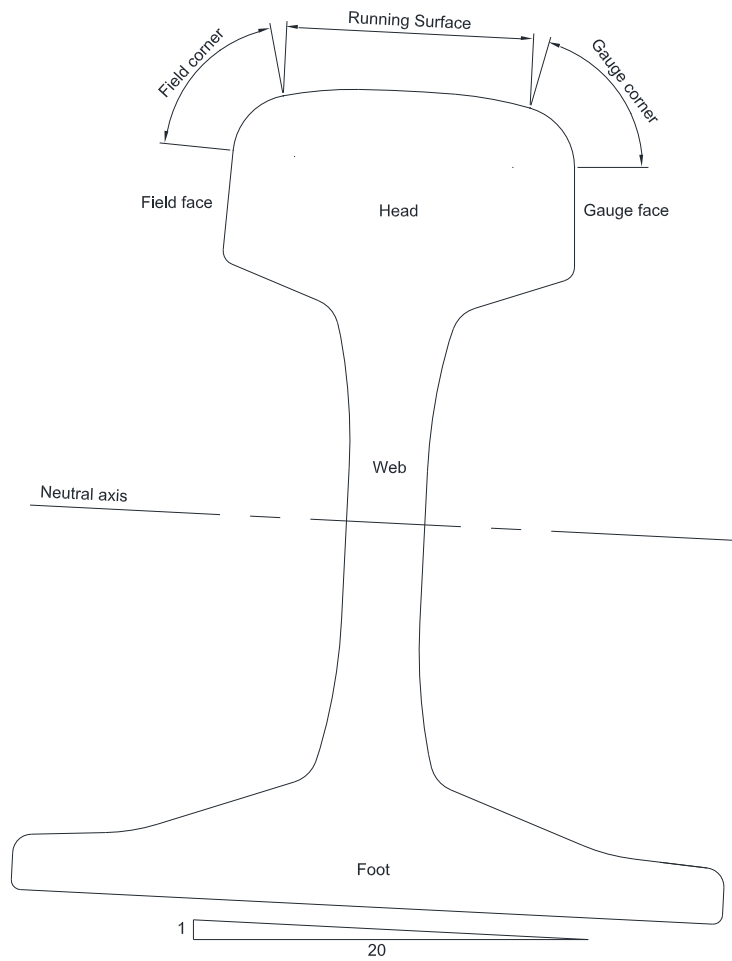


Figure 2.1: Rail profile terminology



Figure 2.2: Rail cross section indicating the vertical, transverse and longitudinal directions

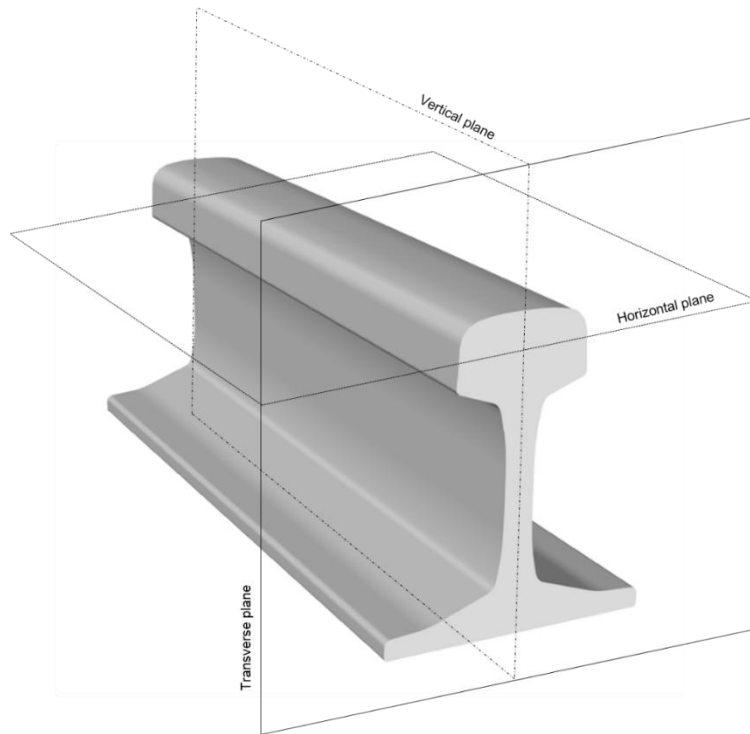


Figure 2.3: Rail cross section indicating the vertical, horizontal and transverse planes

2.1.2 Track Forces

Forces generated at the wheel-rail interface propagate upward into the wheel, axle and suspension before ultimately reaching the sprung body of the vehicle. These forces also propagate downward through the rail, fasteners, pads and ultimately into the substructure (Smith, 2003). The interactions between the components of the railway track structure are important when considering the influence of the condition of one component on the condition of other related components as it is known that the deterioration (and thus the condition of track components) is not independent of the condition of related components (Patra, 2007). The focus of this study is on the rail.

Rail forces are divided into vertical, lateral and longitudinal components as illustrated in Figure 2.4. The potential sources of each force component are indicated in Table 2.1. Depending on the nature of these rail forces they can be further divided into a static and a dynamic component. Dynamic loads are generally a result of (Esveld, 2001):

- Track irregularities and track stiffness variations.
- Discontinuities in the rail running surface such as corrugations.
- Discontinuities at welds, joints and points.
- Hunting of vehicle bogies and wheel flats or out of round wheels.

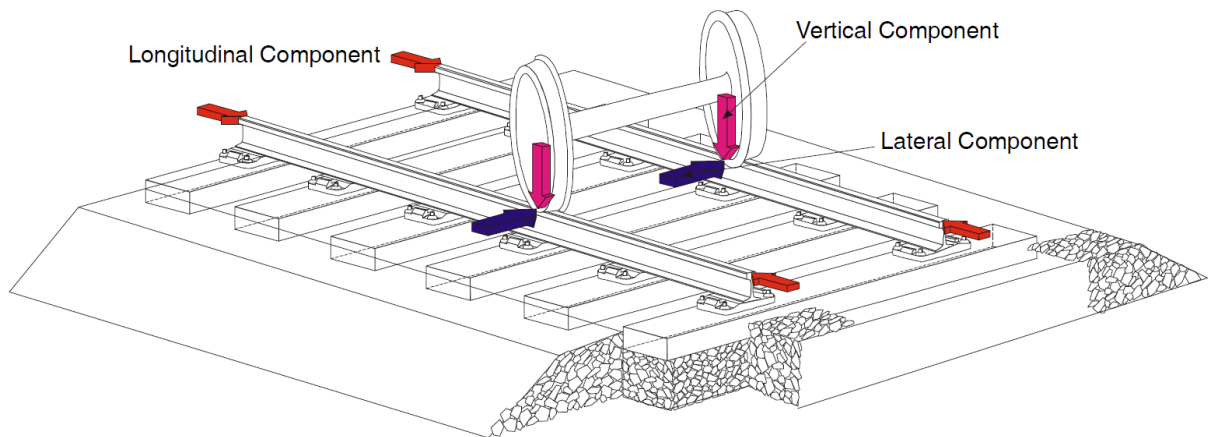


Figure 2.4: Rail force components

Table 2.1: Rail force components and their sources (Selig & Waters, 1994; Esveld, 2001)

<i>Force Component</i>	<i>Sources</i>
Vertical	<ul style="list-style-type: none"> • Vertical wheel force • Uplift force ahead of wheel
Lateral	<ul style="list-style-type: none"> • Lateral wheel force • Buckling reaction force
Longitudinal	<ul style="list-style-type: none"> • Locomotive traction force • Braking force • Thermal force • Rail wave action • Track creep • Shrinkage stresses cause by welding

2.1.3 Introduction to Continuous Welded Rail Concepts

In jointed track thermal expansion and contraction of the rail is permitted at the joints. This results in no or insignificant build-up of longitudinal thermal stresses in the rail as it is allowed to breath at these joints. However, the discontinuous rail running surface at these joints tends to generate high dynamic loads which results in a maintenance intensive system due to the increased deterioration at these bolted joints. A popular alternative is the use of continuous welded rail (CWR) which creates a smoother rail running surface through the elimination of bolted joints in the rail. The advantages of CWR include a reduction in maintenance and an increase in rail service life (Esveld, 2001). CWR consists of sections

of rail manufactured using the rolling process, which are then joined on-site through the use of welds to form long continuous rail segments. The welding techniques used can generally be classified into two main groups namely:

- Flash butt welding
- Alumino-thermic welding

In South Africa rail segments are typically joined off-site at production plants into long sections using flash butt welds (FBWs). These long sections which are approximately 240 m long are then transported to site using a Rail Car. Thereafter, the sections are joined on-site using alumino-thermic welds (ATWs). Mobile flash butt welding machinery was introduced by Plasser & Theurer in 1973 (Esveld, 2001). However, the use of mobile flash butt welding equipment on South African railways is limited. The alumino-thermic welding technique is easier to conduct on-site as mobile flash butt welding equipment is large and expensive. However, from a weld integrity viewpoint FBWs are preferred over ATWs, as FBWs are less prone to the development of internal defects originating from the welding process. Evidence of this can be seen when considering rail break statistics and their association with the different weld types (Sawley & Reiff, 2000; Marais & Mistry, 2003). Figure 2.5 (a) shows a mobile flash butt welding machine from Plasser & Theurer. Figure 2.5 (b) shows a completed FBW. The alumino-thermic welding process as conducted during the installation of the University of Pretoria's Test Track is illustrated in Figure 2.6.

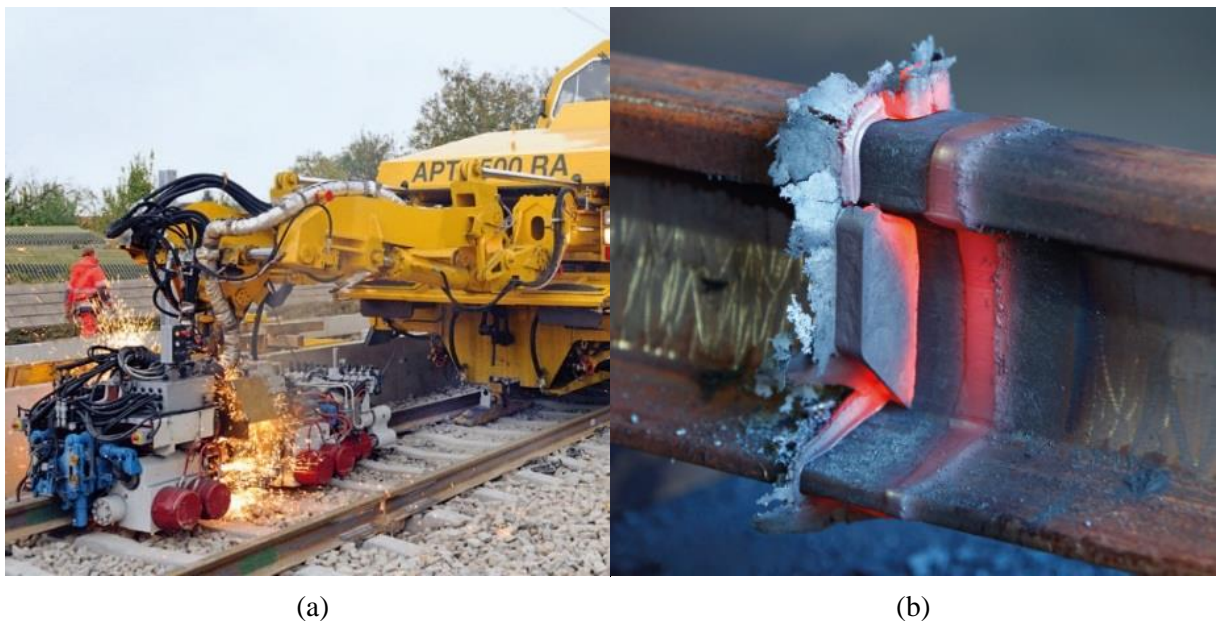


Figure 2.5: (a) Mobile flash butt welding machine (Plasser South Africa, 2015) (b) New FBW (Lekule, 2015)

2.2.1 Rail Damage

The nomenclature on rail damage is poorly standardised as noted by Reinschmidt et al. (2015). This situation is aggravated by the fact that many of the damage mechanisms are interrelated. This makes it challenging to group and compare rail defects according to failure modes and failure effects. Nonetheless, Reinschmidt et al. (2015) suggests a rail damage hierarchy as shown in Figure 2.7.

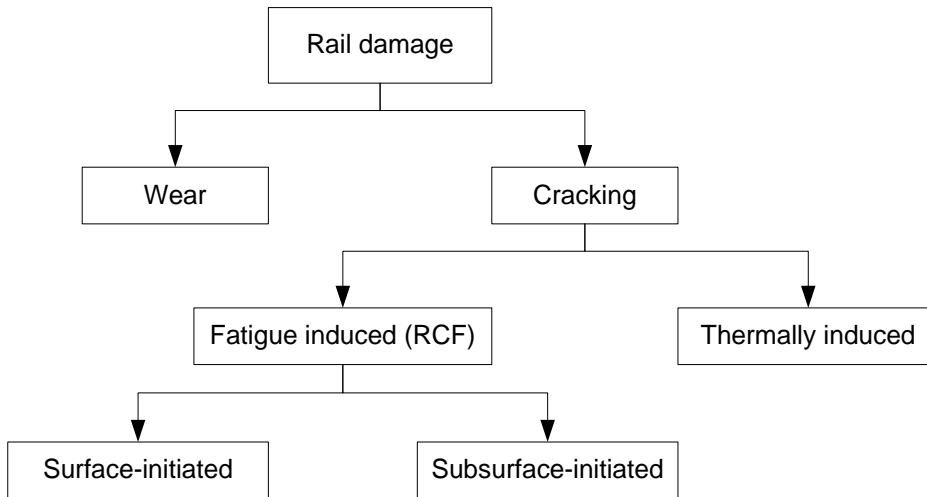


Figure 2.7: Rail damage hierarchy (Reinschmidt et al., 2015)

Fröhling (2007) defines two regimes of rail deterioration in line with Figure 2.7 namely the wear regime and the stress regime. Deterioration in the wear regime is fairly predictable whereas deterioration in the stress regime is non-linear and more difficult to measure, control and predict. With increasing track loading rail deterioration tends towards the stress regime rather than the wear regime. The stress regime is characterised by rolling contact fatigue (RCF) type defects and failures. Only stress regime damage mechanisms are considered for this study.

Steel rails are subject to fatigue-type failures due to the cyclic nature of their loading. Thus, the predominant failure mode of in-service steel rails is fatigue induced. This fatigue-based mechanism is commonly referred to in railway literature as RCF. Cannon et al. (2003) describes three stages by which a fatigue initiated rail failure may occur. These three stages are:

- 1) A fatigue crack or defect must be present within the rail.
- 2) The fatigue crack then grows in size with loading cycles.
- 3) The rail then ultimately breaks if this fatigue crack is not maintained and reaches a critical defect size.

Fatigue crack initiation may occur randomly or be related to the aging of the rail with use. The time or tonnage for a defect to progress from a potential failure (or defect) to a functional failure is termed the P-F interval (Moubray, 1997). Critical defect size for internal transverse defects was studied by Jeong et al. (1997) using a fracture mechanics approach. Some typical rail fatigue defects include but are not limited to (Kumar, 2006; Reinschmidt et al., 2015):

- Shelling
- Tache ovale or kidney defects
- Head checks
- Squats
- Spalling

RCF defects can be subdivided into surface-initiated defects and subsurface-initiated defects (Cannon & Pradier, 1996). Surface-initiated RCF defects can be controlled to a certain extent through the use of preventive grinding measures. Generally, sub-surface RCF defects are a result of non-metallic imperfections originating from the manufacturing process which then grow with RCF load cycles. Examples of such defects are shelling and tache ovale defects.

Shelling initiates near the surface at stress concentrations created by inclusions near the rail surface. Shelling typically occurs under lubricated heavy axle loading conditions where the peak shear stress occurs just below the surface of the rail due to low adhesion conditions. Cracking at the inclusions can cause loss of material in the rail running surface or can ultimately propagate downwards resulting in a transverse fracture of the rail (Reinschmidt et al., 2015). Shelling is commonly found on the gauge corner of the high rail in lubricated curves. Shelling resulting in the loss of material from the rail running surface can be seen in Figure 2.8 (Kumar, 2006).

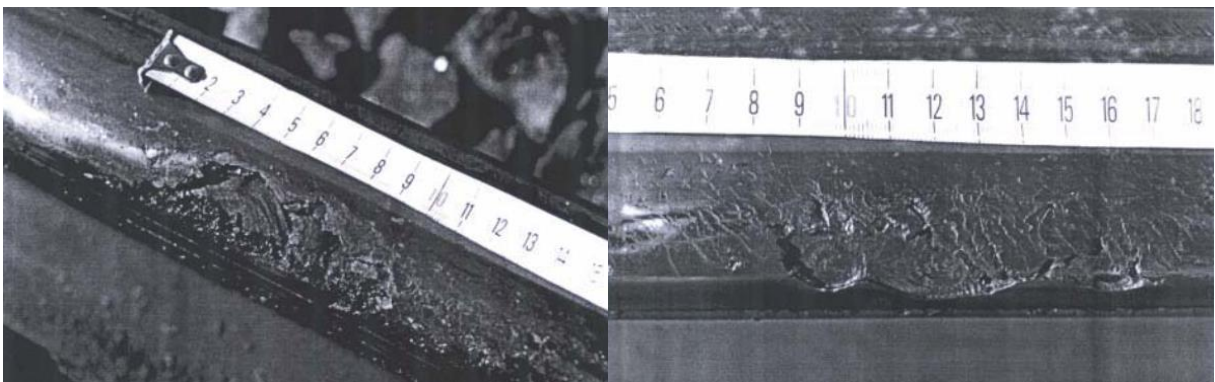


Figure 2.8: Severe shelling on the gauge corner of a high rail (Kumar, 2006)

Tache ovale defects (also called kidney defects) originate at non-metallic inclusions forming during the rail manufacturing or welding process. The number of these defects has substantially been reduced due to the introduction of more advanced rail production processes (Cannon et al., 2003). Figure 2.9 shows a rail break which occurred as a result of a tache ovale defect and describes the phases in the progression to ultimate failure.

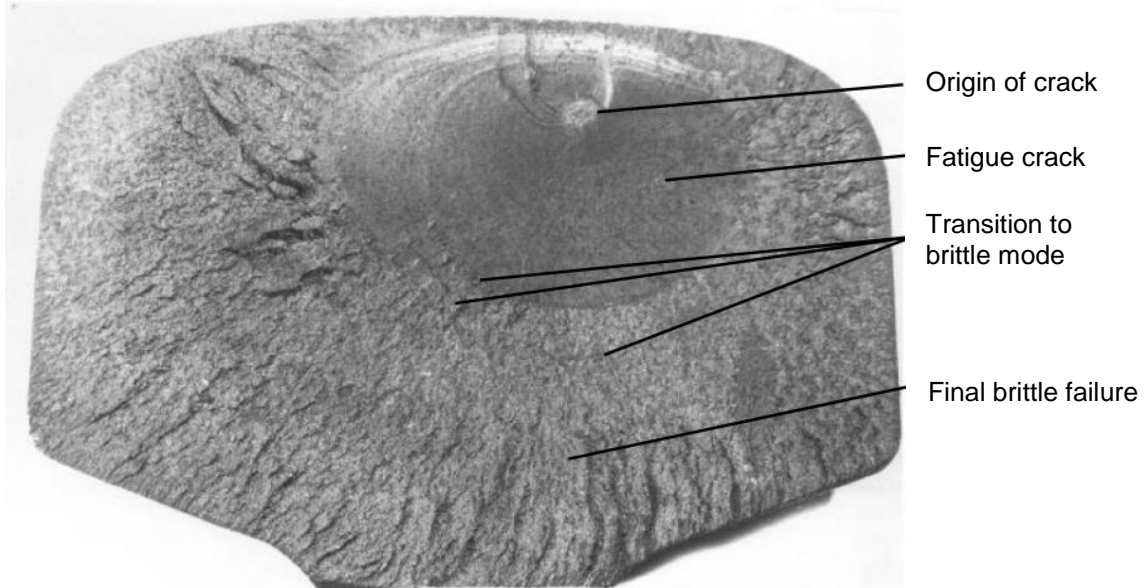


Figure 2.9: Tache ovale defect (Marais & Mistry, 2003)

Surface initiated defects originate either at locations of high shearing stresses at or near the surface of the rail or as a result of slipping at the wheel-rail interface which can create high temperatures and lead to the formation of brittle martensite which subsequently cracks (Reinschmidt et al., 2015). Examples of surface-initiated RCF defects are head checking, squats and spalling.

Head checking commonly occurs under high adhesion conditions. This results in the peak shear stress occurring at the rail surface itself. Head checking or cracking at the surface occurs due to the exhaustion of the steel's fatigue resistance induced either by non-conformal wheel/rail profiles or thermally due to wheel burn or slip (Cannon et al., 2003). One of the consequences of head checking is spalling. This occurs if head checks intersect with other crack paths resulting in the loss of material from the rail running surface (Kumar, 2006). Alternatively, head checks can propagate vertically and cause a transverse fracture of the rail. Figure 2.10 shows an example of head checking. The development of squats is governed by a more complicated mechanism than that of head checks. A thorough review of squat defects is provided by Grassie (2012). Figure 2.11 shows pictures of a squat defect.

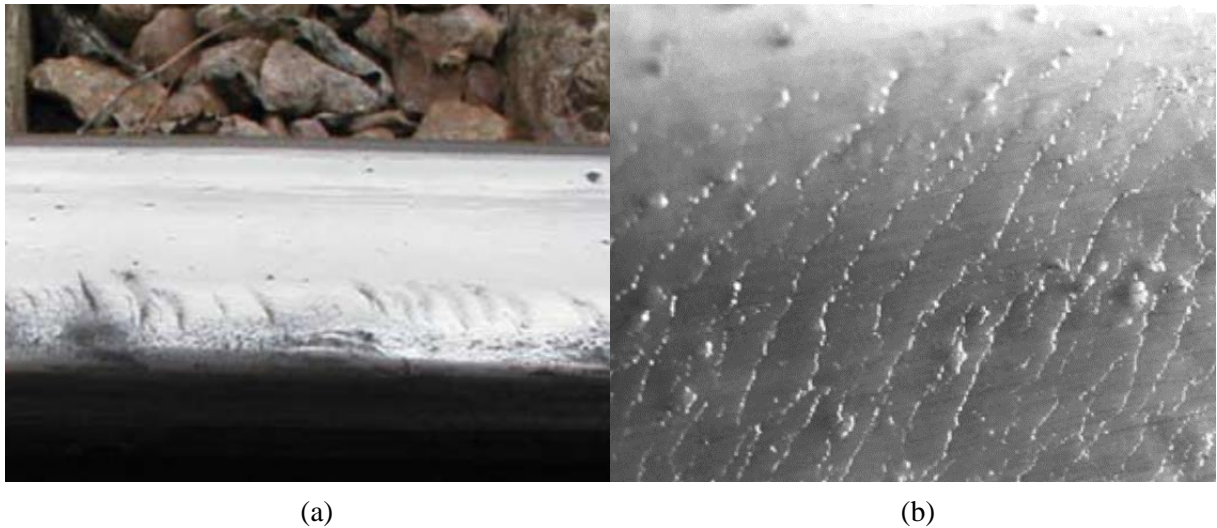


Figure 2.10: (a) Head checks on the gauge corner of a rail (Cannon et al., 2003) (b) Macrograph of head check cracks (Olofsson & Nilsson, 2002)

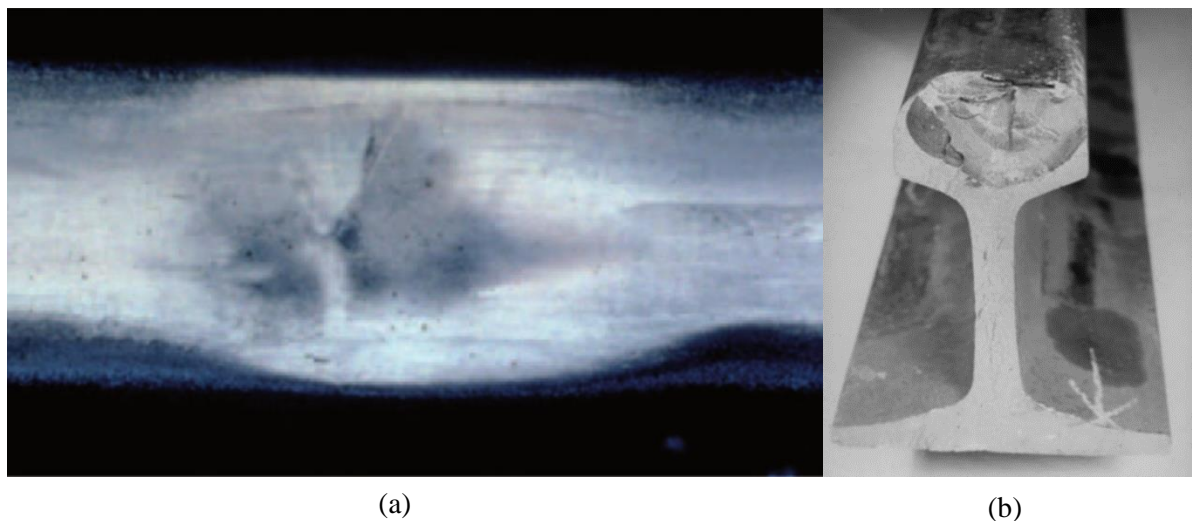


Figure 2.11: (a) Squat defect as seen on the rail running surface (b) Rail break caused by the propagation of a squat defect (Grassie, 2012)

Spalling is the loss of material from the rail running surface. However, spalling is not solely caused by head checks but can also be thermally induced by wheel burn. Muster et al. (1996) describes the difference between shelling and spalling as being origin-based. Shelling originates internally and is related to an existing imperfection in the rail whereas spalling originates on the surface as a result of the exhaustion of the steel's fatigue resistance by either high contact stresses or thermal overloading. Examples of what has been termed rail spalling by Cannon et al. (2003) and Kumar (2006) are shown in Figure 2.12 (a) and (b) respectively.

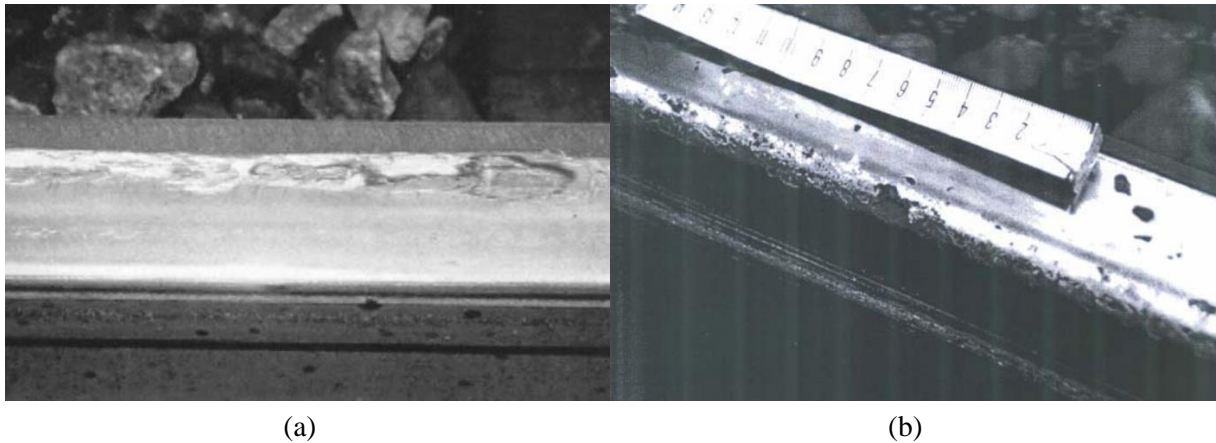


Figure 2.12: (a) Spalling caused by wheelburn (Cannon et al., 2003) (b) Spalling caused by head checking (Kumar, 2006)

Other rail damage mechanisms not discussed include but are not limited to (Kumar, 2006):

- Plastic flow
- Corrugation (short-wave and long-wave)
- Wear

2.2.2 Rail Fatigue Defect Classification

Rail defect classification and reporting is a vital part of improving rail defect modelling. Cannon et al. (2003) acknowledges the importance of rail defect classification and data reporting. With the aid of adequate data and thorough statistical analysis, effective measures can be developed to curb rail breaks from occurring. Numerous rail defect classification systems are in use around the world by different railway operators (Sawley & Reiff, 2000).

In order to model rail defects it is important that they are categorised according to their failure mode type so that they can be grouped into homogeneous populations for statistical analysis. A few examples of rail defect classification as found in the literature is discussed. According to Cannon et al. (2003) rail defects may be classified as follows:

- Defects which originate during rail manufacture. Examples of such defects include tache ovale and shelling defects.
- Defects which originate through improper handling, installation and/or use. An example of such a defect is a wheel burn.
- Defects which originate due to RCF. Examples of RCF defects are head checks, squats and spalling.

Alternatively, Marais & Mistry (2003) classify rail defects into two different groups namely:

- Defects related to rail joints in the form of welds and fish plate joints.
- Defects related to rail quality such as tache ovale defects.

Generally, all the classifications present in reputable literature have merit and are useful in understanding the origin and location of rail defects. The most appropriate rail defect classification system to use for stochastic modelling is one which enables grouping of defects into homogeneous populations according to initiation and deterioration behaviour.

2.2.3 Rail Fatigue Defect Growth and Deterioration Behaviour

In a rail segment or weld, there will generally exist a state at which an anomaly becomes detectable with modern inspection technologies such as ultrasonic, eddy current, magnetic particle or wave inspection before a functional failure occurs. This state shall be referred to as a potential failure or defect (Zhao et al., 2006).

In order to optimise maintenance and inspection activities associated with the rail, the deterioration behaviour of rail defects needs to be studied. The deterioration behaviour of a rail is a complex process which is influenced by numerous factors. A comprehensive study of the rail deterioration process by Kumar (2006) is summarised in the cause and effect diagram in Figure 2.13.

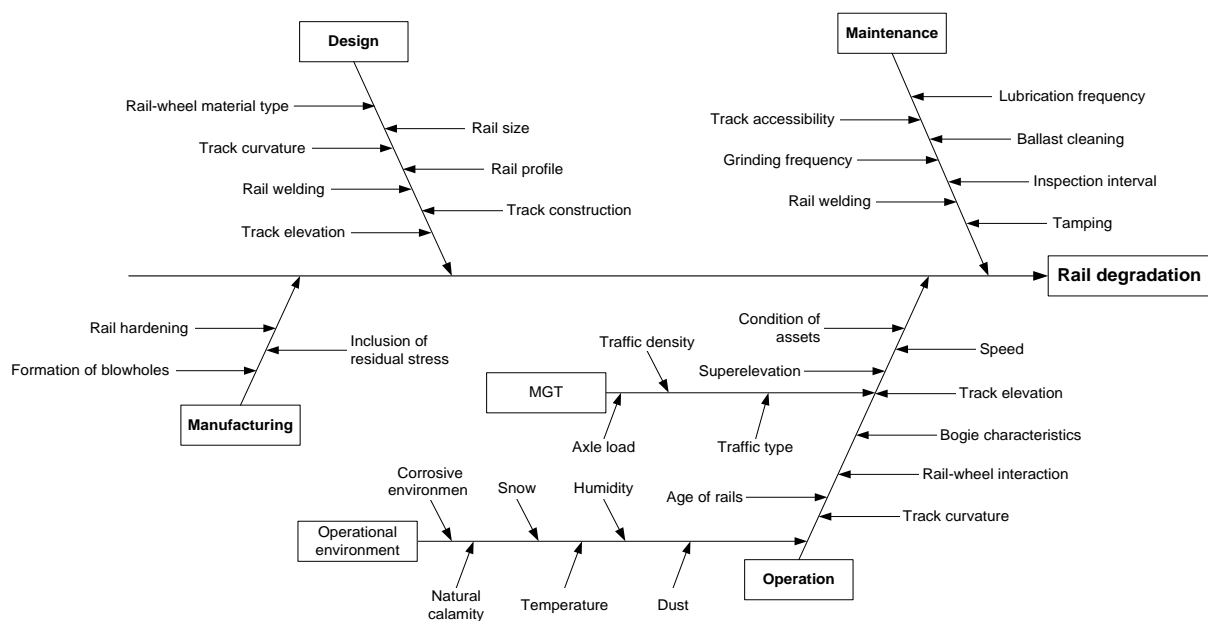


Figure 2.13: Cause and effect diagram for rail deterioration (Kumar, 2006)

Jeong et al. (1997) analysed transverse fractures (detail fractures) in a rail using fracture mechanics. In their research the transverse defect was simplified and modelled as elliptical in shape with an aspect ratio (minor axis over major axis) of 0.7. The effects of bending, thermal and residual stress were considered. Furthermore, factors taking into account local stress gradients and finite boundaries relative to the position of the defect within the rail head were used to determine the stress intensity factor for the transverse fracture as a function of major axis length and rail temperature. This was used to determine the critical defect size at which the rail will completely fracture or fail.

Figure 2.14 shows the general relationship between defect size and the stress intensity factor as a function of rail temperature. From Figure 2.14 it can be seen that for a steel of given fracture toughness the critical defect size in percentage head area (%HA) at which the rail is expected to fail decreases in relation to a decrease in rail temperature. This is due to the increasing tensile stresses in the rail which correspond with decreasing temperatures.

The critical defect size as a function of rail temperature is shown in Figure 2.15 as determined by Jeong et al. (1997) for a specific combination of rail profile type, vertical-lateral load combination, rail curvature, foundation stiffness, rail neutral temperature and fracture toughness of the rail. An approximately linear relationship is observed between rail temperature and critical defect size over the range of temperature values presented with the critical defect size increasing with rail temperature as expected. It should also be noted that dynamic loading may significantly reduce the critical defect size as given by these results (Cannon et al., 2003).

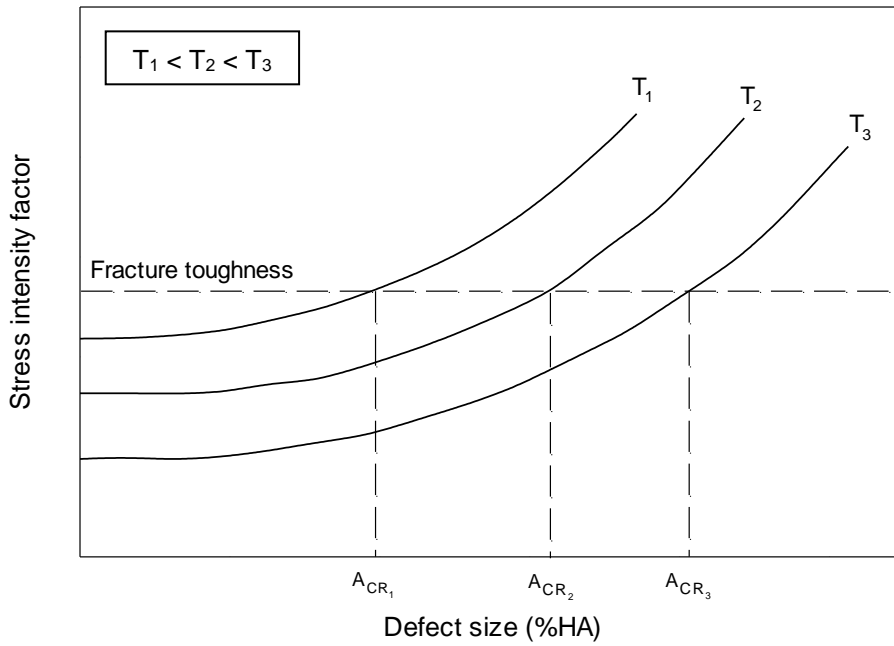


Figure 2.14: The effect of temperature variation on the critical defect size (redrawn from Jeong et al., 1997)

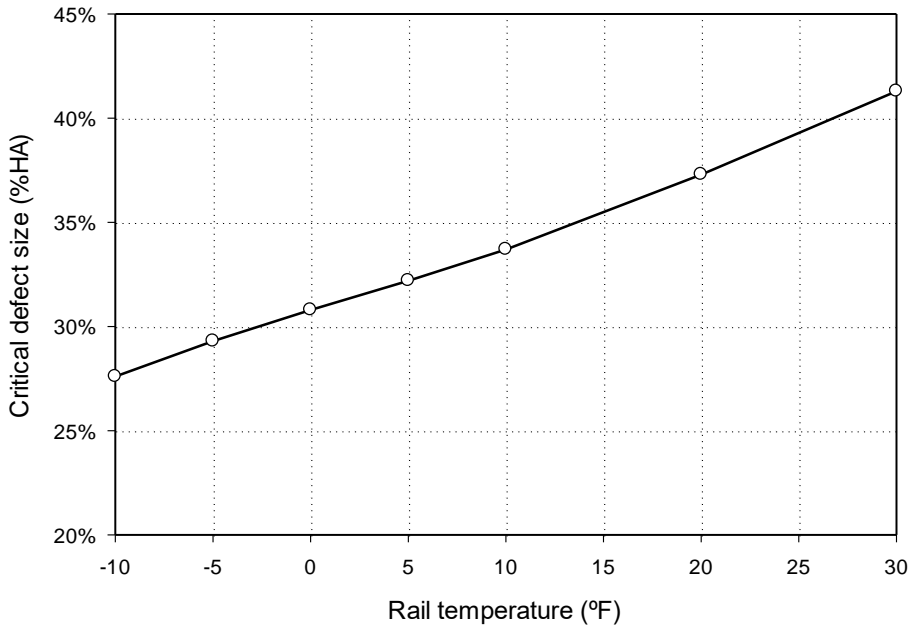


Figure 2.15: Critical defect size versus rail temperature (interpreted from Jeong et al., 1997)

Jeong et al. (1997) also considered the rate of growth of transverse fractures. The growth rate of transverse fractures was modelled using a relationship proposed by Orringer et al. (1984). Jeong et al. (1997) accounted for the fact that fatigue cracks only grow under the tensile portion of a stress cycle and that the stress cycle imparted under two adjacent bogies of two adjacent wagons is not of a constant

stress amplitude. The time for transverse fractures of varying initial size to reach critical defect size as a function of rail temperature is shown in Figure 2.16. It is of interest to note that the range of rail temperatures investigated by Jeong et al. (1997) is relatively low if viewed in the South African context. Furthermore, the traffic intensity used to relate cumulative traffic to safe-growth time in days was 0.5 million gross tonnes (MGT) per day which is relatively high when compared to the tonnage on the South African Heavy Haul Coal Line. Therefore, using the same fracture mechanics approach used by Jeong et al. (1997) a longer safe-growth time for transverse fractures can be expected for a typical railway line in South Africa if all other parameters are kept constant.

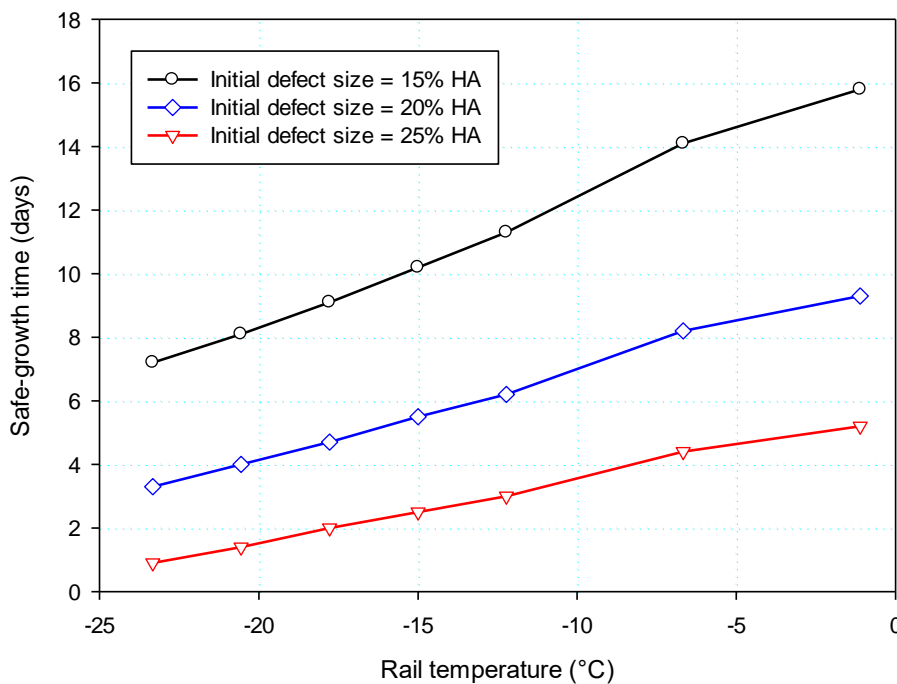


Figure 2.16: The influence of rail temperature on safe-growth time for transverse fractures in rail with varying initial defect size (redrawn from Jeong et al., 1997)

Despite numerous studies on fatigue defect growth and advancements in rail inspection technologies, a great deal of variability and subjectivity are still present when sizing and reporting internal rail defects in-service. Furthermore, longitudinal stress variations in the rail which fluctuate with rail temperature are known to affect apparent defect size due to opening and closing of the fracture by these thermal stresses (Orringer et al., 1984). Thus, on a hot day the apparent defect area as measured by an ultrasonic inspection may be smaller than the true defect size.

Tests by Jablonski et al. (1990) conducted on laboratory compact tension specimens to simulate transverse rail fractures reflected a significant dependency of the crack propagation rate on the loading stress spectra. Furthermore, the crack growth rate was found to be higher in curved track segments due

to the increased lateral load component. The loading cycles were applied to simulate stress spectra from typical train consist configurations in the United States. These load cycles were then rescaled to represent cumulative traffic in million gross tons. Figure 2.17 shows the effect that the loading spectra from the different typical consist configurations had on the rate of crack length propagation in the study by Jablonski et al. (1990). It should be noted that the stress spectra length in time units per train was not constant. The number of loading cycles per simulated train was also not constant and neither was the simulated axle load.

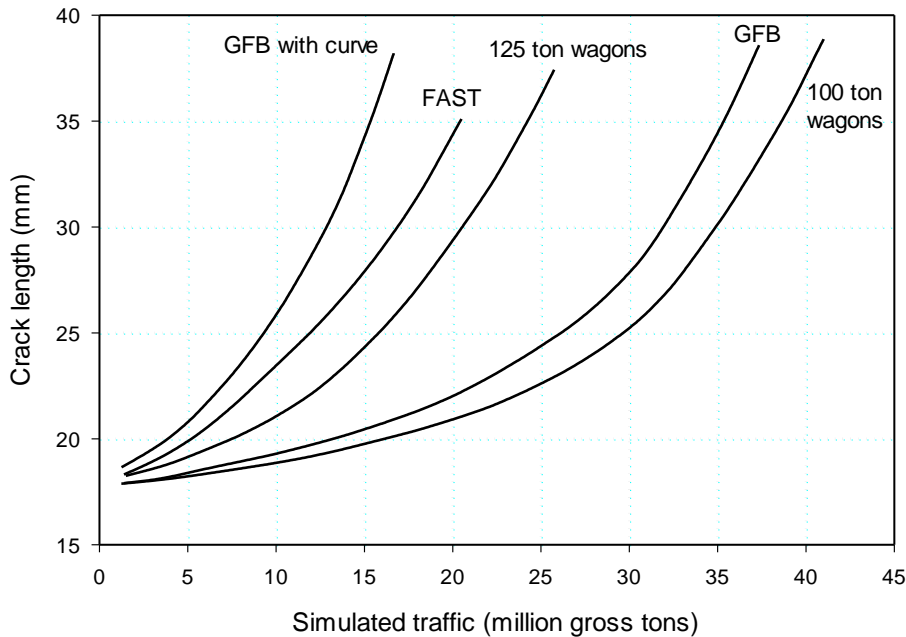


Figure 2.17: The relationship between crack length propagation and simulated traffic as determined in a study by Jablonski et al. (1990) which serves to demonstrate the variability in rail defect growth rate under different loading spectra

The results from these experiments suggest that rail transverse fracture defect growth rates should be determined for the actual loading conditions on the railway track under consideration. Due to the numerous variables which influence rail defect growth from a deterministic approach, a stochastic approach in the form of a P-F interval probability distribution is often used. This stochastic approach is discussed in Section 2.5.3.

2.3 RAIL MAINTENANCE, RENEWAL AND INSPECTION

In this section the maintenance, renewal, inspection and condition monitoring practices for a typical heavy haul line are discussed with regard to the rail of the track. Rail maintenance and renewal (M&R)

practices effect the deterioration of the rail and ballast and thus have an effect on the LCCs associated with rail breaks.

2.3.1 Rail Inspection Practices

A combination of rail inspection methods is generally used on a railway line. These methods include the use of ultrasonic inspection cars, handheld ultrasonic inspection probes as well as visual inspections. Once detected, defects may be either removed as part of a planned maintenance procedure or left in place if deemed non-critical and removed at a later stage (Kumar et al., 2008). Rail inspections are not perfect in the sense that it is possible that defects remain undetected after an inspection. One possible reason for this is attributed to shadowing of deeper defects by superficial defects such as head checks. The concept of shadowing is illustrated in Figure 2.18. Use of technologies such as eddy current inspection and improved ultrasonic probe arrangements can mitigate the shadowing effect (Cannon et al., 2003).

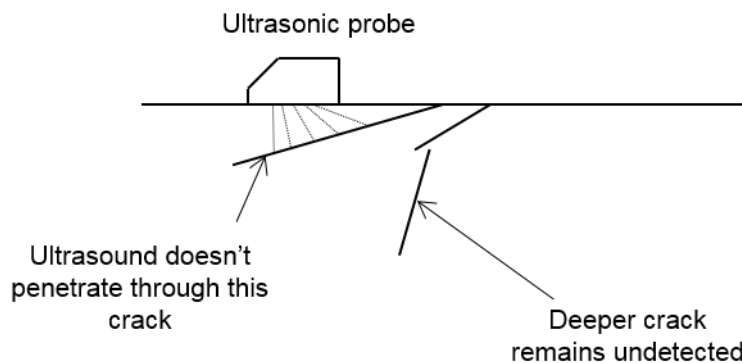


Figure 2.18: Schematic illustration of the shadowing of ultrasound by shallow defects which can cause deeper more severe defects to remain undetected (redrawn from Cannon et al. (2003))

If defects remain undetected for an extended period of time they can reach the critical defect size and cause rail breaks or failures. However, rail breaks do not always manifest themselves as derailments. The use of signalling circuits to detect rail breaks significantly reduces the number of rail breaks which cause derailment (Kumar et al., 2008).

2.3.2 Rail Grinding

Hearle & Johnson (1983) as cited by Magel et al. (2005) state that new RCF cracks on the surface of the rail are only a fraction of a millimetre in size. At this stage of the crack life the growth rate of the crack remains relatively slow. It is only at depths of approximately 1 – 3 mm that the growth rate accelerates.

Natural wear of the rail can truncate these RCF cracks to a certain extent but often not enough. Preventive grinding introduces artificial wear and is aimed at truncating these RCF cracks at an early stage to prevent cracks propagating downward and causing rail fracture. Grinding strategies also ensure conformal wheel/rail profiles reducing stresses in the rail and further mitigating RCF crack initiation amongst other damage mechanisms (Fröhling, 2007).

There are numerous influential factors to consider when assessing the effectiveness of a preventive grinding policy and it is not within the scope of this study to account for these factors within a LCC model. Preventive grinding policy optimisation remains a popular research field on its own due to these complexities. The reader is referred to studies by Magel et al. (2003), Wu et al. (2005) and Zoeteman et al. (2014). However, it is important that grinding should in some manner be incorporated with regard to its effect on surface-initiated RCF defects in a LCC model (Zhao et al., 2007).

2.4 MAINTENANCE STRATEGIES AND MAINTENANCE MODELLING

Modelling processes associated with the maintenance and deterioration of infrastructure are vital in determining LCCs. A deterioration model aims to describe the change in condition of an infrastructure asset over a specific period of time under specific operating conditions. These models may be either deterministic or stochastic, mechanistic or empirical. Furthermore, the complexity of a deterioration model may vary. The model may use either functional condition measures to quantify the health of the system being modelled or structural condition measures to quantify the health of the individual components of which the system is comprised. It should be noted that the deterioration and therefore the condition of interrelated systems or components will not be independent from each other (Andrade, 2008). The focus of this section is to provide an overview of the literature available in the fields of deterioration and maintenance modelling.

2.4.1 Maintenance Hierarchy

A generalised hierarchy of maintenance strategies for a railway is shown in Figure 2.19. Maintenance hierarchies in the literature vary. For example Bevilacqua & Braglia (2000) use a different maintenance hierarchy than that of Van der Westhuizen & Gräbe (2013) and Kumar et al. (2008). Furthermore, the maintenance strategies in Figure 2.19 may be considered very broad with Wang (2002) identifying over 14 different sub-categories of replacement and preventive maintenance policies alone.

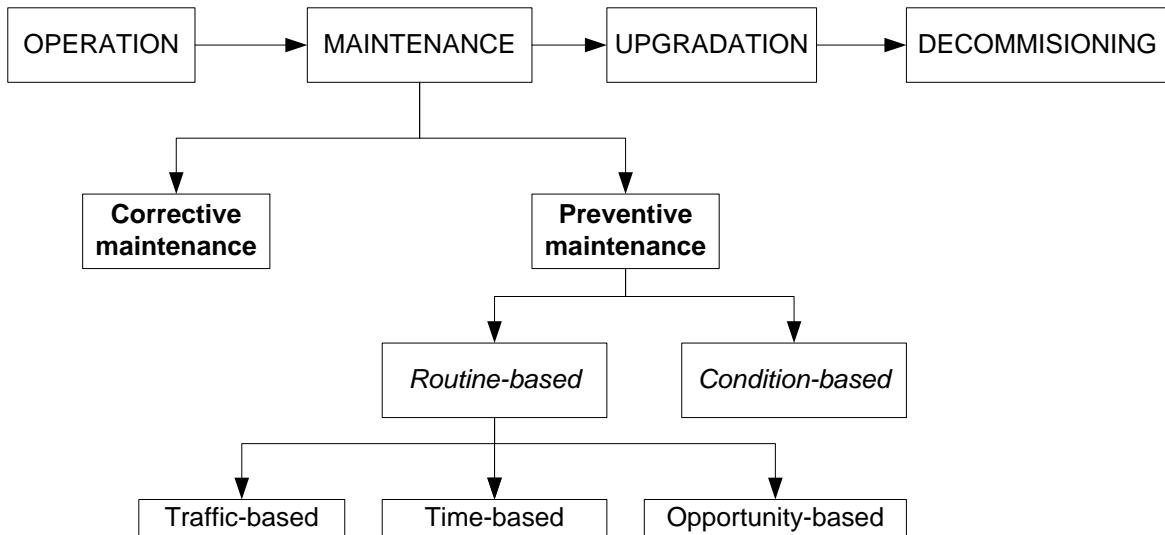


Figure 2.19: Maintenance strategy hierarchy (adapted from Kumar et al. (2008); Van der Westhuizen & Gräbe (2013))

The definitions supplied in *MIL-STD-721B* for corrective maintenance (CM) and preventive maintenance (PM) are adopted for this study. Thence, CM is defined as all actions performed as a result of failure in order to restore an item to a specified condition whereas PM is defined as all actions performed in an attempt to retain an item in a specified condition by providing systematic inspection, detection and prevention of incipient failures (Department of Defense, 1981). CM may be regarded as reactive in the sense that it is only initiated once failure has occurred. This maintenance strategy should only be purposely implemented for non-critical components of a system for which failure would not result in significant loss of reliability, availability or safety. Thus, CM should not be encouraged for the rail, as rail failures generally have significant effects on a railway operation. A rail break is an example of an event which results in CM which should be avoided (Van der Westhuizen & Gräbe, 2013).

PM is aimed at mitigating failure and is subdivided into routine-based maintenance and condition-based maintenance. Routine-based maintenance takes place at time-, traffic- or opportunity-based intervals. Condition-based maintenance, unlike routine-based maintenance is initiated at a specific condition level of the asset or component (Van der Westhuizen & Gräbe, 2013). The condition of the asset can be determined at discrete intervals by inspections or continuously by modern condition monitoring equipment such as the Wayside Intelligent Longstress Management System (WILMA) on the Coal Line. The goal of a maintenance policy should be to minimise corrective maintenance and where possible implement condition-based instead of routine-based preventive maintenance (Kumar et al., 2008).

Figure 2.20 shows the general manner in which different types of PM and CM influence the hazard rate function and assists in understanding the differences between corrective and preventive maintenance. In Figure 2.20 a limiting hazard rate has been set at which preventive maintenance is conducted. A detailed discussion of different maintenance models and their associated influence on the hazard function is given in 2.4.3.

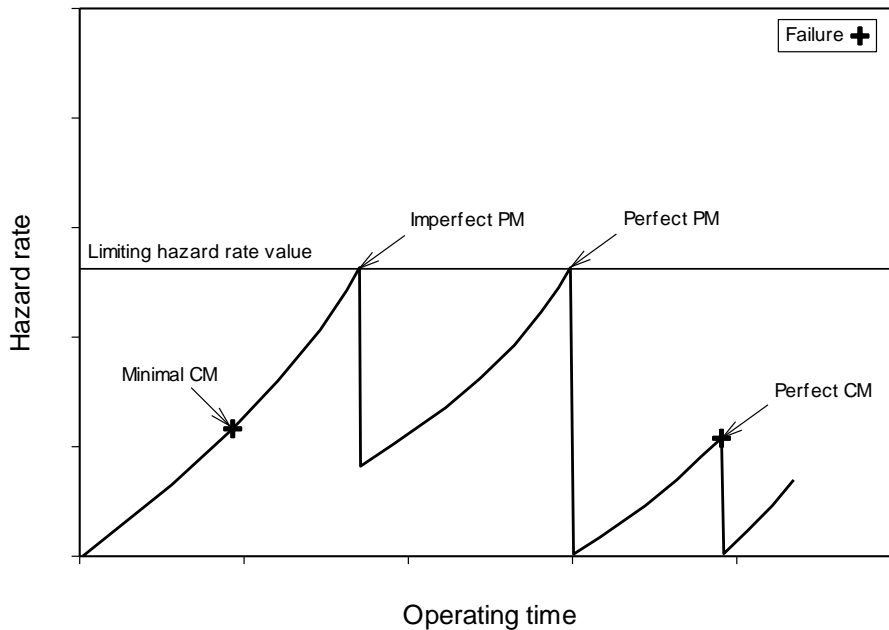


Figure 2.20: Effect of perfect and minimal corrective and preventive maintenance on the hazard rate of a system (Lie & Chun, 1989)

Nguyen & Murthy (1981) provide two important reasons justifying the development and implementation of PM models. Firstly, it is generally more cost effective to repair a non-failed system than it is a failed system and secondly, at some point in the life cycle of a system it becomes more economic to replace the system than to continue repairing it.

A combination of maintenance strategies is often used to maintain an asset-base. This allows the minimisation of maintenance related interference and failures (Van der Westhuizen & Gräbe, 2013). Bevilacqua & Braglia (2000) developed a method for selecting the optimal maintenance strategies for different components of a system using an analytical hierarchy process.

2.4.2 Maintenance Optimisation

Maintenance optimisation models are concerned with determining the balance between costs and benefits of a specific maintenance policy using either a single or a combination of different maintenance

strategies as discussed in Section 2.4.1 (Dekker, 1996). Furthermore, stochastic optimisation models can be constructed either on a risk-basis or an uncertainty-basis. The former assume that the time to failure distribution is known whilst the latter do not. According to Wang (2002) maintenance policies can be optimised in one of four ways:

- Minimisation of the maintenance cost rate.
- Maximisation of the system reliability.
- Minimisation of the maintenance cost rate whilst satisfying system reliability constraints.
- Maximisation of the system reliability whilst satisfying maintenance cost constraints.

Minimising the maintenance cost rate alone is not necessarily the best approach to maintenance policy optimisation. Wang (2002) appreciates the contradiction that many maintenance optimisation models in the literature focus only on cost rate minimisation whilst one of the most important functions of preventive maintenance is the improvement of system reliability. Therefore, it is vital that an optimal maintenance policy addresses both cost and reliability.

2.4.3 Maintenance Modelling

Pham & Wang (1996) identify 5 different categories of maintenance quality according to the condition to which the system is restored by the maintenance activity. These categories are summarised in Table 2.2. When modelling maintenance, Kijima (1989) suggested using perfect maintenance for a system containing a single, structurally simple component and suggested the use of minimal maintenance when modelling a system with many components each with unrelated failure modes. However, this differentiation is subjective and in reality the maintenance activity is likely to fall in-between these two extremes and is subsequently termed imperfect maintenance. Therefore, there is a need to study and model systems using the concept of imperfect maintenance (Pham & Wang, 1996). Worst maintenance occurs when the system fails immediately after maintenance of a component. Worse or worst maintenance can occur when faulty or incorrect parts are used in the maintenance service or if an incorrect part is repaired or replaced (Nakagawa & Yasui, 1987).

Table 2.2: Summary of system maintenance quality classification (Pham & Wang, 1996)

<i>Repair type</i>	<i>Effect on hazard rate of system</i>	<i>Laymen's description</i>
Perfect maintenance	Decreases to same value as when new	As good as new
Imperfect maintenance	Decreases but not to same value as when new	Between good as new and bad as old
Minimal maintenance	Not affected	As bad as old
Worse maintenance	Increases	Worse than old
Worst maintenance	N/A	Instant failure

Pham & Wang (1996) reviewed numerous methods of modelling imperfect maintenance present in the literature at that time. Eight categories were identified by Pham & Wang (1996). The 8 methods for modelling imperfect maintenance all assume that the state of the system is always known with certainty and that the system exists in either an operational state or a failed state. The methods differ predominantly with regard to the manner in which they treat the change in hazard rate of the component or system with preventive or corrective maintenance actions.

The first method assumes that there is a probability that the maintenance action (corrective or preventive) will be either perfect or minimal with the two probabilities being complementary. Thus, the hazard rate is either returned to “as good as new” or “as bad as old”. This method was adopted by Nakagawa (1979). The second method, developed by Block et al. (1985) is similar to the first with the exception that the probabilities of perfect or minimal repair are age-dependant. Thus, as the age of the component increases the probability that CM will restore the component to “as good as new” decreases.

The third method assumes that the effect of maintenance can be represented by a reduction in the hazard rate for the component. This reduction is characterised by a so-called improvement factor which may be constant or proportional to the hazard rate at which the maintenance takes place (Malik, 1979; Chan & Shaw, 1993). The fourth model uses the concept of virtual age. By this model the probability distribution for the time until the next failure is dependent on the virtual age of the component when it previously failed (Kijima, 1989).

The fifth model type does not work on the principle of hazard rate. The fifth model type is referred to as a shock model whereby a new component (with zero damage) is subjected to increments of shock at random times. With each shock increment a random damage increment is imposed. This damage accumulates until the component's damage limit is reached and the component fails. The damage level remains constant between shock increments (Kijima & Nakagawa, 1991). The sixth method assumes

that after each maintenance activity the remaining life probability distribution of the component will be reduced by some fraction of the remaining life distribution immediately before the maintenance activity (Wang & Pham, 1996).

The seventh method is simply a multivariate equivalent of the first method whereby multiple components of a system are modelled simultaneously. The eighth and final method discussed by Pham & Wang (1996) is a method proposed by Nakagawa & Yasui (1987) whereby the reduction in hazard rate after a maintenance activity is modelled as a function of the quantity of some resource consumed during the maintenance activity as well as an additional parameter.

These are classical maintenance models which continue to form part of the basis for ongoing research in the field of maintenance modelling and reliability engineering. Two more recent publications are:

- Lui et al. (2012), who considered goodness-of-fit tests to aid in determining the most accurate PM model using failure data.
- Kim et al. (2007), who used the improvement factor approach as described in the third method above to develop an adaptive PM policy using Bayesian inference.

Pham & Wang (1996) proceeded to identify 6 different preventive maintenance policy types in the literature namely:

- Age-dependant – the component is maintained either preventively at predetermined age intervals or correctively if it fails before reaching the predetermined age.
- Periodic – similar to the age-dependant policy except that the component is maintained at predetermined time intervals.
- Failure limit – the component is preventively maintained at a predetermined hazard rate.
- Sequential – in a sequential policy the component is maintained at predetermined intervals which reduce as the component ages.
- Repair limit – at failure the component is repaired if the repair cost is below a predetermined threshold value and replaced if it is above this value.
- Multicomponent.

In order to justify selection of one model or policy over another it is necessary to analyse failure data and conduct a failure modes and effects analysis. A failure mode is the event associated with the cause of a functional failure whereas a failure effect describes what happens when a failure mode occurs. Studying failure modes is justified as it enables development of efficient preventive maintenance through identification of root causes (Moubray, 1997). As an elementary example, head checking may be regarded as a failure mode whereas a derailment would be the associated failure effect.

2.5 STOCHASTIC MODELLING

Stochastic modelling of rail defects is not uncommon in the literature. Some basic principles thereof are introduced and discussed in this section. The use of the Weibull probability distribution for rail defect modelling is discussed with regard to different types of models present in the literature and their associated assumptions. Thereafter, the concept of a hazard function as well as a P-F interval are discussed.

2.5.1 The Weibull Distribution and Poisson Process for Rail Defect Modelling

Weibull methods used in the literature often differ with regard to their underlying assumptions and are thus not identical. The significance of this is that parameter values estimated for a particular Weibull method cannot be used as the parameter values for a different Weibull method due to their different underlying assumptions (Roth, 2008). The probability density function (PDF) and cumulative distribution function (CDF) for a generic 2-parameter Weibull probability distribution of a random variable (RV) X are shown in Equations 2.1 and 2.2 respectively (Kobayashi, Mark & Turin, 2012):

$$f_X(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha}, x > 0 \quad (2.1)$$

With:

- x = a specific value of the RV X
- α = shape parameter (dimensionless)
- β = scale parameter (same units as X)

$$F_X(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha}, x > 0 \quad (2.2)$$

Three of the four Weibull methods reviewed by Roth (2008) were concluded as being rational and providing a good fit to trial data from a heavy haul line. Only the three methods considered adequate by Roth (2008) are discussed further. Table 2.3 shows a summary of some of the limitations as well as references to the Weibull methods discussed further. It should be noted that Weibull Method 3 is actually proposed by Roth (2008) and is not a review of a method found in previous literature.

Table 2.3: Summary of Weibull methods used to model rail defects

<i>Weibull Method #</i>	<i>Number of defects limited</i>	<i>Location simulated</i>	<i>References</i>
1	Yes	Yes	<ul style="list-style-type: none"> • Besuner et al. (1978) • IHHA (2001) • Orringer et al. (1999)
2	No	No	<ul style="list-style-type: none"> • Chattopadhyay et al. (2005) • Chattopadhyay & Reddy (2007) • Zhao et al. (2007)
3	No	Yes	<ul style="list-style-type: none"> • Roth (2008)

Weibull Method 1 divides the rail into equally sized segments. Each segment can develop at most a single defect. The tonnage at which a specific rail segment will develop this defect is assumed to follow a Weibull probability distribution. As a result the state of each segment follows Boolean logic. Therefore, the total number of defects over the total length of rail considered follows a binomial distribution with the probability an event (defect or no defect) occurring specified by the Weibull distribution.

Weibull Method 2 assumes that defects arrive according to a non-homogenous Poisson process (NHPP) with an intensity function given by the hazard function of the Weibull probability distribution. Defects can be grouped into homogenous populations so that they can be modelled by a Weibull distribution with calculated parameter values. The hazard function for all defects can then be summed and used as the intensity function in the NHPP. The Poisson probability distribution is a discrete probability distribution which describes a counting process $N(t)$ for the number of events n which occur during a period t at an expected rate λ (also known as the intensity function). Arrivals following a NHPP may be described using a probability mass function (PMF) as given in Equation 2.3 with an intensity function given by Equation 2.4:

$$p_n(t) = P[N(t) = n] = \frac{\mu(t)^n}{n!} e^{-\mu(t)} \quad (2.3)$$

Where:

$$\mu(t) = \int_0^t \lambda(\tau) d\tau \quad (2.4)$$

With:

- $p_n(t)$ = the probability that n events occur during time t
 $\mu(t)$ = the expected number of occurrences during interval $(0, t)$
 $\lambda(\tau)$ = the hazard function for the Weibull probability distribution
 t = time or tonnage
 τ = dummy integration variable

Weibull Method 3 is proposed by Roth (2008) and is based on the continuous approximation of Weibull Method 1. Instead of dividing the rail into equal segments which limits the maximum number of defects, it is assumed that potential defects are space according to an exponential distribution. Thus, the number of potential defects along a length of rail follows a Poisson process. The potential defects then develop into actual defects at a tonnage characterised by a Weibull distribution. This model has not been applied in the literature and Roth (2008) suggests that this model requires further investigation.

2.5.2 The Hazard Function and Maintainable versus Non-Maintainable Failure Modes

The probability distribution of the failure time of a component from a homogeneous population may be described in three different ways namely the survivor function (SF), the PDF or the hazard function. The most useful of these is the hazard function as it is able to fully specify both the PDF and the SF (Kalbfleisch & Prentice, 2002). A hazard function $\lambda_T(t)$ (also commonly called a failure rate function or the force of mortality) is defined such that $\lambda_T(t) \cdot dt$ represents the probability that a component's life will end in the time interval $(t, t + dt]$ given that it has survived up to age t . Thus, the hazard function for a RV T with a PDF $f_T(t)$, CDF $F_T(t)$ and SF $S_T(t)$ is defined as (Kobayashi et al., 2012):

$$\lambda_T(t) = \frac{f_T(t)}{S_T(t)} = \frac{f_T(t)}{1 - F_T(t)} \quad (2.5)$$

Specifically, for the 2-parameter Weibull probability distribution shown in Equation 2.1, the corresponding hazard function and survivor function is given by Equations 2.6 and 2.7 respectively:

$$\lambda_T(t) = \frac{\alpha}{\beta^\alpha} t^{\alpha-1} \quad (2.6)$$

$$S_T(t) = e^{-\left(\frac{t}{\beta}\right)^\alpha} \quad (2.7)$$

With:

- α = the shape parameter (dimensionless)
 β = the scale parameter (same units as t)

It should be noted that the hazard function as shown in Equation 2.6 is synonymous with the power law process (Rigdon & Basu, 2000). Failure modes can be divided into those which are maintainable and those which are not. Maintainable failure modes will exhibit some change to their hazard function with maintenance. Non-maintainable failure modes are generally those which exhibit a random failure pattern displaying no identifiable trend with age or usage. This implies that the hazard function is constant. The hazard rate function of a non-maintainable failure mode will remain constant regardless of whether preventive maintenance is conducted or not (Endrenyi et al., 2001).

2.5.3 The P-F Interval

Failure modes can be separated into those which display an identifiable relationship between the probability of defect initiation occurring and the age or usage of the component and those which do not i.e. defects initiation occurs at random with no identifiable trend. For the former failure mode type the hazard function would be a strictly increasing function and for the latter failure mode type the hazard function would be constant (Endrenyi et al., 2001). Regardless of this relationship, failure modes are generally detectable before they result in functional failure. The time period between defect initiation and functional failure is referred to as the P-F interval (Moubray, 1997). A generic P-F interval is displayed in Figure 2.21. Point P in the diagram denotes the point at which an anomaly becomes detectable and is therefore subsequently termed a potential failure or defect and Point F denotes the point at which functional failure occurs, hence the name P-F interval. It is important to note that there will be a time period when the failure process has already initiated but is not yet detectable. This time period exists between point A and P in Figure 2.21.

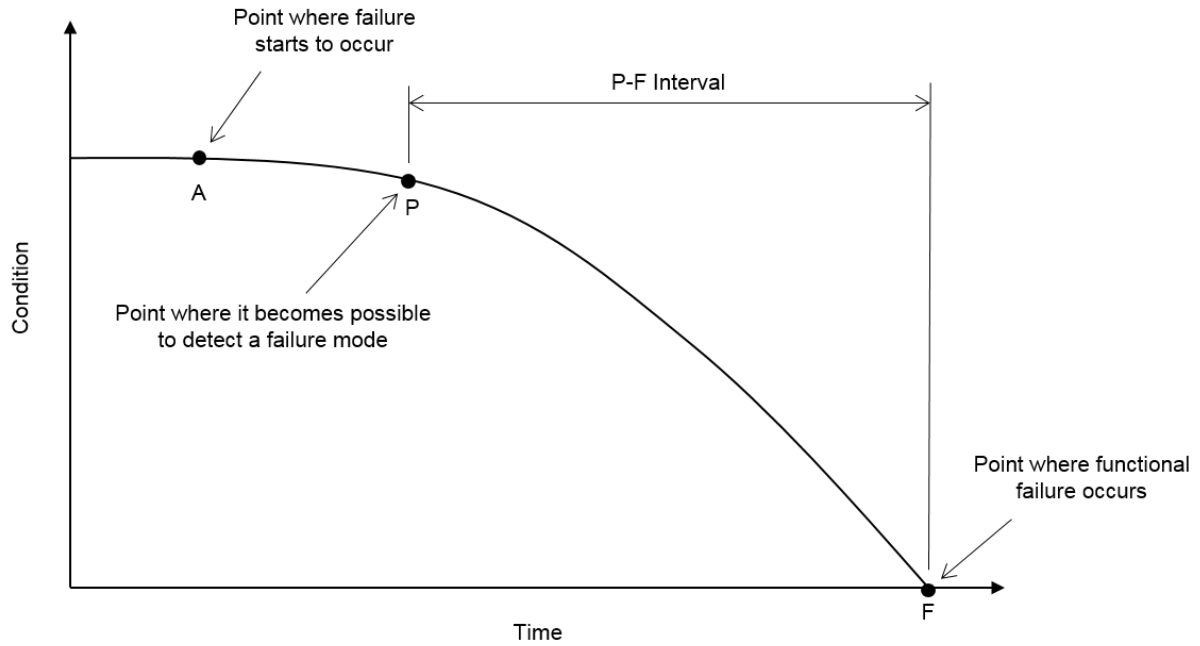


Figure 2.21: Illustration of the concept of a P-F interval for a deteriorating component (Moubray, 1997)

The P-F interval is a vital concept when considering condition-based maintenance strategies. The P-F interval length dictates the inspection interval length. If the inspection interval is longer than the P-F interval then there is a chance that the failing component will not be detected before functional failure occurs. Regardless of inspection interval length there also exists a chance that inspection will not reveal defects. Therefore, multiple inspections within the P-F interval may be warranted to increase the chance of detecting defects (Vatn et al., 2003). However, if the inspection interval is much shorter than the P-F interval then unnecessary inspection costs will result (Moubray, 1997). Therefore, a balance is required between the inspection interval length and the P-F interval length. Thus, the concept of the nett P-F interval is introduced.

The nett P-F interval is the time available after detection of a defect for maintenance or renewal works to be conducted which will permit mitigation or prevention of the failure effects associated with functional failure (Moubray, 1997). The length of the nett P-F interval is a function of the P-F interval itself as well as the inspection interval. Through the concept of a nett P-F interval it is conceivable that for certain failure modes condition-based maintenance may not be technically feasible. Moubray (1997) lists criteria to be used in order to determine whether condition-based maintenance is technically feasible:

- It must be possible to clearly identify a state of potential failure i.e. a defect.
- The P-F interval should be reasonably consistent for a specific failure mode.

- The condition of the component must be measurable at intervals which allow reasonable chance of detecting defects before functional failure occurs.
- The nett P-F interval must be sufficient so that maintenance or renewal works can be conducted before functional failure occurs.

The P-F interval for a failure mode can be lengthened through improved inspection and condition monitoring techniques which allow earlier detection of failing components or improved materials and manufacturing techniques. According to Moubray (1997) a lengthened P-F interval has two advantages namely increased possibility to avoid or mitigate failure effects and longer required inspection intervals both of which have the potential to reduce life cycle costs.

A popular approach in the literature is to model the P-F interval stochastically using a probability distribution. Vatn et al. (2003) model the P-F interval of rail cracks using a gamma distribution. In their study a single P-F interval probability distribution is used to model all rail cracks independent of failure mode. Furthermore, Vatn et al. (2003) note that the P-F interval for rail crack defects will most likely be dependent on other operational factors such as train speed, axle load, seasonal temperature variations, rail quality and superstructure quality but do not investigate the effects thereof (some of these factors were discussed in Section 2.2.3). Zhao et al. (2006) model P-F intervals using an exponential probability distribution. Zhao et al. (2006) expand on Vatn et al.'s (2003) approach by using different P-F interval probability distributions depending on the failure mode type. Both studies assume that the P-F interval is independent of defect initiation time.

2.6 LIFE CYCLE COST ANALYSIS AND UNCERTAINTY

Track M&R decisions should be made on the basis of their associated LCC. Andrade (2008) states that if the principle of life cycle cost analysis (LCCA) is applied to M&R decisions, then the long-term cost of ownership of railway infrastructure can be minimised. The principle of using LCCA in M&R decisions is reinforced by Patra, Söderholm & Kumar (2008) who state that LCCA allows the inclusion of operating costs, maintenance costs, energy costs and taxes in addition to the initial capital cost. The outcomes of decision making processes which include LCC as opposed to ones which only include capital costs may well differ as the operation and maintenance costs of railway infrastructure form a significant portion of the total LCC. Thus, LCCA may be used as a decision support system (DSS) for M&R actions (Zoeteman, 2001).

Zoeteman (2001) developed a framework for a DSS based on a LCC approach in order to combine the costs associated with construction, maintenance, financing and transport operations. The calculation

steps involved in the proposed LCC plan are summarised in Figure 2.22. A similar approach could be used when conducting a LCCA for a single infrastructure component such as the rail.

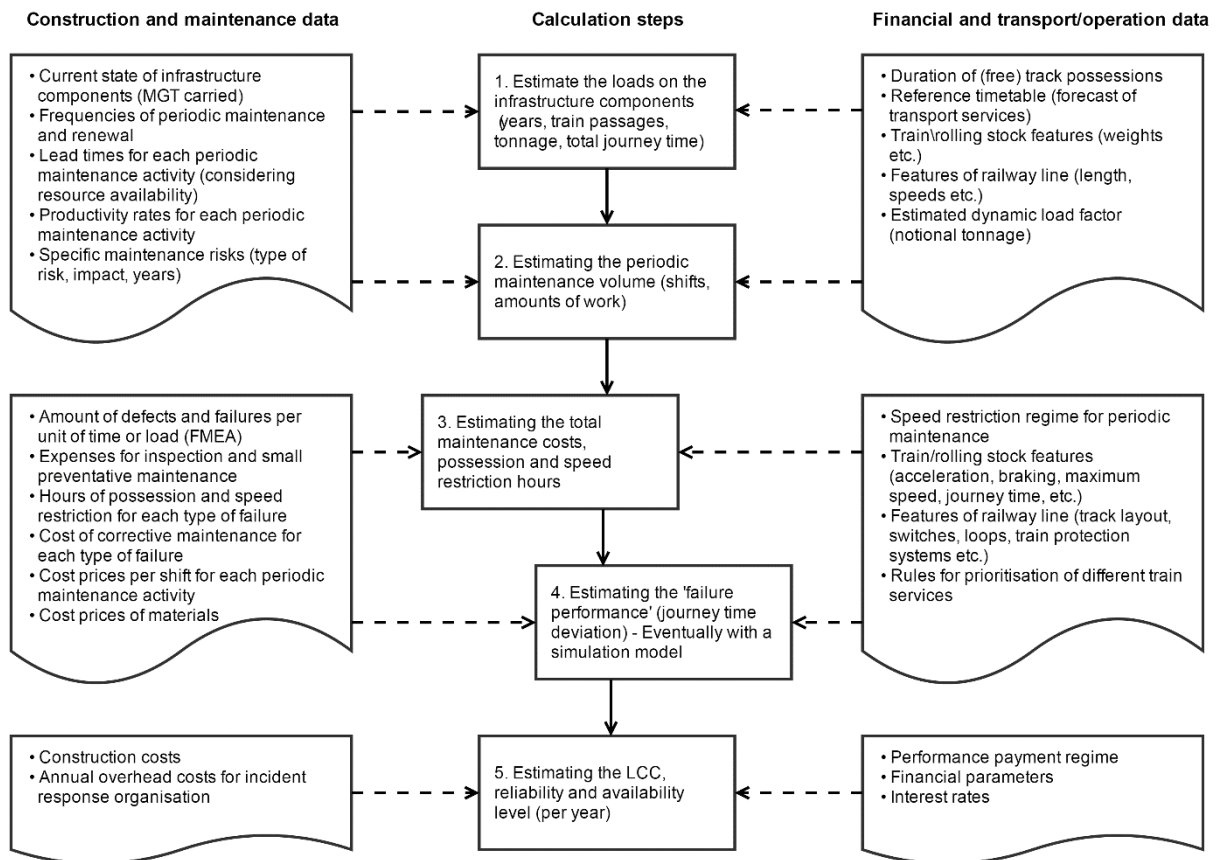


Figure 2.22: Calculation steps and inputs required to develop a LCC plan (redrawn from Zoeteman (2001))

For a DSS based on a LCCA to be sufficiently accurate, adequate data needs to be available linking the relationships between transport volumes, infrastructure quality, maintenance, availability and reliability (Esveld, 2001). Thus, a great deal of good quality data is required to conduct a LCCA. Zoeteman (2001) identifies four components which must be established in order to analyse the effectiveness of different M&R policies:

- Asset registration
- Identification of maintenance concepts
- Life cycle cost analysis
- Computer-assisted work planning

Nonetheless, the reliability of most data does infer certain uncertainties into the calculation of LCC. Thus, the robustness of a model should be tested. Zoeteman (2001) suggests the use of a sensitivity

analysis or an uncertainty analysis. Maintenance costs are considered the most complex costing component of a LCCA (Patra et al., 2009). Maintenance costs should include the following components within the maintenance cost itself:

- Cost of materials, equipment and labour.
- Cost of condition monitoring and inspections.
- Cost related to track possession time.

2.6.1 Uncertainty Estimation

Uncertainty estimation related to LCCA has been studied by Patra (2007) and Patra et al. (2009). Patra et al. (2009) identified two levels of uncertainty namely level 1 uncertainty and level 2 uncertainty. Level 1 uncertainty is associated with external risk as a result of costs associated with train delay and derailment. Level 2 uncertainty is defined as originating internally due to variable contributions of uncertainty in the reliability and availability parameters to the total LCC. These concepts are not unrelated. In order to assess uncertainty in LCCA Patra et al. (2009) used a combination of design of experiment principles and Monte Carlo simulation.

It is understood that the Bootstrap Method was used by Patra et al. (2009) in order to calculate confidence intervals for the reliability and maintainability parameters; namely: time to failure and time to repair. These point estimates were then used to infer corresponding point estimates for the LCC. The method of interval inference via point estimates is called the Point Estimation Method (PEM). Through the use of the above methods Patra et al. (2009) could minimise the number of simulations required to attain reliable results. However, this technique is only applicable to solutions where the LCC can be expressed as a closed-form solution.

2.7 REVIEW OF PREVIOUS MAINTENANCE MODELS

This section of the literature review describes the methodology, assumptions and limitations of some previous maintenance and renewal models identified in the literature.

2.7.1 Age-Usage Renewal Model

Shafiee et al. (2016) developed what they consider the first published hybrid age-usage maintenance model for railway infrastructure. In their research a cumulative damage shock model (see Qian et al. (2005)) was used to model the damage process of the railway track. The system is considered to be in a state of failure when the cumulative damage reaches a predetermined critical level. According to Shafiee

et al. (2016) the train movements are analogous to the shocks to the system and arrive according to a NHPP. The damage to the system is analogous to the gross tonnage borne by the rails. The magnitude of the damage associated with each train movement or shock is then proportional to the gross mass of the train in million gross tonnes (MGT). The variation in MGT associated with each train movement is modelled using an exponential distribution.

Maintenance then takes place according to the following scheme:

- If the cumulative usage reaches the predetermined failure usage-level of the system then corrective replacement is conducted.
- If the cumulative usage reaches a critical usage-level (less than the failure usage-level) then unplanned perfect PM is conducted. This critical usage-level is one of the decision variables to be optimised in order to minimise the expected long-run maintenance cost for an infinite time span.
- If the operational age of the system reaches a critical age-level before reaching the critical usage-level then planned PM is conducted. This critical age-level is also a decision variable which should be optimised.

The limitations of such a model is that the behaviour of the entire system is modelled in relation to the tonnage borne. This does not allow for individual consideration of different failure modes and different maintenance policies for the failure modes. No P-F interval is included and a failure cumulative tonnage has to be assumed. This failure tonnage is a vague concept.

2.7.2 Stochastic Rail Defect Model

The model discussed in this section is the model developed by Zhao et al. (2006). Zhao et al. (2006) studied the influence of inspection intervals on the LCC of rail. The LCC was calculated per km of rail and normalised against tonnage borne. The study was focused on the modelling of fatigue defects and ignored rail wear and corrugation. A schematic representation of the development and consequence of different types of rail failures is shown in Figure 2.23. From Figure 2.23 it can be seen that an interconnected relationship exists between maintenance activities and the occurrence of defects. For example, preventive rail grinding will influence rail wear and the occurrence of rail corrugations, however, it will also influence the occurrence of fatigue defects through truncation of RCF cracks as explained in Section 2.3.2. Rail grinding and inspections may prove beneficial to the maintenance of multitudinous defect types. However, the whole cost thereof is included in the maintenance of fatigue defects by Zhao et al. (2006). The LCC was minimised by finding an optimised inspection interval and

renewal tonnage for the rail. Important assumptions and components incorporated in the model may be summarised as follows:

- It is assumed that rail fatigue defects arrive according to a Weibull probability distribution.
- The inter-arrival tonnage for fatigue defects is independent and identically distributed.
- The P-F interval for each defect type is modelled using an exponential probability distribution.
- An increase in hazard rate for ATW defects was modelled in order to account for the fact that the number of ATWs in the rail is likely to increase with cumulative tonnage due to maintenance operations.
- Imperfect inspections were modelled using a constant probability of detection per defect type.
- The effect of preventive rail grinding practices was taken into account through a reduction of the associated RCF defect hazard rate.

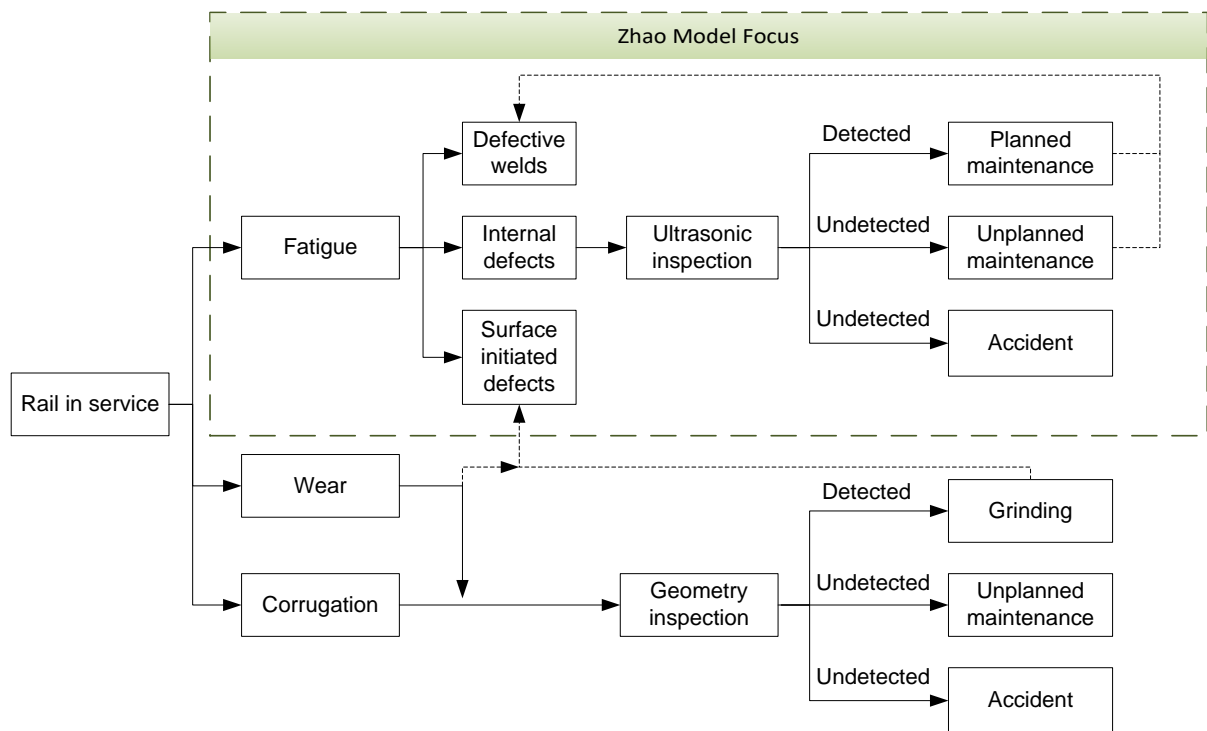


Figure 2.23: The development and consequence of rail failures (Zhao et al., 2006)

The maintenance costs modelled by Zhao et al. (2006) include:

- Rail renewal cost
- Planned maintenance cost
- Unplanned maintenance cost
- Derailment cost
- Ultrasonic inspection cost
- Rail grinding cost

Zhao et al. (2006) define a state before failure and refer to this as a potential failure or defect. A defect is considered to be detectable by modern ultrasonic inspection equipment thus enabling it to be repaired before a failure occurs. During the repair of defect in CWR, a new rail segment (called a closure rail) is typically installed and fixed into position using either ATWs or FBWs. Zhao et al. (2006) assumed that only ATWs are used for maintenance. Defects are categorised into two types by Zhao et al. (2006); namely: Type A and Type B defects.

Type A defects consist only of ATW defects. All other rail defects are classified as Type B defects. Type B defects include FBW defects, tache ovale defects, shell defects and squat defects. Removing a Type A defect will introduce 1 additional ATW weld into the rail as the existing ATW will be removed and 2 additional ATWs be introduced resulting in a net increase in ATW count of 1. It follows intuitively that if a Type B defect is removed, 2 additional ATWs will consequently be introduced into the rail system. It is important to distinguish between *additional* and *new* ATWs. Removal of both Type A and Type B defects will introduce 2 *new* ATWs into the system, but only Type B defects will introduce 2 *additional* ATWs into the system.

Under the assumptions stated above, Zhao et al. (2006) derived Equation 2.8 for the hazard rate of Type A defects taking into account the increase in ATWs due to maintenance activities:

$$v_a(t) = \lambda_a(t) \cdot e^{\int_0^t \lambda_a(\tau) d\tau} \left[n_0 + 2 \int_0^t v_b(\tau) \cdot e^{-\int_0^\tau \lambda_a(x) dx} d\tau \right] \quad (2.8)$$

Where:

$$v_b(t) = \sum_{j=1}^{h_j} \lambda_{b_j}(t) \quad (2.9)$$

$$\lambda(t) = \frac{\alpha}{\beta^\alpha} t^{\alpha-1} \quad (2.10)$$

With:

$\lambda_a(t)$ = hazard function of a single Type A defect at tonnage t

$\lambda_b(t)$ = hazard function of a single Type B defect at tonnage t

$v_a(t)$ = hazard function of Type A defects taking into account the increase in ATWs at tonnage t

$v_b(t)$ = hazard function of all Type B defects, which is simply the sum of individual hazard functions of all individual Type B defects

n_0 = initial number of ATWs in the system at $t = 0$

- h_j = the number of Type B defects modelled
 α = Weibull shape parameter for the specific defect under consideration
 β = Weibull scale parameter for the specific defect under consideration (MGT)

It is important to note that the hazard function of a single Type A defect $\lambda_a(t)$ must be specified with α and β parameters such that it is representative of a single ATW. Contrary, the hazard function of a Type B defect $\lambda_b(t)$ should be specified with α and β parameters such that it is representative of the entire km section of rail modelled. This is because the hazard function of Type A defects is assumed to be proportional to the number of ATWs whereas the hazard function of a Type B defect is assumed to be proportional to the length of rail modelled.

Subsequently, Zhao et al. (2006) derived Equation 2.11 for the expected number of failures as a function of renewal tonnage $N_f(T_R)$. Equation 2.11 incorporates imperfect inspections and an exponentially distributed P-F interval for each defect type:

$$N_f(T_R) = \sum_{j=1}^h \sum_{i=1}^{m+1} \sum_{k=1}^i \left\{ (1 - \eta_j)^{i-k} \int_{t_{k-1}}^{t_k} v_j(\tau) [G_j(t_i - \tau) - G_j(t_{i-1} - \tau)] d\tau \right\} \quad (2.11)$$

Where:

$$G_j(x) = 1 - e^{-\left(\frac{x}{\mu_j}\right)} \quad (2.12)$$

With:

- h = total number of defect types modelled (Type A and Type B)
 m = number of ultrasonic inspections conducted during the interval $(0, T)$
 η_j = the probability that a defect of type j will be detected by ultrasonic inspection
 t_k = tonnage at which the k^{th} inspection occurred
 t_i = tonnage at which the i^{th} inspection occurred
 $v_j(\tau)$ = the hazard function for defect type j which is representative of the entire km section modelled
 μ_j = the expected P-F interval length for defect type j

Once the expected number of failures $N_f(T_R)$ has been calculated, the expected number of defects detected by inspections $N_d(T_R)$ can be calculated using Equation 2.13:

$$N_d(T_R) = \sum_{j=1}^h \left[\int_0^T v_j(\tau) d\tau \right] - N_f(T_R) \quad (2.13)$$

Where:

h = total number of modelled defect types

$v_j(\tau)$ = the hazard function for the j th defect type

$N_f(T_R)$ = expected number of failures for a life cycle with a renewal tonnage T_R

The LCC as a function of renewal tonnage $LCC(T_R)$ is subsequently calculated using Equation 2.14:

$$LCC(T_R) = \left\{ c_R + \frac{c_I T_R}{s_I} + \frac{c_g T_R}{s_g} + [(1 - \xi)c_f + \xi c_d] N_f(T_R) + c_p N_d(T_R) \right\} \frac{1}{T_R} \quad (2.14)$$

Where:

c_R = cost of rail renewal (£/km)

c_I = cost of ultrasonic inspection (£/km)

s_I = interval of inspection (MGT)

c_g = cost of rail grinding (£/km)

s_g = interval of rail grinding (MGT)

ξ = probability that a rail failure will cause a derailment

c_f = cost of unplanned maintenance (£/repair)

c_d = cost of a derailment (£/derailment)

c_p = cost of planned maintenance (£/repair)

Zhao et al. (2006) used an exponential distribution to define a specific P-F interval for each defect type; namely: ATW, FBW, tache ovale and squat defects. Zhao et al. (2006) based the μ_j parameter for the exponential distribution on research by Jeong et al. (1997) and Vatn et al. (2003). The study by Jeong et al. (1997) was discussed in Section 2.2.3.

The LCC as calculated by Zhao et al. (2006) is shown in Figure 2.24. The stochastic parameters used are shown in Table 2.4 and the costs as well as the non-stochastic parameters are shown in Table 2.5. The LCC can be split into a renewal cost component and a maintenance cost component. The maintenance cost component comprises the grinding cost, inspection cost, planned maintenance cost, unplanned maintenance cost and derailment cost. The renewal cost decreases with an increase in renewal tonnage T_R whereas the maintenance cost increases with an increase in renewal tonnage T_R . These two costs oppose each other and the optimum tonnage for renewal occurs at the optimum trade-off between these two costs. The optimum LCC in Figure 2.24 occurs at approximately 308 MGT and has a value of £791.8/MGT/km.

Table 2.4: Stochastic parameters used by Zhao et al. (2006)

<i>Defect</i>	α^1	β^2	μ^3	η^4	ξ^5
ATW defects	1.01	315.8	10	0.7	0.00056
FBW defects	2.00	286.6	10	0.7	
Tache ovale defects	2.55	191.8	5	0.6	
Squat defects	2.17	182.3	7	0.7	

¹ Weibull shape parameter

² Weibull scale parameter

³ Expected P-F interval length

⁴ Probability of detection by ultrasonic inspection

⁵ Probability that a rail failure will cause a derailment

Table 2.5: Costs and non-stochastic parameters used by Zhao et al. (2006)

Renewal cost, c_R (£/km)	87000
Cost of a functional failure, c_f (£/failure)	4850
Cost of repairing a detected defect, c_d (£/defect)	680
Cost of a single ultrasonic inspection, c_I (£/km)	100
Cost of a derailment, c_x (£/derailment)	2720000
Cost of rail grinding, c_g (£/km)	1860
Initial number of ATWs, n_0	22
Hazard rate reduction factor for grinding, γ	0.4
Interval of grinding, s_g (MGT)	10
Interval of inspection, s_I (MGT)	2.5

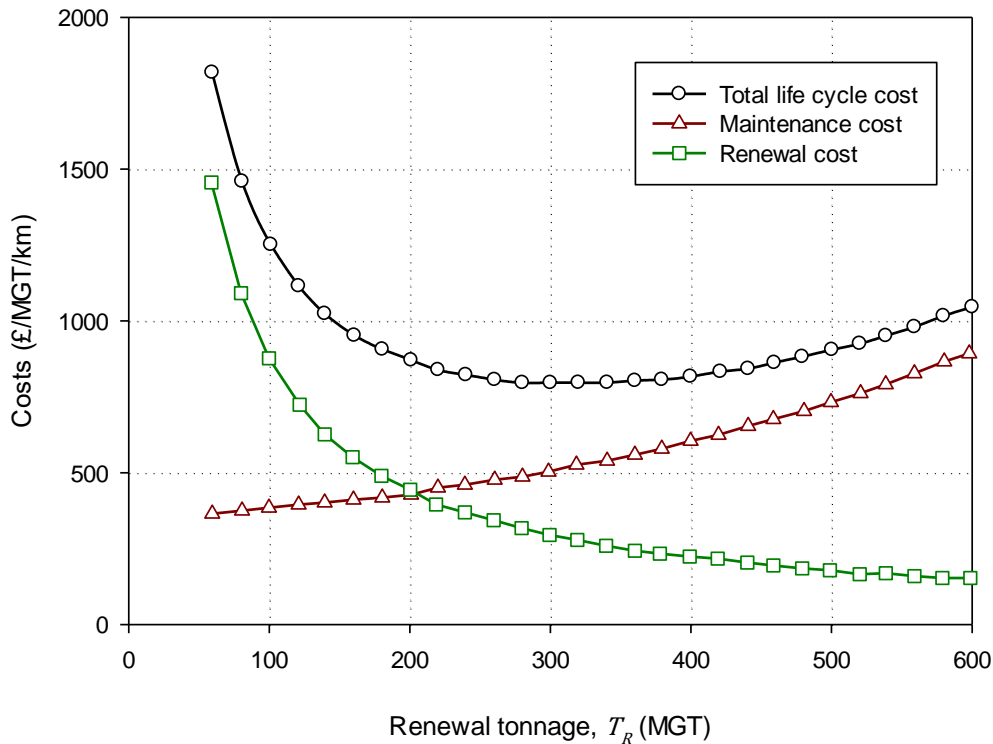


Figure 2.24: LCC expressed as the sum of maintenance cost and renewal cost (redrawn from Zhao et al. (2006))

Figure 2.25 illustrates the effect of varying the inspection interval length s_I on the LCC curve (Zhao et al., 2006). The LCC curve remains steady near its minimum over a larger range of renewal tonnages when the inspection interval length is reduced. This allows a larger window of opportunity to renew the rail whilst still achieving a desirable LCC. Figure 2.26 illustrates the effect of varying the inspection interval length on the minimum attainable LCC. From Figure 2.26 it can be seen that reducing the interval of inspection beyond a critical point will cause an increase in the minimum attainable LCC. The optimum interval of inspection for the case considered by Zhao et al. (2006) is approximately 1.1 MGT.

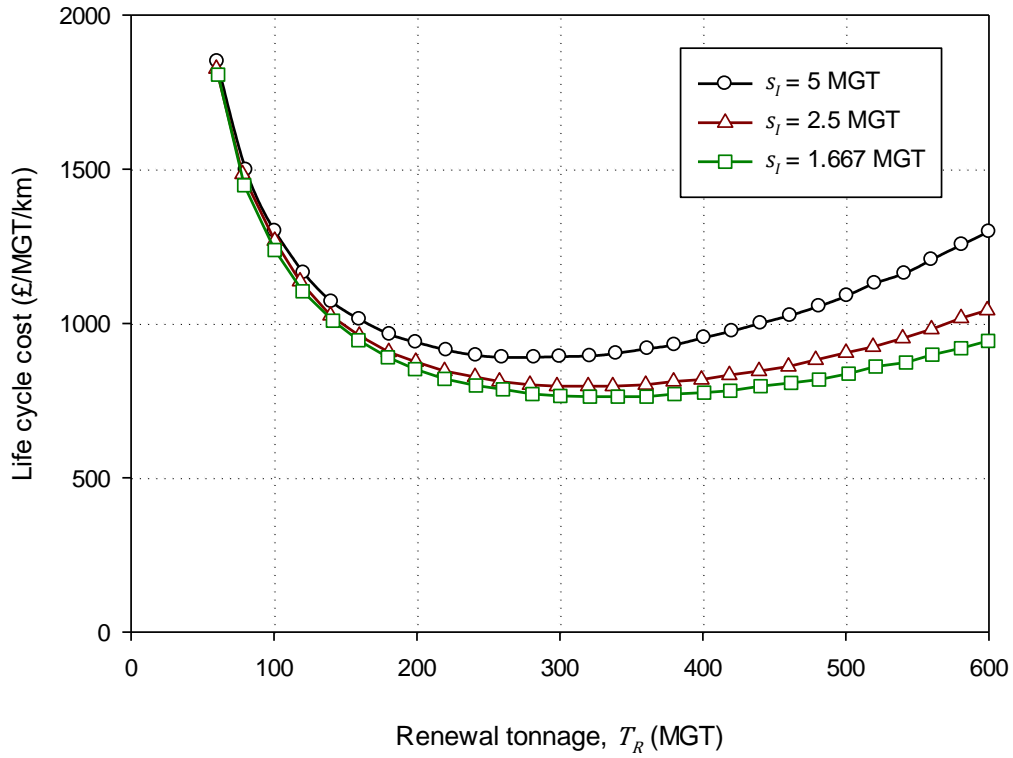


Figure 2.25: Influence of inspection interval length s_I on the LCC curve (redrawn from Zhao et al. (2006))

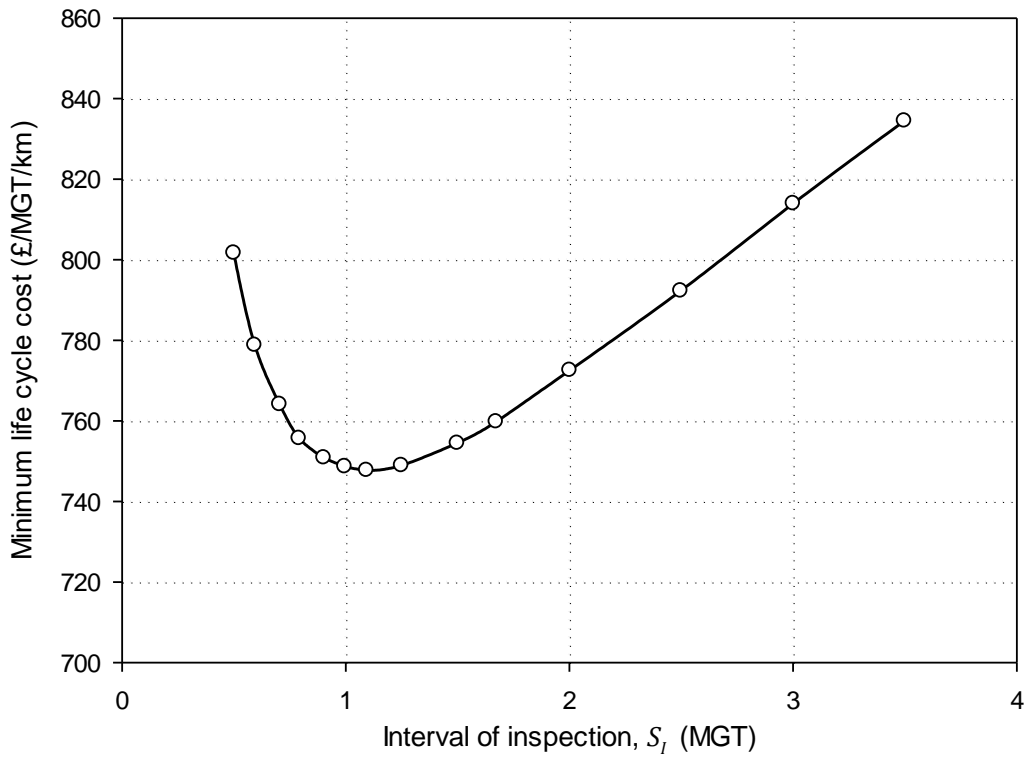


Figure 2.26: Influence of inspection interval length s_I on the minimum attainable LCC (redrawn from Zhao et al. (2006))

The model developed by Zhao et al. (2006) was modified and used in another paper by Zhao et al. (2007). The new model developed by Zhao et al. (2007) varies in the following ways from the model developed by Zhao et al. (2006):

- A probability parameter was introduced to account for the possibility that a maintenance action may replace an existing adjacent ATW.
- The hazard function of ATWs was divided into two separate functions. The first function accounted for early defects and the second for fatigue type defects.
- Values for the hazard rate reduction factor for grinding γ were related to specific grinding rates in mm/year.

2.8 SUMMARY

Section 2.2 shows that the rail deterioration process can be divided into fatigue defects, wear and corrugations. The focus of this study is on fatigue defects only. Rail fatigue defects can be categorised according to their longitudinal location on the rail (i.e. at joints or on plain rail sections) and according to whether or not their hazard rate is effected by rail grinding. A short review of an analytical study by Jeong et al. (1997) showed that the progression of a transverse rail defect into a functional failure is influenced by the location, orientation, size, vertical-lateral load combination, rail curvature, foundation stiffness, rail neutral temperature and fracture toughness of the rail. Thermal stresses further affect the rate of defect propagation during loading cycles. Jablonski et al. (1990) used a field study to show that the loading stress spectra of various train consists also affect rail fatigue defect propagation. Due to these numerous influencing factors, modelling of rail fatigue defects for LCCA lends itself to a stochastic approach.

Section 2.3 shows that inspection practices are not perfect and that a maintenance model should account for the probability that a defect is not detected by ultrasonic inspection. Furthermore, studies were cited indicating that RCF defects are mitigated through preventive rail grinding. Thus, a maintenance model should include some reduction in the hazard rate of RCF defects to model the influence of preventive rail grinding.

A background to maintenance modelling concepts was provided in Section 2.4 followed by a mathematical introduction to stochastic maintenance modelling in Section 2.5. In Section 2.6 it was argued that LCCA is an important DSS for M&R decisions. Furthermore, it is important that LCCA takes into account the inherent uncertainties related to reliability and maintainability parameters. The literature study was concluded with a review of two maintenance models specific to rail maintenance developed by Shafiee et al. (2016) and Zhao et al. (2006) respectively.

Shafiee et al. (2016) used a cumulative damage shock model which requires an arbitrary predetermined critical cumulative damage level to be defined at which the system is assumed to have failed. The damage shock model considers the entire rail as a single system and does not provide any insight into specific defects and underlying M&R parameters which can influence the LCC. Furthermore, inspections are not explicitly taken into account by the damage shock model. No explicit derailment probability has been included by Shafiee et al. (2016).

Zhao et al. (2006) explicitly model the effect of inspections on the LCC. Furthermore, a probability for derailment for undetected failures is also included. However, due to the closed-form nature of the LCC model derived by Zhao et al. (2006), certain assumptions are made which may inadvertently influence the final LCC calculated. Some of these assumptions are:

- All closure rail is installed using ATWs. The derivation of the model does not allow the influence of using FBWs instead of ATWs for maintenance on the LCC to be quantified.
- The hazard rate of “new” ATWs is not reset i.e. all maintenance is minimal maintenance. The author is of the opinion that this may not be representative of reality.
- FBWs are modelled with a hazard rate proportional to the length of the rail (they are defined as Type B defects; see Section 2.7.2). This is counter intuitive as the hazard rate of ATWs is modelled in proportion to the number of welds present in the modelled section.
- Due to the fact that FBWs are modelled as Type B defects by Zhao et al. (2006) their hazard rate does not reduce over time despite the fact that FBWs will be maintained and replaced by ATWs.
- The probability of a derailment occurring due to an undetected failures is constant across all defect types.
- The hazard function of Type A defects $v_a(t)$ assumes that new welds are installed immediately when a defect initiates. Therefore, the P-F interval is essentially ignored and the hazard rate is essentially larger than it should be when the weld is actually installed. This could have a progressive effect with an increase in tonnage.

The results of both the model by Shafiee et al. (2016) and Zhao et al. (2006) do not explicitly indicate the effect of any uncertainty related to the input parameters on the estimated LCC. The outputs of both these models are point estimates of a fairly complicated underlying process. A rail maintenance model is presented in Section 3 which aims to overcome the stated limitations of the models discussed.

3 NUMERICAL MODEL SETUP

This section describes the operational logic and setup of the developed numerical model. The stochastic simulation model was developed based on the Monte Carlo method. In short, Monte Carlo simulation is the name given to any simulation which involves the use of random or pseudorandom numbers to introduce randomness in the underlying model (Rubinstein & Kroese, 2008). The model was programmed in MATLAB. All modelling was done for a single representative kilometre of track. It should be noted that the terms tonnage and time are used interchangeably in this section but should be understood to both mean tonnage. This is because the rail deterioration process due to fatigue is directly related to the tonnage borne by the rail and not the time which has passed. This study considers only the uncertainty related to technical parameters and not economical parameters such as the discount rate and inflation. Thence, no discount rate has been used in the model. This is common practice and the same assumption has been made by Zhao et al. (2006) and Shafiee et al. (2016).

3.1 INTRODUCTION TO THE NUMERICAL MODEL

Two unique types of rail fatigue defects are identified with regard to the manner in which they are modelled, namely: Category A and Category B defects. Category A defects are related to weld quality whereas Category B defects are related to longitudinal rail quality. No bolted joints are considered in the model. The hazard rate of Category A defects $\lambda_A(t)$ must be specified for a single weld whereas the hazard rate of Category B defects $\lambda_B(t)$ must be specified in proportion to the length of rail modelled. Category A defects are sub-divided into alumino-thermic weld (ATW) defects and flash butt weld (FBW) defects only. However, Category B defects are sub-divided into defects such as tache ovale defects, squat defects, shelling, spalling and head checking. The number of Category B defect types which can be modelled is unlimited. The reason for defining these two categories of defects is in order distinguish between the manner in which their hazard functions are specified. The hazard functions of Category A defects are specified per defect whereas the hazard functions of Category B defects are specified per unit length of rail. The basic assumptions with regard to the modelling procedure may be summarised as follows:

- Rail defect arrival is governed by a 2-parameter Weibull distribution as found by Zarembski (1991) and which has been used in numerous studies as highlighted by Roth (2008). These two parameters are α (the shape parameter) and β (the scale parameter).
- The defect inter-arrival tonnages t_d for a specific defect type are independent and identically distributed and hence follow a non-stationary exponential probability distribution with a mean value function specified by the hazard function of the Weibull distribution. This is also known as the power law process.

- The P-F interval t_{P-F} follows an exponential distribution as assumed by Zhao et al. (2006).
- A defect of type j is considered detectable with probability η_j upon ultrasonic inspection. This accounts for imperfect inspections.
- The hazard rate of rolling contact fatigue (RCF) defects is reduced by a factor $\gamma(q_g)$ which is dependent on the annual grinding rate q_g in mm/year. This is in order to model the effect of preventive rail grinding practices. This approach is suggested by Zhao et al. (2007).
- The maintenance of defects can be divided into 3 categories with regard to the state in which the defect or weld was found at the time of maintenance. These 3 categories are planned maintenance, unplanned maintenance and renewal maintenance. Planned maintenance is conducted when the defect is detected before failure occurs i.e. during the P-F interval of the defect. Unplanned maintenance is conducted when a failure occurs. Renewal maintenance is conducted when a weld or defect is removed due to renewal of the rail at the end of its life cycle.
- All maintenance actions are assumed to take place without delay.
- The effect of maintenance on the hazard rate of the system can be modelled using either minimal maintenance (no effect on the hazard rate) or perfect maintenance (hazard rate decreases to same value as when new) as defined in Table 2.2 in Section 2.4.3.
- The joining of the closure rail during maintenance work may be done using either aluminothermic welding or mobile flash butt welding technology.

A single simulation produces N_{sim} virtual life cycles using unique values for the inter-arrival tonnage t_d and P-F interval length t_{P-F} for each defect. These unique values are attained through Monte Carlo sampling and the Inverse Transform Method as explained in Section 3.2. The model requires that the initial number of ATWs $n_{0_{ATW}}$ and initial number of FBWs $n_{0_{FBWs}}$ be specified. This initial number of welds in combination with the hazard function of all Category A and Category B defects represents the condition of the rail as installed at tonnage $T = 0$ MGT. It should be noted that tonnage relative to the life cycle being simulated is denoted by a capital letter 'T' and tonnage relative to the defect or weld under consideration is denoted by a small letter 't'.

Calculation points need to be specified for the model. These calculation points are specified in the form of an array of increasing renewal tonnages \vec{T}_R . A life cycle cost (LCC) is calculated for each renewal tonnage in array \vec{T}_R . The calculation procedure is described in detail using flow charts in Section 3.3.

3.2 INVERSION AND THE GENERATION OF RANDOM VARIABLES

In the stochastic simulation model developed, the Inverse Transform Method is used to generate random variables according to specified probability distributions. The Inverse Transform Method is described in detail by Robert & Casella (2004). In short, the Inverse Transform Method states that if a random variable U follows a uniform distribution on the open interval $(0,1)$ then the inverse of the cumulative distribution function F , $F^{-1}(U)$ has the probability distribution F . This assumes that F is a non-decreasing function. The Inverse Transform Method is illustrated schematically in Figure 3.1 for a single variate u . The Inverse Transform Method is used to derive expressions describing the arrival behaviour of defects and the length of the P-F interval in Section 3.2.1 and 3.2.2 respectively.

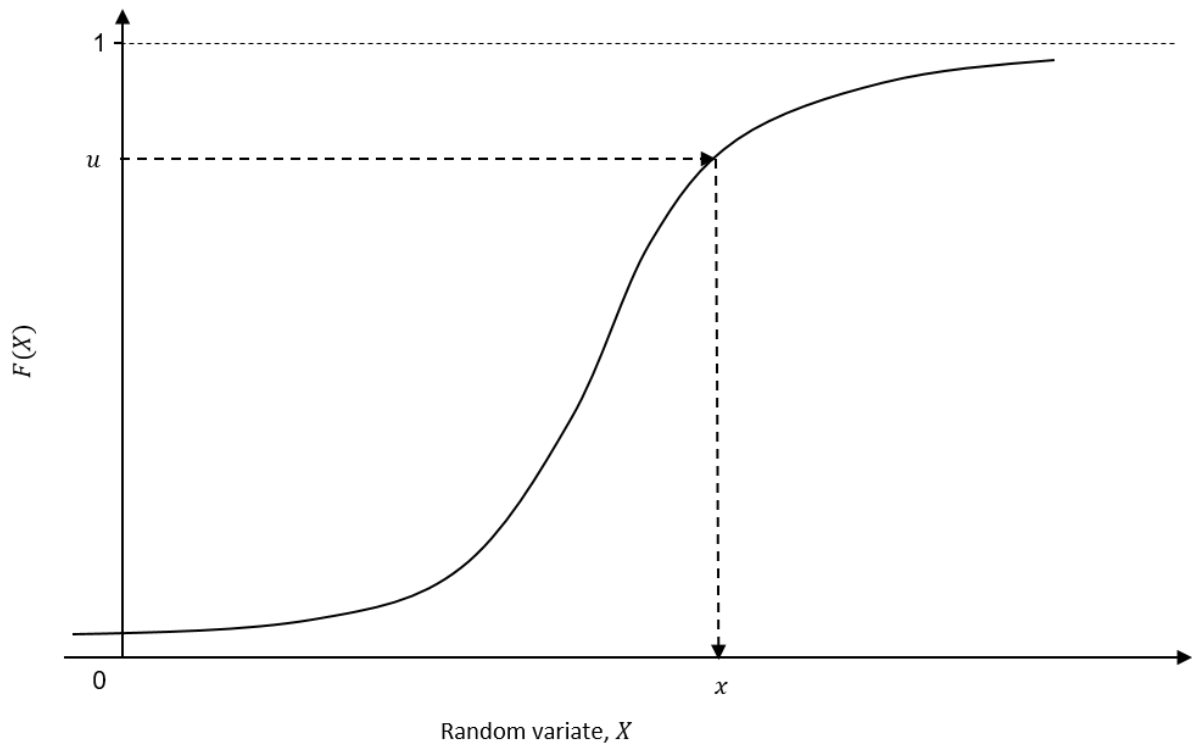


Figure 3.1: The Inverse Transform Method for generating a random variable (redrawn from Rubinstein & Kroese (2008))

It is important that the pseudorandom numbers generated in the Monte Carlo simulation do not have a significant correlation which would inadvertently influence the output LCC distribution. It may be warranted in certain cases to generate pseudorandom numbers with a specified correlation in order to appropriately model a phenomenon which displays correlation in the physical world. However, in the current model developed no specific correlation was assigned to the pseudorandom numbers as no evidence exists for a relationship between the inter-arrival tonnage of defects t_d , the length of the P-F interval t_{P-F} and the probability of detection η .

Latin hypercube samples were used in order to control pseudorandom number generation and ensure repeatability for the model in MATLAB. Latin hypercube sampling uses a stratified method to sample pseudorandom numbers. Therefore, a more uniform distribution of numbers is attainable in comparison to a congruential generator which means that fewer simulations need to be conducted. A built-in function in MATLAB allows the user to reduce the correlation of the numbers within the Latin hypercube sample by generating consecutive Latin hypercube samples and then selecting the sample with the smallest correlation coefficient ρ .

Latin hypercube samples were generated of adequate length and with sufficient dimensions to account for each defect type which had to be modelled. The random numbers were then sampled consecutively from the Latin hypercube sample. Figure 3.2 shows the evolution of the Pearson's linear correlation coefficient ρ for a 4-dimensional Latin hypercube sample. The black dashed lines in Figure 3.2 indicate correlation coefficients of -2 % and 2 % respectively. The correlation is relatively high for pseudorandom samples of length less than 300 numbers but stabilises and remains less than $|0.02|$ for pseudorandom samples of length greater than 20 000 numbers.

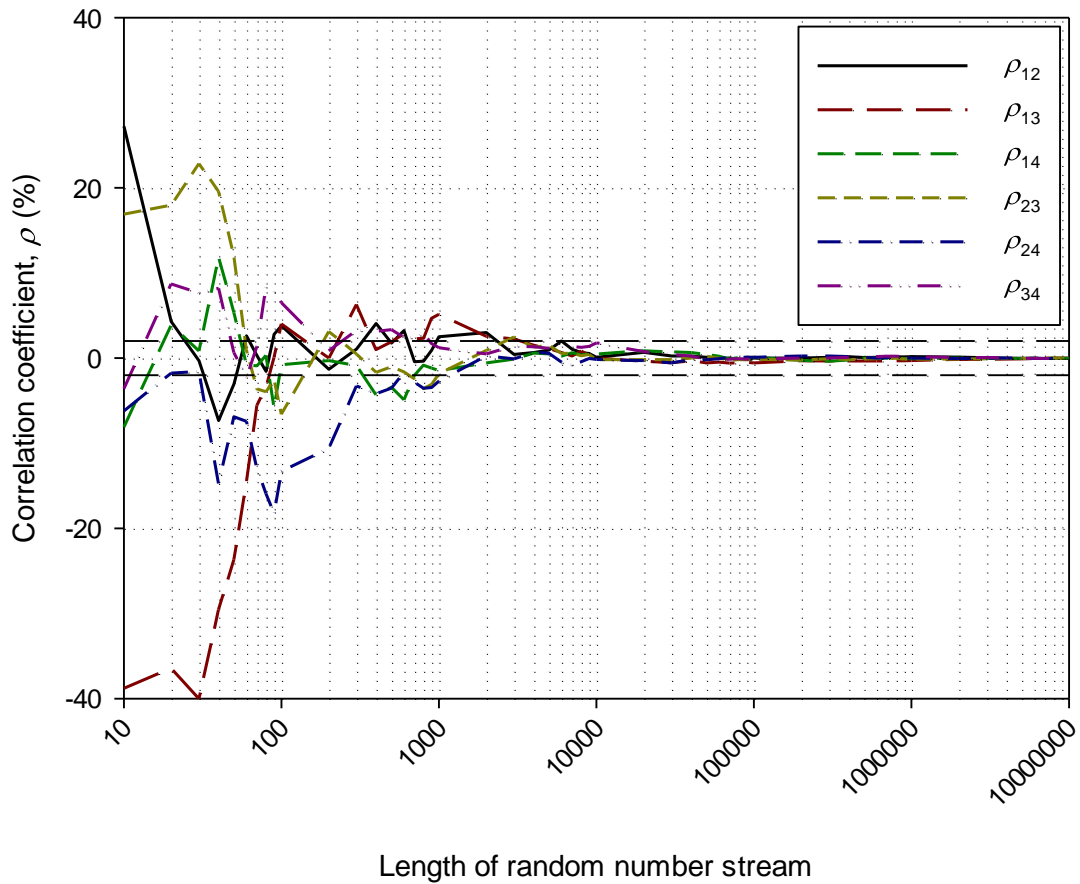


Figure 3.2: Correlation of a 4-dimensional Latin hypercube sample versus the length of the pseudorandom sample

The 1st dimension of the pseudorandom sample is used for the generation of the inter-arrival tonnage of defects t_d for ATWs. The 2nd is used for FBWs and so on. A similar Latin hypercube sample was generated for use during the Inverse Transform Method for the P-F interval length for each defect type modelled.

3.2.1 Modelling the Arrival of Defects

It is assumed that the arrival of rail defects is adequately modelled by a 2-parameter Weibull probability distribution and that the inter-arrival tonnages t_d are independent and identically distributed. Thus, it follows that the inter-arrival tonnages of rail defects t_d can be modelled using an exponential distribution with a mean value function $\mu(t)$ given by the integral of the hazard function of the Weibull distribution. A probability distribution of this nature has a cumulative distribution function (CDF) as shown in Equation 3.1:

$$F(t) = 1 - e^{-\mu(t)} \quad (3.1)$$

Where:

$$\mu(t) = \int_0^t \lambda(\tau) d\tau \quad (3.2)$$

$$\lambda(\tau) = \frac{\alpha}{\beta^\alpha} \tau^{\alpha-1} \quad (3.3)$$

With:

- t = the inter-arrival tonnage of defects (MGT)
- $\mu(t)$ = the expected number of arrivals at tonnage t
- $\lambda(\tau)$ = hazard function of the Weibull distribution at tonnage τ
- α = the Weibull shape parameter
- β = the Weibull scale parameter (MGT)
- τ = a dummy integration variable

Equation 3.3 needs to be modified in order to model both minimal maintenance and perfect maintenance. The variable τ is replaced with the expression $\tau + a$, where a represents the horizontal shift in the hazard rate for minimal maintenance. The variable a is analogous to the life of the defect which has essentially been lost through minimal maintenance. Furthermore, a grinding reduction factor $\gamma(q_g)$ is introduced into Equation 3.3 to represent a reduction in the hazard rate of defects due to preventive grinding maintenance. Thus, Equation 3.3 becomes:

$$\lambda(t) = (1 - \gamma(q_g)) \cdot \frac{\alpha}{\beta^\alpha} (\tau - a)^{\alpha-1} \quad (3.4)$$

With:

- a = the model tonnage at which the weld was installed
- $\gamma(q_g)$ = the grinding reduction factor which is a function of the grinding rate q_g in mm/year

Substituting Equation 3.4 into Equation 3.2 and subsequently Equation 3.1 and solving for t_d gives Equation 3.5:

$$t_d(U_1) = \left[a^\alpha - \frac{\beta^\alpha \ln(1 - U_1)}{1 - \gamma(q_g)} \right]^{\frac{1}{\alpha}} - a \quad (3.5)$$

With:

U_1 = a uniformly distributed variable sampled using Monte Carlo simulation

Equation 3.5 is used to calculate the defect inter-arrival tonnages t_d . U_1 is sampled using Monte Carlo simulation from the appropriate Latin hypercube sample. This produces values of t_d distributed according to the desired non-stationary exponential distribution.

Figure 3.3 illustrates the behaviour of the hazard function $\lambda(t)$ shown in Equation 3.4. For all the curves in Figure 3.3, $\alpha = 1.5$ and $\beta = 200$ MGT. The solid curve illustrates the case when $a = 0$ and $\gamma(q_g) = 0$. This case is associated with Category A defects for welds installed during maintenance activities which are modelled using perfect maintenance. The dashed curve illustrates the hazard rate curve when $a = 200$ MGT and $\gamma(q_g) = 0$. This case is associated with the modelling of all Category B defects as well as Category A defects which are modelled using minimal maintenance. The dash-dot curve in Figure 3.3 demonstrates the effect of the hazard rate reduction factor for modelling the influence of rail grinding. For this case $\gamma(q_g) = 0.4$.

Figure 3.4 shows plots of the CDF in Equation 3.1 for each respective hazard function shown in Figure 3.3. The Inverse Transform Method is applied to these CDFs in order to determine the defect inter-arrival tonnage t_d .

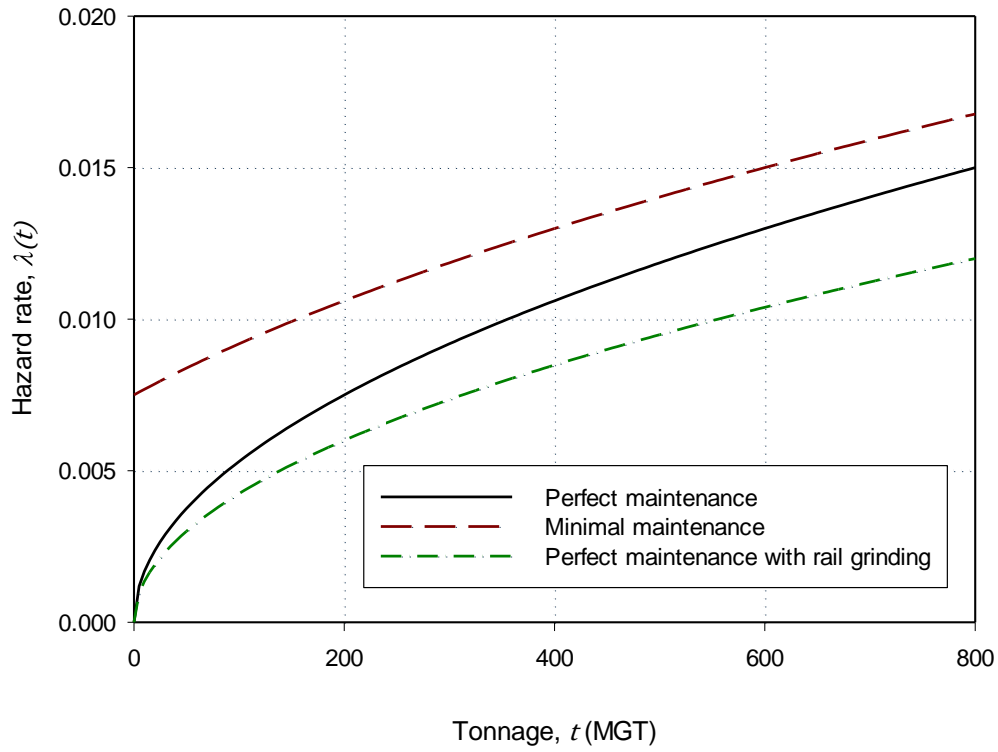


Figure 3.3: Comparison of the hazard rate for a defect modelled using perfect maintenance, minimal maintenance and perfect maintenance with a grinding reduction factor

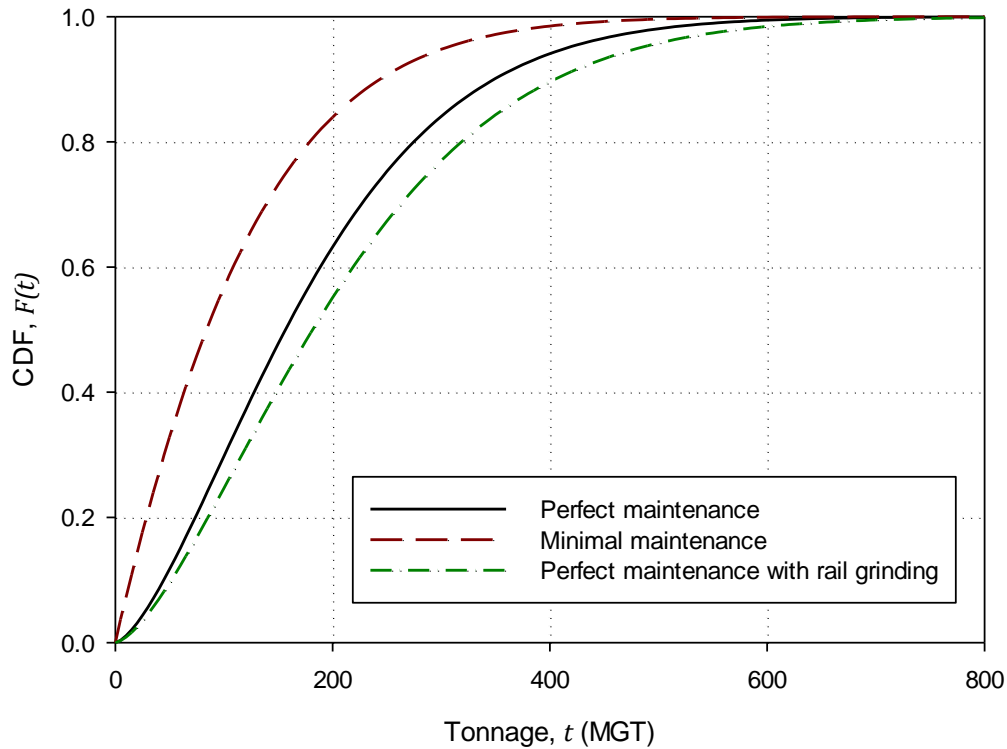


Figure 3.4: CDF for the defect inter-arrival tonnage t_d for a defect modelled using perfect maintenance, minimal maintenance and perfect maintenance with a grinding reduction factor

Figure 3.5 is a schematic representation of the behaviour of the total system hazard rate $\nu(T)$ as influenced by different maintenance modelling assumptions. The figure represents a system with a single weld at tonnage $T = 0$ MGT. The system hazard rate then increases according to the hazard function for that single weld $\lambda(t)$. At a tonnage t_d determined through Monte Carlo sampling the weld will birth a defect. It is important to note that the hazard function plays no part in determining the tonnage between the point where the defect initiates t_d and the point where it is maintained T_m . This is fully governed by the length of the P-F interval t_{p-f} as well as the number of inspections conducted during the P-F interval in combination with the probability of detection for that defect η_j . This will be illustrated clearly in the flow charts in Section 3.3.

Three different curves are shown in Figure 3.5. The solid line is the system hazard rate $\nu(T)$ before maintenance occurs and is equal to the weld hazard rate $\lambda(t)$. Two different scenarios are possible after the maintenance action has occurred at $T = 200$ MGT. If maintenance is modelled using minimal maintenance then the system hazard rate $\nu(T)$ will follow the dash-dot curve. This hazard rate will be equal to $2\lambda(t)$ with $a = 200$. If the maintenance is modelled using perfect maintenance then the system hazard rate $\nu(T)$ will follow the dashed curve. This hazard rate will be equal to $2\lambda(t)$ with $a = 0$. The reason that the hazard rate is multiplied by two is because two new welds are introduced by the maintenance activity.

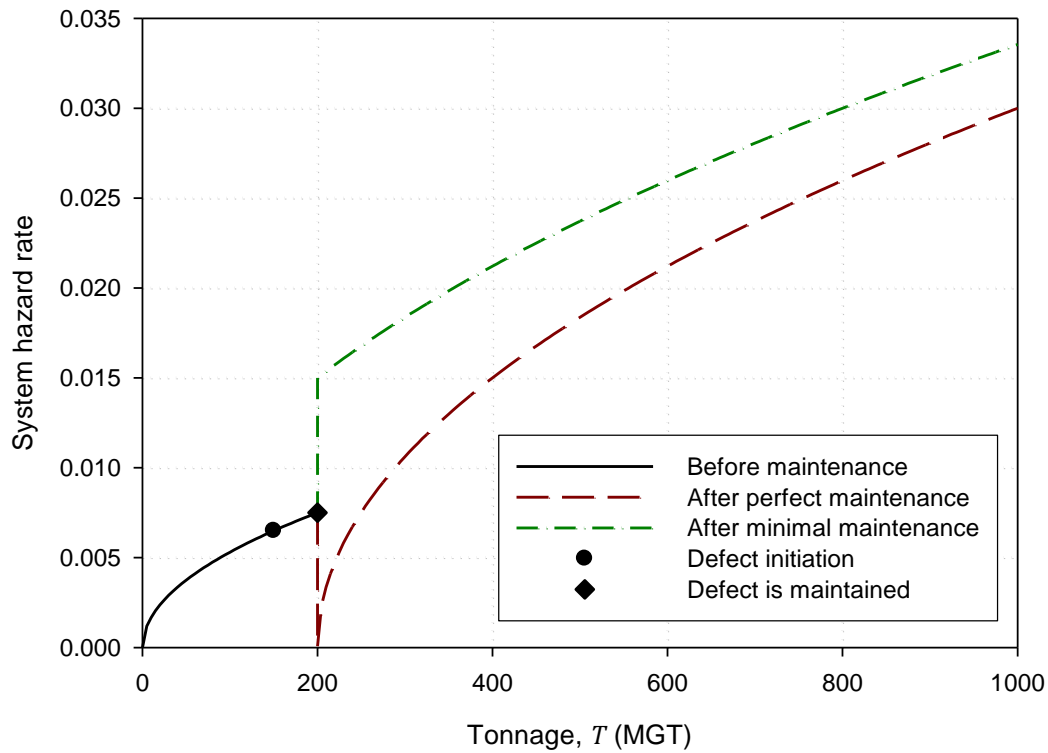


Figure 3.5: The effect of maintenance on the total system hazard rate for maintenance modelled using perfect maintenance and minimal maintenance

3.2.2 Modelling the P-F Interval

It is assumed that the P-F interval for fatigue defects is adequately modelled by a stationary exponential distribution as shown in Equation 3.6:

$$F(t) = 1 - e^{-\frac{t}{\mu}} \tag{3.6}$$

With:

μ = the expected P-F interval length (MGT)

Rearranging Equation 3.6 for t gives Equation 3.7:

$$t_{P-F}(U_2) = -\mu \ln 1 - U_2 \tag{3.7}$$

With:

t_{P-F} = the P-F interval length (MGT)

U_2 = a uniformly distributed variable sampled using Monte Carlo simulation

Equation 3.7 is used to calculate the P-F interval length t_{P-F} . U_2 is sampled using Monte Carlo simulation from the appropriate Latin hypercube sample. This produces values of t_{P-F} distributed according to the desired stationary exponential distribution.

3.3 PROCEDURAL LOGIC OF THE MODEL

This section describes the procedural logic of the model developed for this study. Flow charts are used to describe the procedures implemented within MATLAB. The nomenclature used in this section is outlined below:

- Arrays are denoted using a $\vec{}$.
- Arrays are indexed using a subscript outside of the $\vec{}$.
- The subscripts A and B are used to denote Category A and Category B defects respectively.
- The symbols i, j, k and p are reserved for use in indexing array elements.
- The symbol T denotes tonnage relative to the virtual life cycle under consideration i.e. $T = 0$ MGT at the beginning of the virtual rail life.
- The symbol t denotes a tonnage relative to the life of the defect or weld under consideration i.e. if a weld is installed at $T = 200$ MGT and a defect initiates for that weld at $T = 250$ MGT, then the defect initiation tonnage may be described as $T_d = 250$ MGT or $t_d = 50$ MGT.

The simulation procedure begins by simulating N_{sim} virtual life cycles with a renewal tonnage equal to \vec{T}_{Rend} . Each virtual life cycle simulated produces varying quantities of Category A and Category B defects due to Monte Carlo sampling. Data for all the Category A defects and Category B defects are saved for each virtual life cycle. The data saved for Category A and Category B defects respectively are summarised in Table 3.1. There is no T_{B_l} parameter in Table 3.1 because Category B defects are modelled with hazard functions proportional to the length of the rail and thus do not have an associated installation tonnage as is the case with welds (Category A defects).

Table 3.1: Data stored for Category A and Category B defects for each virtual life cycle modelled

<i>Variable Description</i>	<i>Symbol Used for Category A defects (Welds)</i>	<i>Symbol Used for Category B defects</i>
Defect inter-arrival tonnage (relative to the defect or weld)	t_{Ad}	t_{Bd}
P-F interval length (relative to the defect)	t_{AP-F}	t_{BP-F}
Virtual life cycle tonnage at defect initiation	T_{Ad}	T_{Bd}
Virtual life cycle tonnage at failure	T_{Af}	T_{Bf}
Virtual life cycle tonnage at weld installation	T_{AI}	-
Virtual life cycle tonnage at which defect was maintained	T_{Am}	T_{Bm}
Maintenance case number	C_A	C_B

The deterministic input parameters required to run the model are:

- Initial number of FBWs $n_{0_{FBW}}$ and initial number of ATWs $n_{0_{ATW}}$ at $T = 0$ MGT.
- An array of tonnages at which ultrasonic rail inspections will be conducted, \vec{I} .
- An array of tonnages at which preventive rail grinding will be conducted, \vec{I}_g .
- The hazard rate reduction factor for rail grinding, γ .
- The weld type used for maintenance.
- Maintenance modelling type, namely: perfect maintenance or minimal maintenance.
- Number of virtual life cycles to simulate, N_{sim} .
- An array of renewal tonnages at which the LCC is calculated, \vec{T}_R .

The following stochastic parameters need to be specified for each Category A and Category B defect modelled:

- The Weibull parameters α and β .
- The Expected P-F interval length, μ .
- The probability that a defect will be detected by ultrasonic inspection, η .

Whether or not grinding reduces the hazard rate of a defect can be set for each individual defect type modelled. The following costs need to be specified and should include all material, labour and delay costs:

- Cost of planned maintenance using ATWs for a detected defect, $c_{p_{ATW}}$.
- Cost of planned maintenance using FBWs for a detected defect, $c_{p_{FBW}}$.
- Cost of unplanned maintenance using ATWs for a failed defect, $c_{f_{ATW}}$.

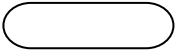

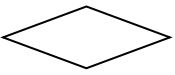
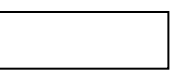

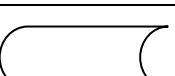
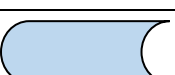
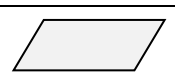
- Cost of unplanned maintenance using FBWs for a failed defect, $c_{f_{FBW}}$.
- Cost of ultrasonic inspection, c_I .
- Cost of rail grinding, c_g .
- Rail renewal cost, c_R .

The calculation procedure for a single virtual life cycle is described in Sections 3.3.1 and 3.3.2. The process described in these sections is repeated N_{sim} times to produce data for N_{sim} virtual life cycles. The calculation of the associated costs for each life cycle is described in Section 3.3.3. The entire procedure can be briefly summarised as follows:

- Category B defects are created using pseudorandom numbers and Monte Carlo sampling.
- The Category B defects are maintained using a combination of inspection interval data and the probability of detection for that defect type.
- The initial Category A defects are created from the welds present at $T = 0$ MGT.
- Category A defects are created for welds installed during the maintenance of Category B defects.
- All Category A defects are now maintained and new welds installed in a process which repeats until the cumulative tonnage of all the welds present in the system has reached the renewal tonnage.

The flow charts used in the proceeding sections make use of different shapes and colours as summarised in Table 3.2.

Table 3.2: Description of the shapes used in the flow charts

<i>Shape</i>	<i>Description</i>
	Indicates a starting condition or the start of a procedure.
	Indicates the end of a procedure.
	Indicates that a decision must be made.
	Indicates the assignment of a value to a variable.
	Indicates the use of an equation or function of which the procedural logic is explained in a separate flow chart.
	Indicates that values are stored in memory.
	Indicates that values are stored in memory and that the end of the procedure has been reached.
	Indicates the input data required for a function or decision.

3.3.1 Calculation Procedure for the Last Renewal Tonnage

The first step in calculating the defects which will originate in a virtual life cycle is to determine the Category B defects which will form during the interval $(0, \overline{T}_{R_{end}})$. Figure 3.6 illustrates this process. The formation of Category B defects is considered to be independent of Category A defects. This assumption implies that the quality of any closure rail installed during maintenance activities is the same as the existing rail to which it is welded. The variable j indexes the different Category B defect types to model ($j \in \mathbb{N}$). The variable k indexes the quantity of Category B defects across all j types. The defect inter-arrival tonnage for the k^{th} Category B defect $\overline{t}_{B_{d_k}}$ is calculated using the j^{th} dimension of the appropriate Latin hypercube sample as well as the input parameters indicated in Figure 3.6. Category B defects of type j are created while $\overline{T}_{B_{d_k}} < \overline{T}_{R_{end}}$. This means that defects in each category indexed by j are created whilst their cumulative defect initiation tonnage is less than the renewal tonnage. This is logical in the sense that any defects which would initiate after the renewal tonnage would not influence the current system. For every defect created an associated P-F interval length $\overline{T}_{B_{P-F_k}}$ is calculated using the appropriate dimension of the corresponding Latin hypercube sample and Equation 3.7. The theoretical failure tonnage for the k^{th} Category B defect is then calculated as $\overline{T}_{B_{f_k}} = \overline{T}_{B_{d_k}} + \overline{T}_{B_{P-F_k}}$.

Each defect created is given a unique identification number $\overrightarrow{ID}_{B_k}$ which is used for the procedures explained in Section 3.3.2. Furthermore, each defect can be categorised as either active or inactive. This is denoted by array $\overrightarrow{\psi}_B$ which can take on only logical values of true or false; true indicating that the defect is active. This is used to identify which defects still need to be maintained and which have already been maintained and have thus been removed from the system. Once all Category B defects have been created, arrays \overrightarrow{T}_{B_d} , \overrightarrow{T}_{B_f} , $\overrightarrow{\psi}_B$ and \overrightarrow{ID}_B are saved for use at a later stage in the simulation.

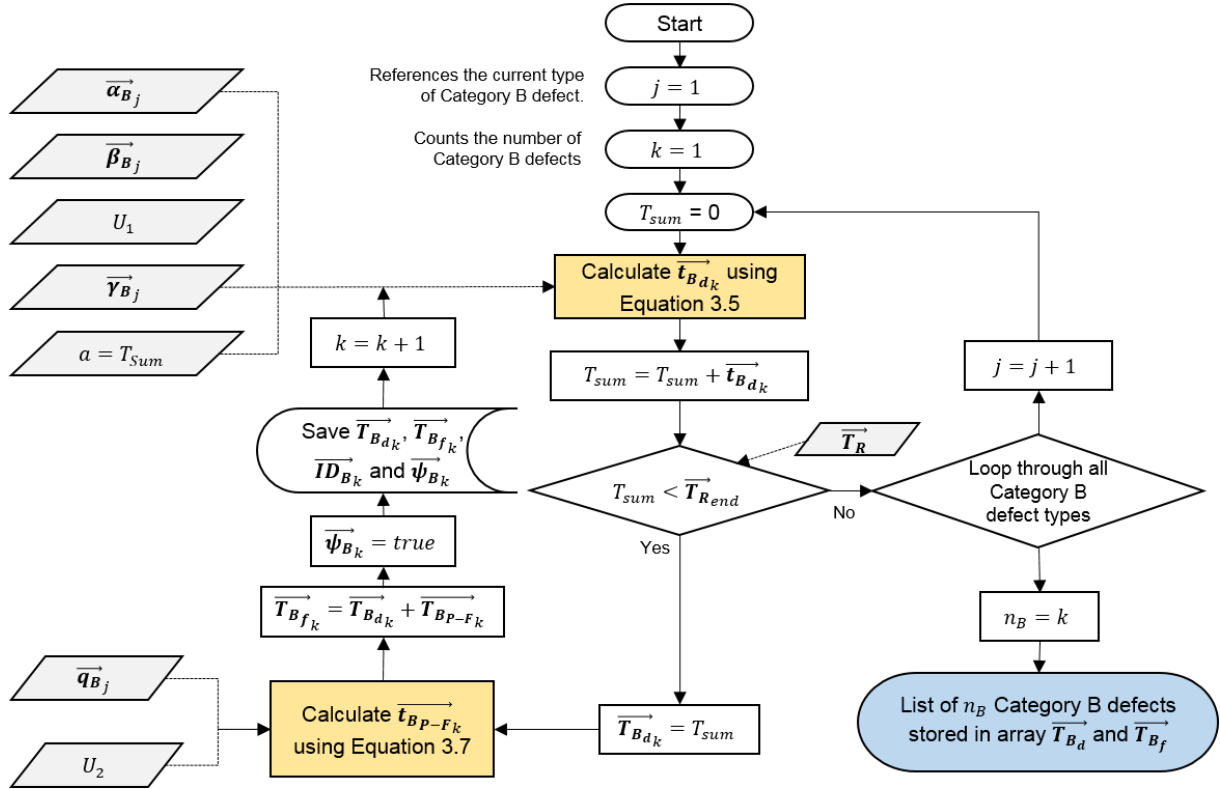


Figure 3.6: Flow chart illustrating the process of creating Category B defects in the model

The next stage involves simulating the maintenance of all the Category B defects created. This procedure is summarised in Figure 3.7. The cumulative defect initiation tonnage array \overrightarrow{T}_{B_d} and the cumulative failure tonnage array \overrightarrow{T}_{B_f} as well as the inspection interval array \vec{l} are used to determine the maintenance case number and the cumulative tonnage at which the defect is maintained. The determination of the maintenance case number for the k^{th} defect \overrightarrow{C}_{B_k} and cumulative tonnage at which the k^{th} defect is maintained $\overrightarrow{T}_{B_{m_k}}$ is calculated within the Maintenance Function illustrated in Figure 3.8. As each Category B defect is maintained its active state $\overrightarrow{\psi}_{B_k}$ is changed to false (inactive). Within the Maintenance Function is another function namely the Detection Function. The Detection Function determines if (and then subsequently when) a defect is detected by an ultrasonic inspection and is illustrated in Figure 3.9.

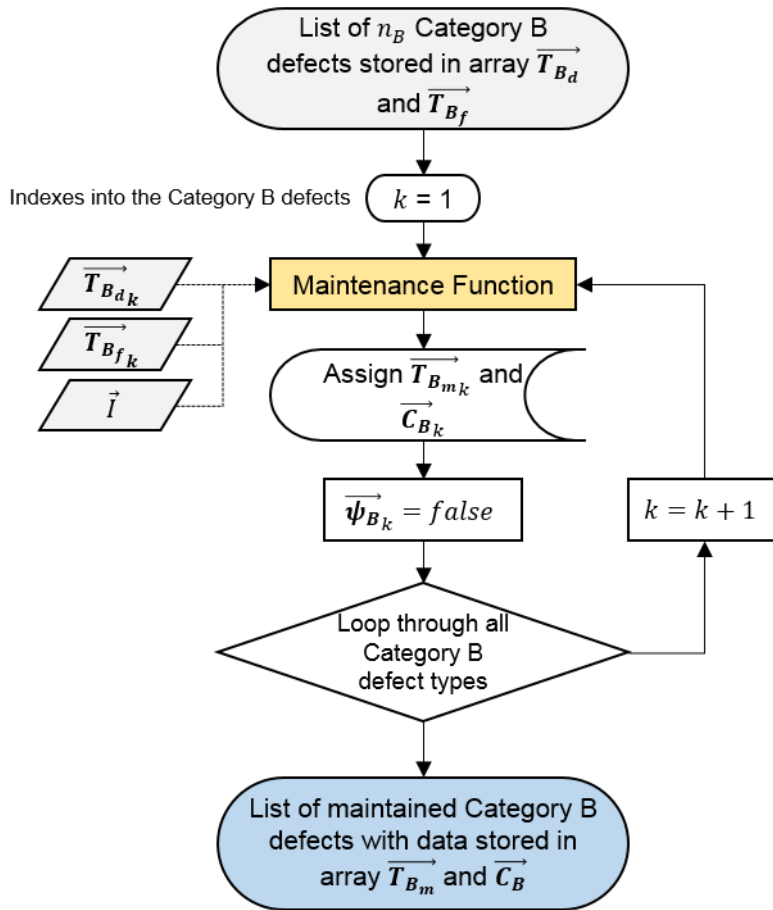


Figure 3.7: Flow chart illustrating the process of determining the tonnage at which maintenance of Category B defects occurs \vec{T}_{B_m} as well as the maintenance case number \vec{C}_B

The Maintenance Function illustrated in Figure 3.8 assigns a maintenance case number \vec{C}_{B_k} ranging from 1 to 7 according to specific criteria shown in the figure. Despite the existence of 7 different maintenance case numbers, only 3 categories of maintenance exist. These categories are planned maintenance, unplanned maintenance and renewal maintenance. The different maintenance case numbers exist only to separate the modes by which one of the 3 maintenance types occur in the model.

The first decision in Figure 3.8 determines whether the weld becomes defective before the renewal tonnage $\vec{T}_{R_{end}}$ is reached. This decision only applies to Category A defects. If the weld only becomes defective after $\vec{T}_{R_{end}}$ then it will be removed during renewal and thence be classified as renewal maintenance. Maintenance case number $\vec{C}_{A_i} = 7$ is assigned to such a case.

The next decision splits defects according to whether they fail before or after $\vec{T}_{R_{end}}$. If a defect fails before $\vec{T}_{R_{end}}$ then it can be further categorised according to whether it's cumulative tonnage at defect

initiation $\overrightarrow{T_{A_{d_i}}}$ occurs within the same inspection interval as the defect's cumulative tonnage at failure $\overrightarrow{T_{A_{f_i}}}$. $\overrightarrow{T_d}$ is a two element array indicating the inspection interval within which the defect originates. Similarly, $\overrightarrow{T_f}$ is a two element array indicating the inspection interval within which the defect fails. If $\overrightarrow{T_d} = \overrightarrow{T_f}$ then there is no chance that the defect may have been detected by an inspection and thus a functional failure occurs. Maintenance case number $\overrightarrow{C_{A_i}} = 1$ is assigned to such a case.

If $\overrightarrow{T_d} \neq \overrightarrow{T_f}$ it indicates that the defect was present during at least a single inspection. Subsequently, the Detection Function determines whether the defect is detected or not. If it is detected then planned maintenance occurs and the maintenance case number $\overrightarrow{C_{A_i}} = 2$ is assigned to that defect. If it remains undetected then maintenance case number $\overrightarrow{C_{A_i}} = 3$ is assigned.

If the defect fails after $\overrightarrow{T_{R_{end}}}$ ($\overrightarrow{T_{A_{f_i}}} > \overrightarrow{T_{R_{end}}}$) then it can be further categorised according to whether the defect originated within the last inspection interval (meaning that the second element of array $\overrightarrow{T_d}$, $\overrightarrow{T_{d_2}}$ is equal to $\overrightarrow{T_{R_{end}}}$) or not. If the defect originated within the last inspection interval then there is no chance for detection and thus renewal maintenance will be conducted and maintenance case number $\overrightarrow{C_{A_i}} = 4$ is assigned. Alternatively, a maintenance case number of $\overrightarrow{C_{A_i}} = 5$ or 6 will be assigned according to whether the defect was detected before $\overrightarrow{T_{R_{end}}}$ or not which corresponds with planned maintenance and renewal maintenance respectively.

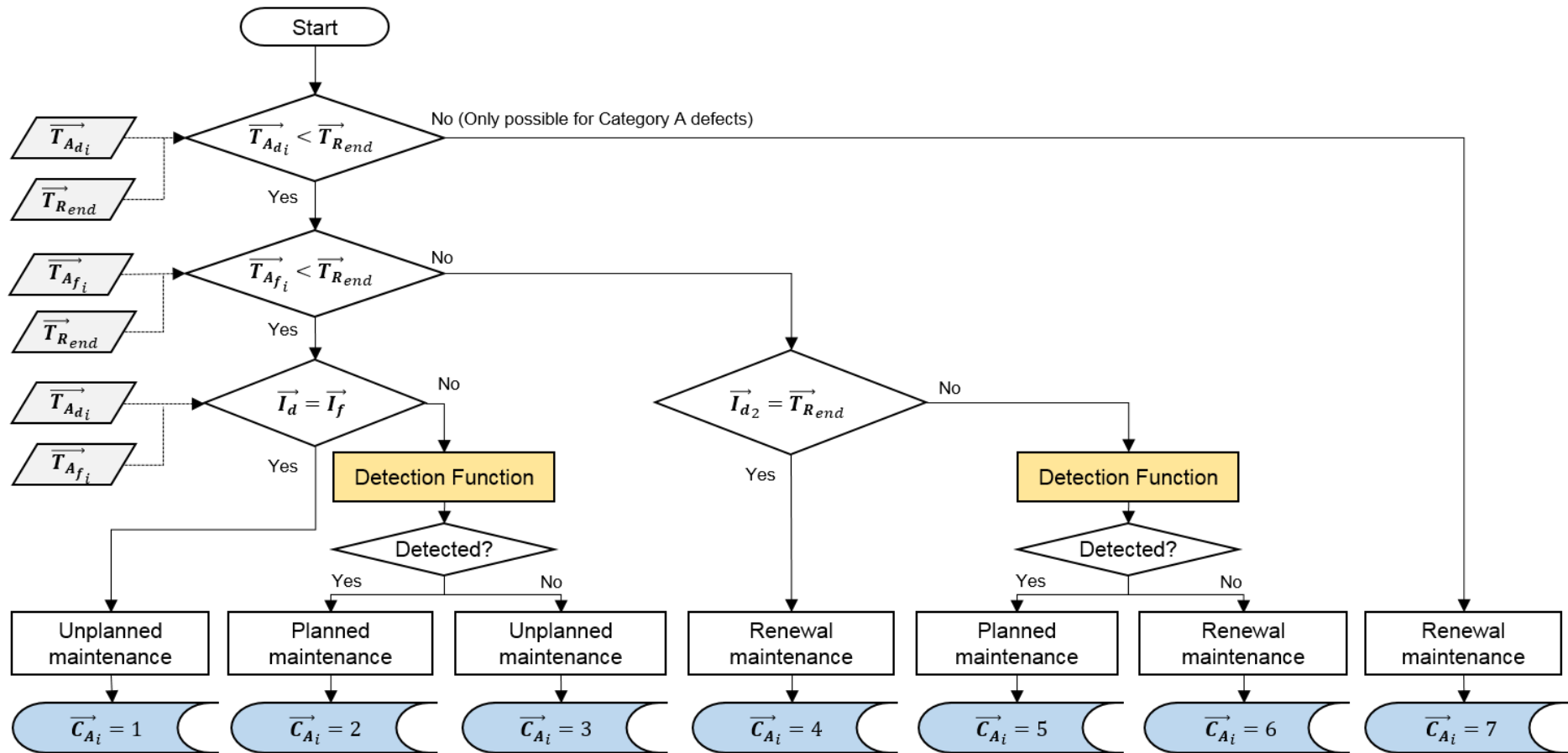


Figure 3.8: Flow chart illustrating the procedural logic contained within the Maintenance Function used to determine the tonnage at which maintenance is conducted \vec{T}_M as well as the maintenance case number \vec{C}

The Detection Function procedural logic illustrated in Figure 3.9 determines how many opportunities exist for the defect to be detected during the P-F interval. For each detection opportunity a random number is sampled using the Mersenne Twister pseudorandom number generator. If this number is less than the detection probability for the defect under consideration $\vec{\eta}_{A_j}$ then the defect is considered to have been detected by that specific inspection.

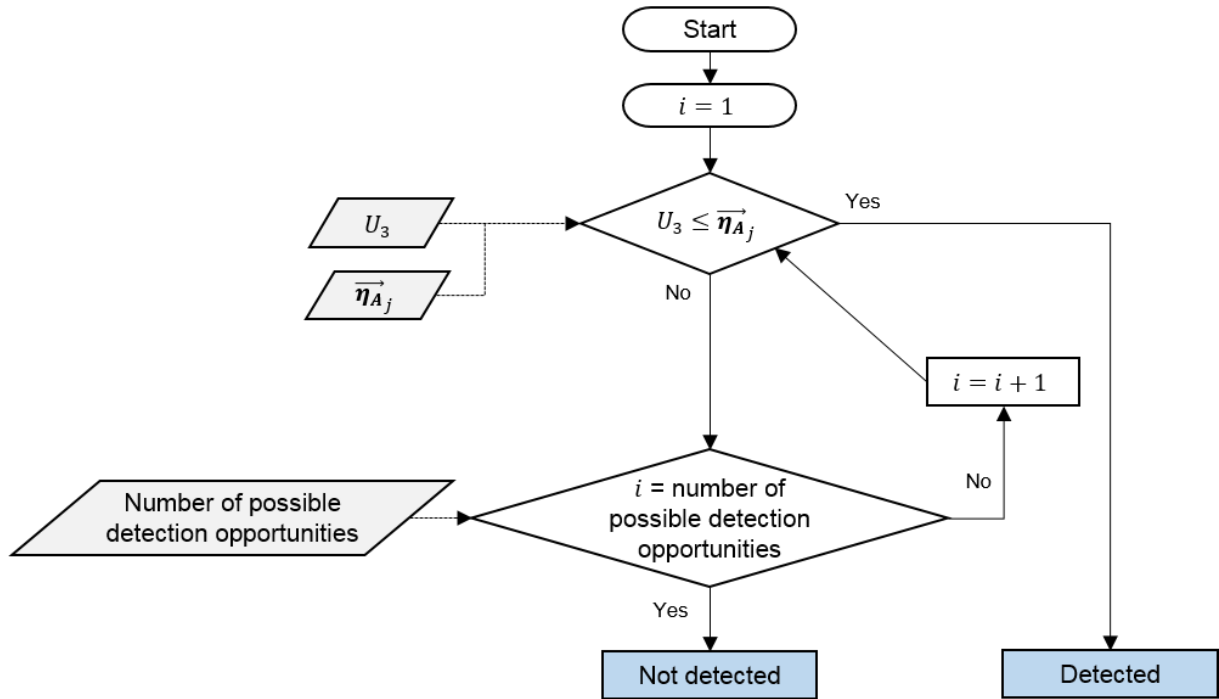


Figure 3.9: Flow chart illustrating the procedural logic contained within the Detection Function used to determine at which inspection (if at all) the defect is detected

It should be noted that subscripts denoting Category A defects are shown for the functions illustrated in Figure 3.8 and Figure 3.9. However, these functions are used for Category B defects as well and the subscript shown is used for illustrative purposes only.

The Category A defects for welds present at $T = 0$ MGT are now modelled. Figure 3.10 illustrates this process. Each weld has a unique defect inter-arrival tonnage \vec{t}_{Ad_i} and P-F interval length \vec{t}_{AP-Fi} . The procedure is similar to the process described at length in Figure 3.6. The difference in this process is that a limited number of Category A defects are created according to the specified number of ATWs $n_{0_{ATW}}$ and FBWs $n_{0_{FBW}}$. Furthermore, $a = 0$ in Equation 3.5 during this process as all these welds are installed at $T = 0$ MGT.

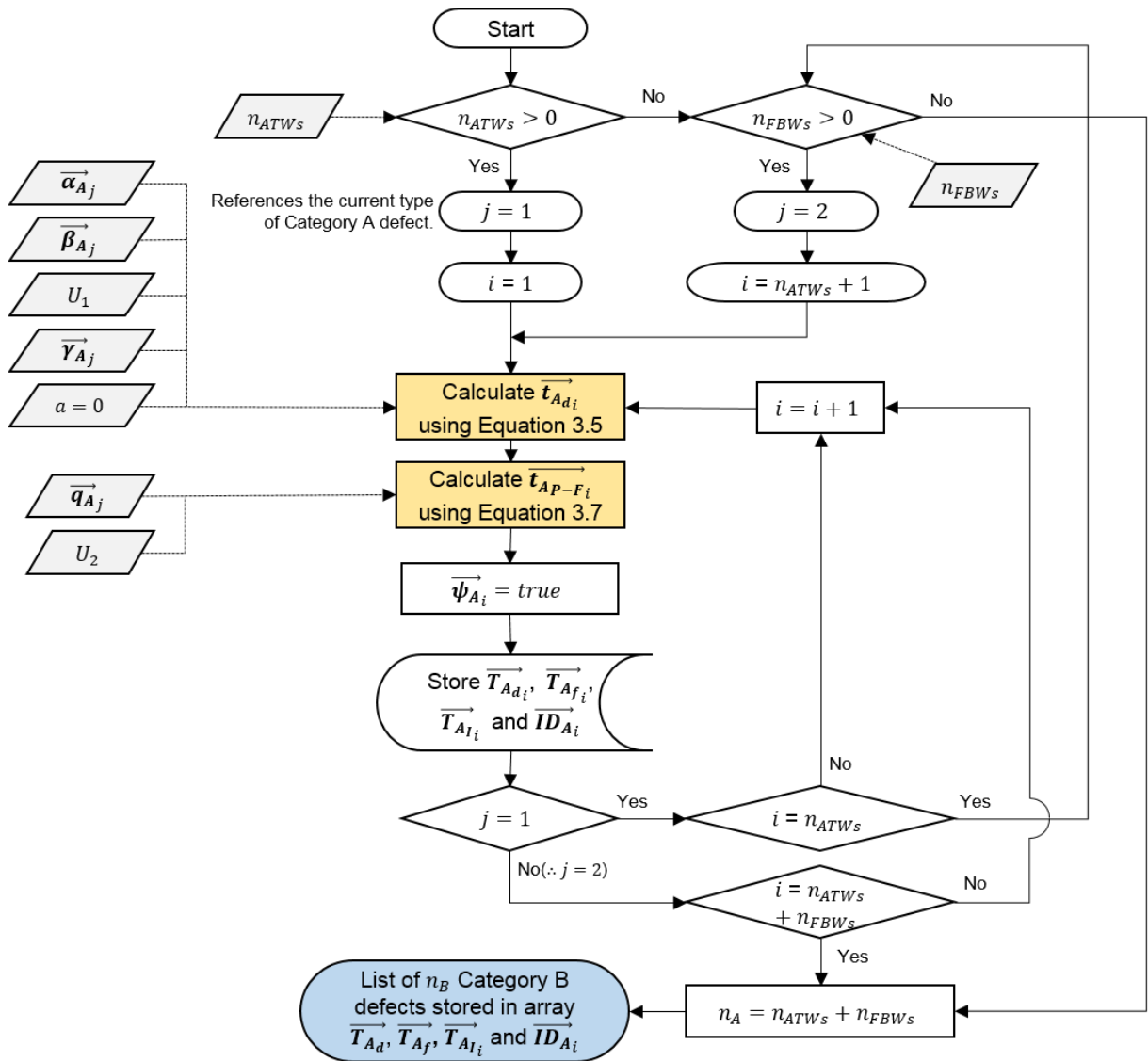


Figure 3.10: Flow chart illustrating the creation of the Category A defects for welds present at $T = 0$ MGT

Category A defects for the welds installed during maintenance of the Category B defects need are now modelled. This is illustrated in Figure 3.11. This process is similar to Figure 3.10. The differences are:

- Welds are only created for Category B defects which triggered planned or unplanned maintenance. This includes all maintenance case numbers except $\vec{C}_{B_k} = 4$ or 6 .
- The value of a in Equation 3.5 is either equal to 0 or $\vec{T}_{B_{mk}}$ depending on whether the maintenance type is modelled as perfect maintenance or minimal maintenance respectively.

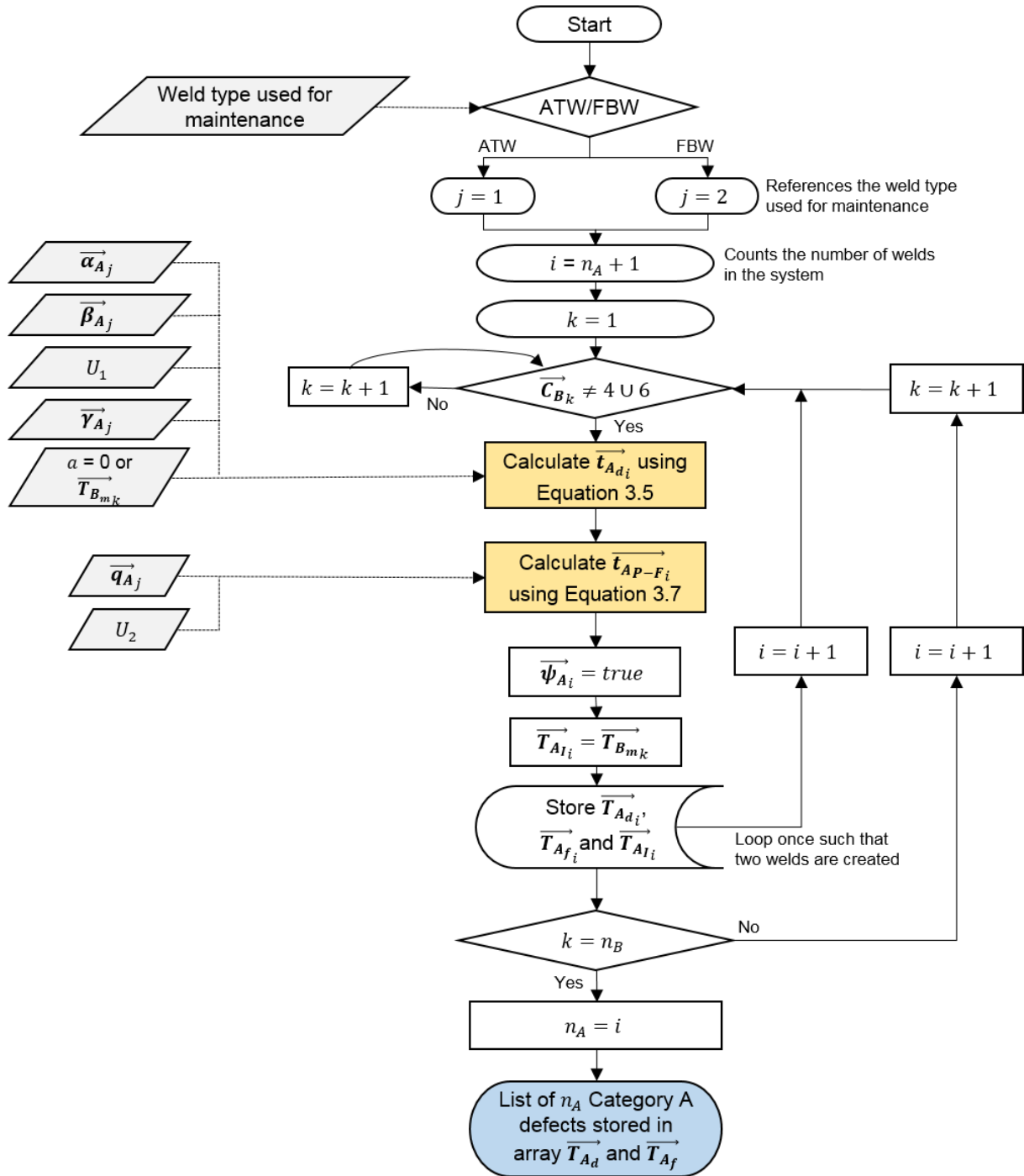


Figure 3.11: Flow chart illustrating the process whereby Category A defects are created for welds created by maintaining Category B defects

Now only Category A defects remain which need to be maintained and new welds installed accordingly. Figure 3.12 illustrates this procedure. For each weld maintained using either planned or unplanned maintenance two new welds are created. The Category A defects for these new welds are then created with their associated $\vec{T}_{A_{d_i}}$ and $\vec{T}_{A_{f_i}}$. This process continues until the renewal tonnage is reached by all active welds in the system.

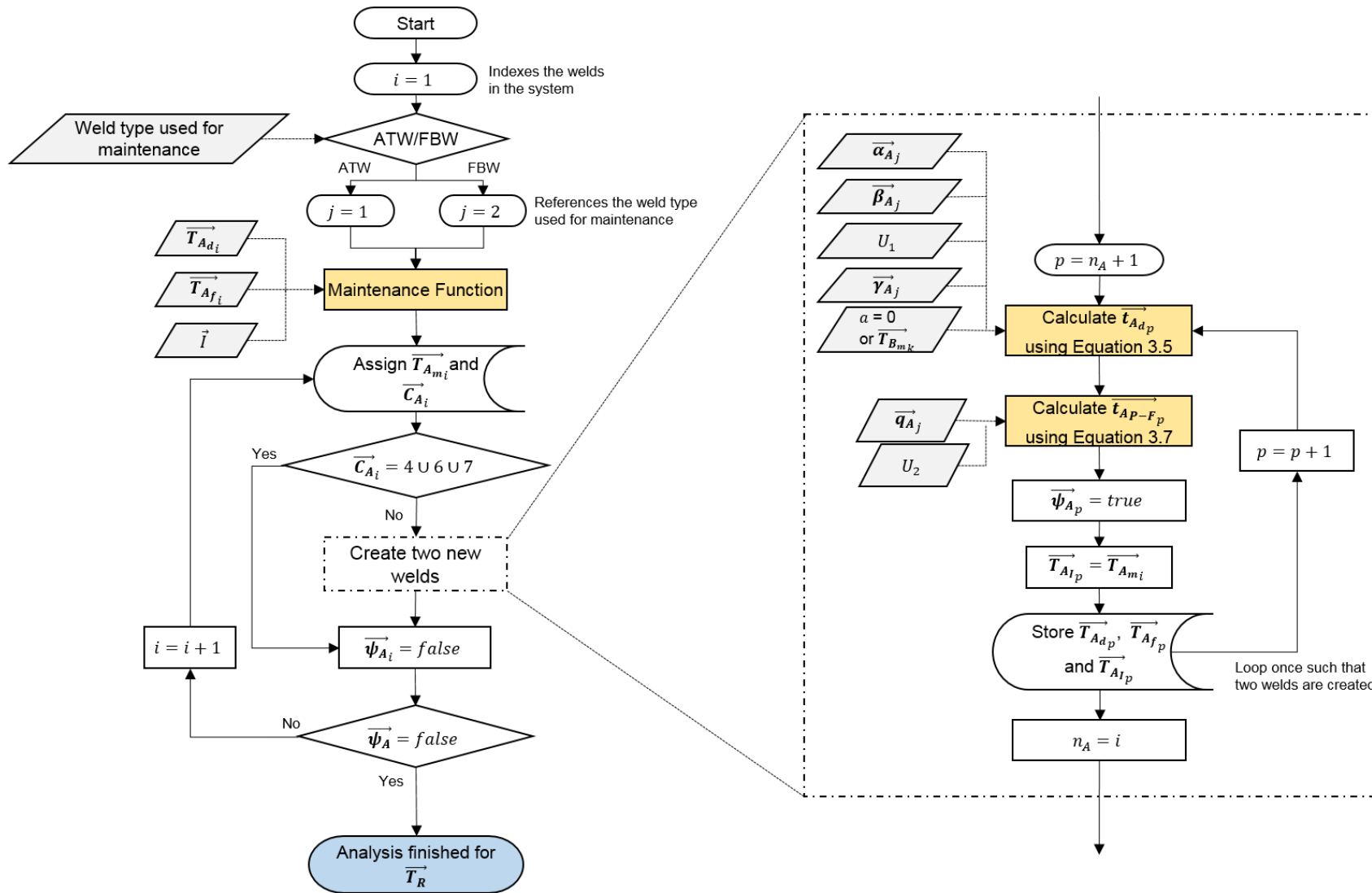


Figure 3.12: Flow chart illustrating the process of modelling the maintenance of all Category A defects and the creation of additional welds as a result thereof

3.3.2 Calculation Procedure for the Remaining Renewal Tonnages

The process described in Section 3.3.1 simulates all the rail defects for a renewal tonnage of $\vec{T}_{R_{end}}$. The defects present at the renewal tonnages other than $\vec{T}_{R_{end}}$ in array \vec{T}_R must be determined in order to calculate the associated LCC for all renewal tonnages specified. Therefore, the defects produced during the simulation process described in Section 3.3.1 must be truncated accordingly and their maintenance case numbers changed if so required. In this procedure the renewal tonnages in \vec{T}_R are looped through in descending order and are indexed by the variable p .

This process of truncation begins with the removal of any Category B defects which initiated at tonnages larger than the current renewal tonnage under consideration \vec{T}_{R_p} . Thereafter, any subsequent welds which originated from the Category B defects which have now been removed are also removed. This is illustrated in Figure 3.13.

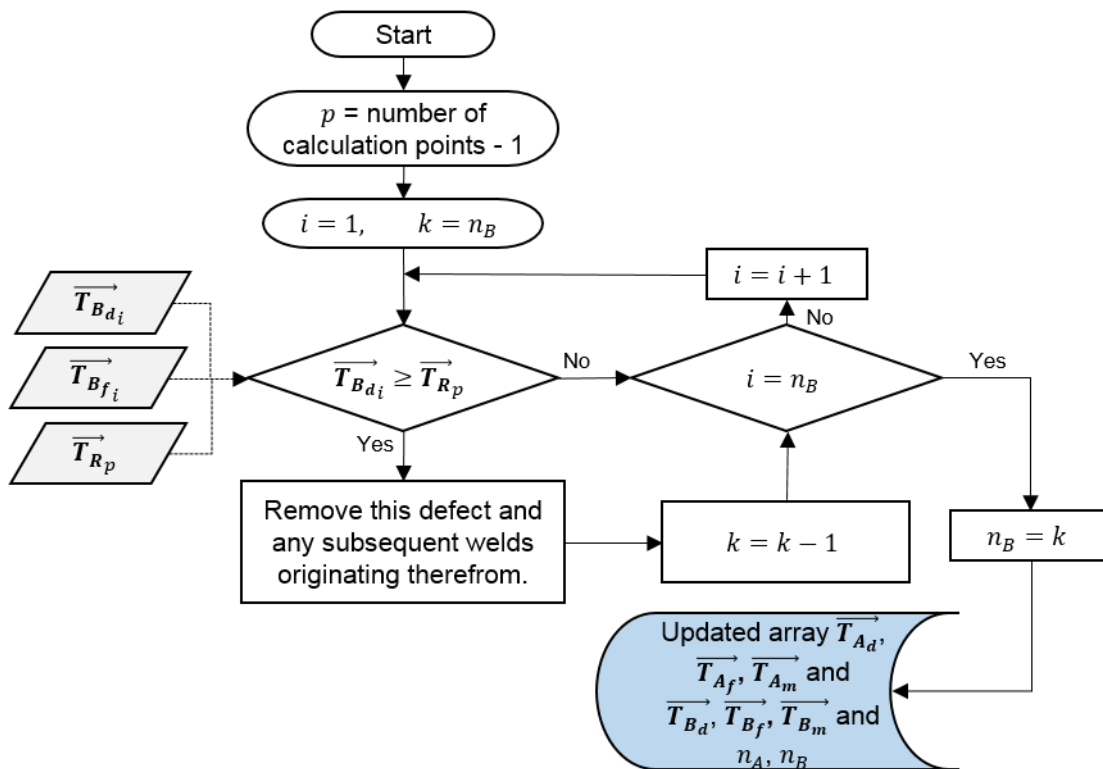


Figure 3.13: Flow chart illustrating the procedural logic followed to remove Category B defects which initiated at a tonnage larger than \vec{T}_{R_p} as well as any subsequent welds and Category A defects

The remaining Category B defects are now assessed to in order to ascertain whether any of their associated maintenance tonnages $\vec{T}_{B_{m_i}}$ are larger than the current renewal tonnage under consideration \vec{T}_{R_p} . If $\vec{T}_{B_{m_i}} > \vec{T}_{R_p}$ the tonnage at which maintenance is conducted is changed to the renewal tonnage. The maintenance case number is also changed accordingly. Thereafter, any child welds which originated from defects which had $\vec{T}_{B_{m_i}} > \vec{T}_{R_p}$ are removed. This is illustrated in Figure 3.14.

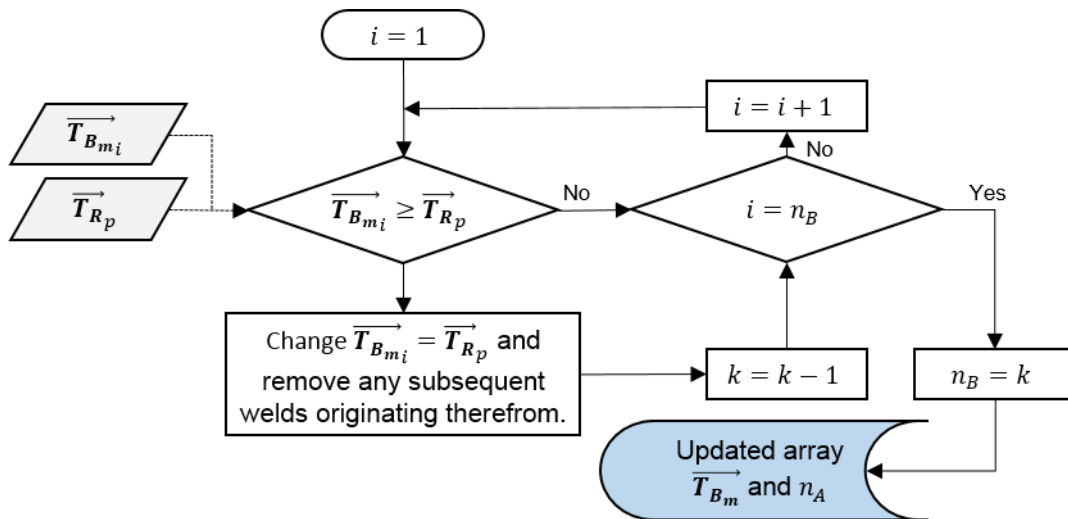


Figure 3.14: Flow chart illustrating the procedural logic followed to change the maintenance type of any remaining Category B defects with $\vec{T}_{B_{m_i}} > \vec{T}_{R_p}$ and removed any subsequent welds and Category A defects

Any welds which were installed after \vec{T}_{R_p} are removed. This is shown in Figure 3.15. Now a new set of arrays with the correct data is stored for the renewal tonnage under consideration. The process described in this section is repeated for all specified renewal tonnages in \vec{T}_{R_p} in descending order.

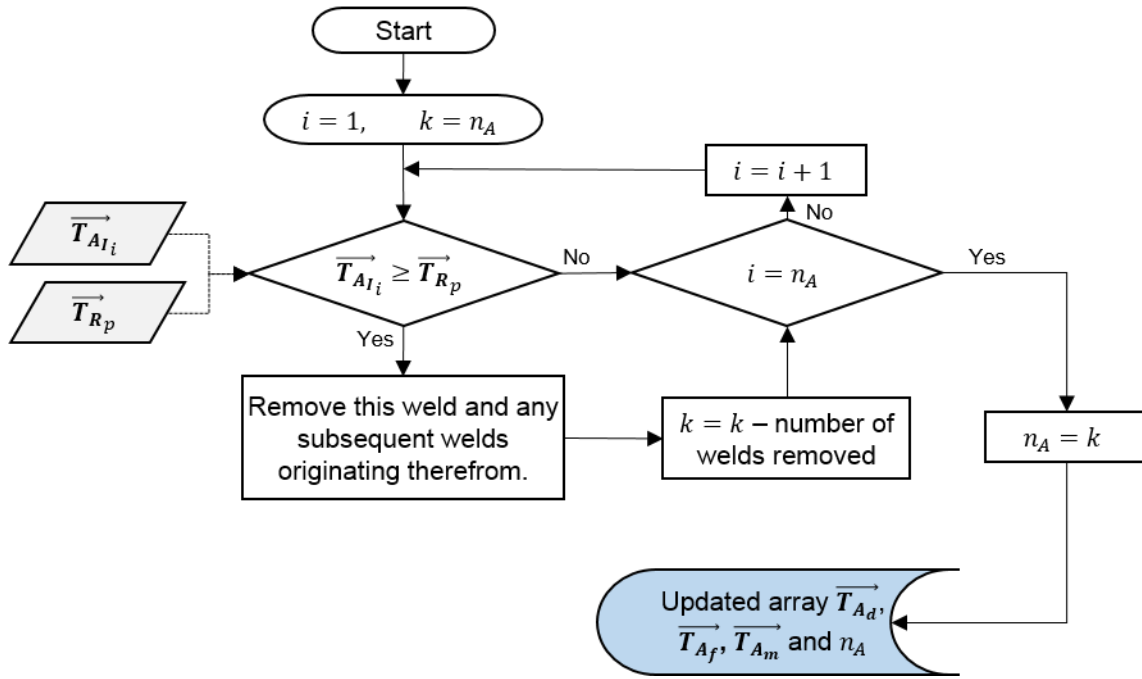


Figure 3.15: Flow chart illustrating the procedural logic used to remove any welds installed at tonnages larger than $\overrightarrow{T_{Rp}}$ and subsequent defects

3.3.3 Life Cycle Cost Calculation

The LCC for virtual life cycle N at renewal tonnage $\overrightarrow{T_{Rk}}$, $\overrightarrow{c_{LCC_{N,k}}}$ is calculated using the defect data for that respective life cycle and renewal tonnage. Arrays $\overrightarrow{C_A}$ and $\overrightarrow{C_B}$ are used in order to determine the planned and unplanned maintenance costs. For each maintenance case number within $\overrightarrow{C_A}$ and $\overrightarrow{C_B}$ a corresponding maintenance cost of $c_{p_{ATW}}$, $c_{p_{FBW}}$, $c_{f_{ATW}}$ or $c_{f_{FBW}}$ is added to the LCC according to whether the maintenance case number corresponds with planned or unplanned maintenance and whether ATWs or FBWs are used for maintenance respectively.

The ultrasonic inspection cost for a given virtual life cycle is equal to the number of inspections multiplied by the cost of a single inspection c_i . The rail grinding cost is calculated in a similar manner. The rail renewal cost is equal to c_R for any given virtual life cycle. All costs are normalised against the tonnage borne by the rail. Therefore, if a single km of rail is modelled the associated LCC will be expressed in units of cost/MGT/km.

The LCCs for a simulation of N_{sim} virtual life cycles across a series of renewal tonnages specified by $\overrightarrow{T_R}$ may be specified as a two dimensional array of size $N_{sim} \times \text{length of } \overrightarrow{T_R}$. Let N index the virtual life

cycle under consideration and let k index the renewal tonnage under consideration $\overrightarrow{T_{Rk}}$. Then the LCC array $\overrightarrow{c_{LCC}}$ may be calculated using Equation 3.8:

$$\overrightarrow{c_{LCC_{N,k}}} = \frac{(c_R + n_{I_k}c_I + n_{g_k}c_g + c_p \sum_{Planned} (\overrightarrow{C_{A_{N,k}}} + \overrightarrow{C_{B_{N,k}}}) + c_f \sum_{Unplanned} (\overrightarrow{C_{A_{N,k}}} + \overrightarrow{C_{B_{N,k}}}))}{\overrightarrow{T_{Rk}}} \quad (3.8)$$

Where:

$\sum_{Planned} (\overrightarrow{C_{A_{N,k}}} + \overrightarrow{C_{B_{N,k}}}) \triangleq$ the number of planned maintenance activities during the tonnage interval $(0, \overrightarrow{T_{Rk}})$

$\sum_{Unplanned} (\overrightarrow{C_{A_{N,k}}} + \overrightarrow{C_{B_{N,k}}}) \triangleq$ the number of unplanned maintenance activities during the tonnage interval $(0, \overrightarrow{T_{Rk}})$

With:

c_R = the cost of renewing the rail per km of track

n_{I_k} = the quantity of ultrasonic inspections conducted during the tonnage interval $(0, \overrightarrow{T_{Rk}})$

c_I = the cost of an ultrasonic inspection per km of track

n_{g_k} = the quantity of grinding activities conducted during the tonnage interval $(0, \overrightarrow{T_{Rk}})$

c_g = the cost of rail grinding per km of track

c_p = the cost of a planned maintenance activity

c_f = the cost of an unplanned maintenance activity

$\overrightarrow{T_{Rk}}$ = the k^{th} renewal tonnage in the renewal tonnage array $\overrightarrow{T_R}$

The terms in the numerator of Equation 3.8 divide the LCC calculation procedure into its individual cost components respectively:

- The cost of renewing the rail at the end of its life cycle.
- The cost of inspecting the rail.
- The cost of rail grinding for the rail.
- The cost of all planned maintenance actions.
- The cost of all unplanned maintenance actions.

No salvage values are considered for the analyses as this model considers costs only. Any analysis period length may be used. However, an analysis period length of $T_R = 800$ MGT was used in most analyses presented in Section 4 as this is large enough to demonstrate the behaviour of the uncertainty within the

LCC and this analysis window includes the renewal tonnage at which the minimum mean LCC is achieved. The LCC calculated can be analysed and interpreted in different ways and this is described in Section 4.2. The calculation procedure has now been fully defined.

4 ANALYSIS AND DISCUSSION OF RESULTS

This section provides a thorough interpretation and discussion of the results of the simulation model developed and described in Section 3. A reference case analysis is defined in Section 4.1. Section 4.2 shows the raw data produced by the model and discusses various methods to present and interpret the data. Section 4.3 is concerned with ascertaining the adequacy of the posterior distributions in the analyses run. Thereafter, the model is critically compared to a model developed by Zhao et al (2006) in Section 4.4. The results of various distribution analyses is presented in Section 4.5. Finally, the influence of the inspection interval length, the weld type used for maintenance as well as the effect of derailments is discussed in Sections 4.6, 4.7 and 4.8 respectively.

4.1 REFERENCE CASE ANALYSIS

A reference case analysis is defined such that comparisons can be made without specifying the entire input parameter set for each analysis run. The parameters which define the reference case analysis are as specified in Table 4.1 through Table 4.3. Table 4.1 shows the stochastic parameters used in the reference case analysis and are taken from Zhao et al. (2006). Table 4.2 shows the cost parameters used. These cost parameters are representative of typical costs which can be expected for a South African heavy haul scenario. Table 4.3 shows any deterministic parameters and further criteria specified for the reference case analysis. The values of the parameters in Table 4.3 were taken from Zhao et al. (2006) where applicable so that the two models could be compared.

Table 4.1: Stochastic input parameters used for the reference case analysis

<i>Defect</i>	α	β (MGT)	μ (MGT)	η
ATW ¹ defects	1.01	315.8	10	0.7
FBW ² defects	2.00	286.6	10	0.7
Squat defects	2.50	191.8	5	0.6
Tach ovale defects	2.17	182.3	7	0.7

¹ alumino-thermic weld

² flash butt weld

Table 4.2: Cost parameters used for the reference case analysis

<i>Cost parameter</i>	<i>Symbol</i>	<i>Value</i>
Planned repair cost using ATWs per defect	c_{pATW}	R 16 000.00
Unplanned repair cost using ATWs per failure	c_{fATW}	R 115 000.00
Ultrasonic inspection cost per km	c_I	R 2 400.00
Rail grinding cost per km	c_g	R 45 000.00
Rail renewal cost per km	c_R	R 2 050 000.00

Table 4.3: Deterministic input parameters used for the reference case analysis

<i>Parameter</i>	<i>Symbol</i>	<i>Value</i>
Ultrasonic inspection interval (MGT)	s_I	2.5
Rail grinding interval (MGT)	s_g	10.0
Initial number of ATWs	$n_{0_{ATWs}}$	0
Initial number of FBWs	$n_{0_{FBWs}}$	22
Hazard rate reduction factor due to rail grinding	γ	0.4
Number of simulated life cycles	N_{sim}	20 000
Weld type used for maintenance	-	ATW
Maintenance modelling type	-	Perfect maintenance

4.2 TYPICAL RESULTS

The results of the simulation model allow the life cycle cost (LCC) c_{LCC} to be interpreted as either a point estimate (for example the mean LCC $\overline{c_{LCC}}$) or as a probability distribution with an estimated parameter set each with their own associated confidence intervals. $N_{sim} = 20\,000$ virtual life cycles were simulated with parameters specified by the reference case analysis. Thereafter, the resulting LCC for each life cycle was grouped and interpreted in various ways. Grouping of the LCC by renewal tonnage and then calculating the mean LCC value at each renewal tonnage produces a plot as shown in Figure 4.1. The curve in Figure 4.1 shall be referred to as the mean LCC curve. This interpretation of the LCC is comparable to that of Zhao et al. (2006) shown in Figure 2.24. From Figure 4.1 it can be seen that a minimum attainable mean LCC $\overline{c_{LCC_{min}}}$ occurs at a critical renewal tonnage $T_{R_{cr}}$ for a set of given input parameters and conditions.

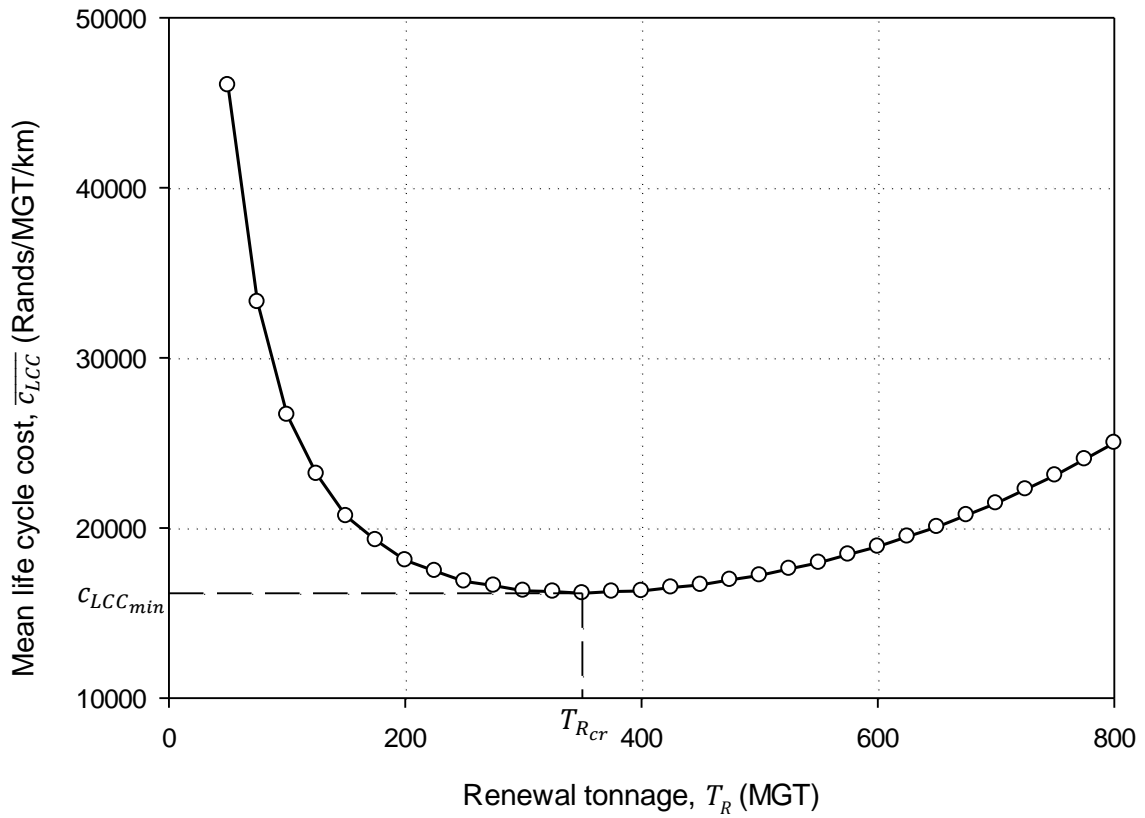


Figure 4.1: Mean LCC versus renewal tonnage T_R for the reference case analysis

An alternative approach to interpret the data is to analyse the distribution of the LCC values c_{LCC} at a fixed renewal tonnage T_R . This provides an indication of the uncertainty associated with the LCC at a given renewal tonnage. This approach is illustrated in Figure 4.2. Plots of both the relative frequency as well as the cumulative relative frequency are shown in Figure 4.2 (a) and (b) respectively.

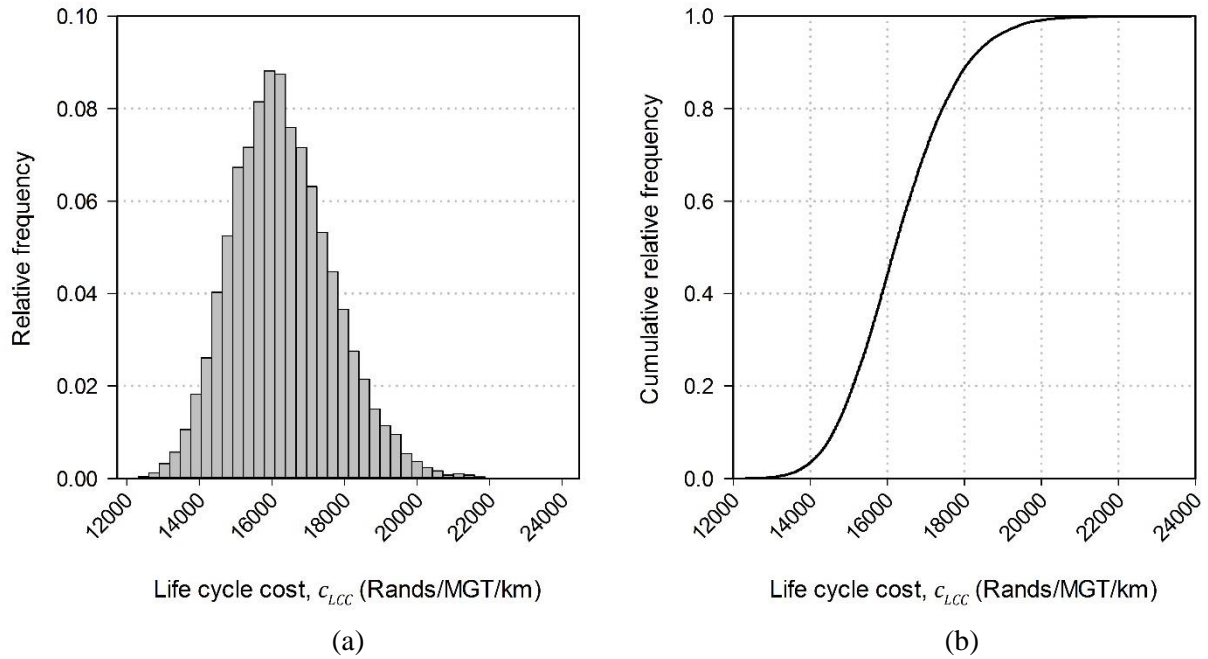


Figure 4.2: Distribution of LCC at a renewal tonnage of 400 MGT for the reference case analysis presented as (a) a relative frequency plot and (b) a cumulative relative frequency plot

The plots shown in Figure 4.1 and Figure 4.2 are in fact two-dimensional representations of a three-dimensional distribution. The three-dimensional distribution of LCC c_{LCC} versus renewal tonnage T_R is illustrated in Figure 4.3 using a bivariate histogram. Figure 4.2 is the result of taking a section of constant renewal tonnage T_R in the life cycle cost – relative frequency plane of Figure 4.3. Taking the mean (or expected) value of each two-dimensional distribution section as shown in Figure 4.2 and plotting it on axes of mean life cycle cost $\overline{c_{LCC}}$ versus renewal tonnage T_R results in Figure 4.1.

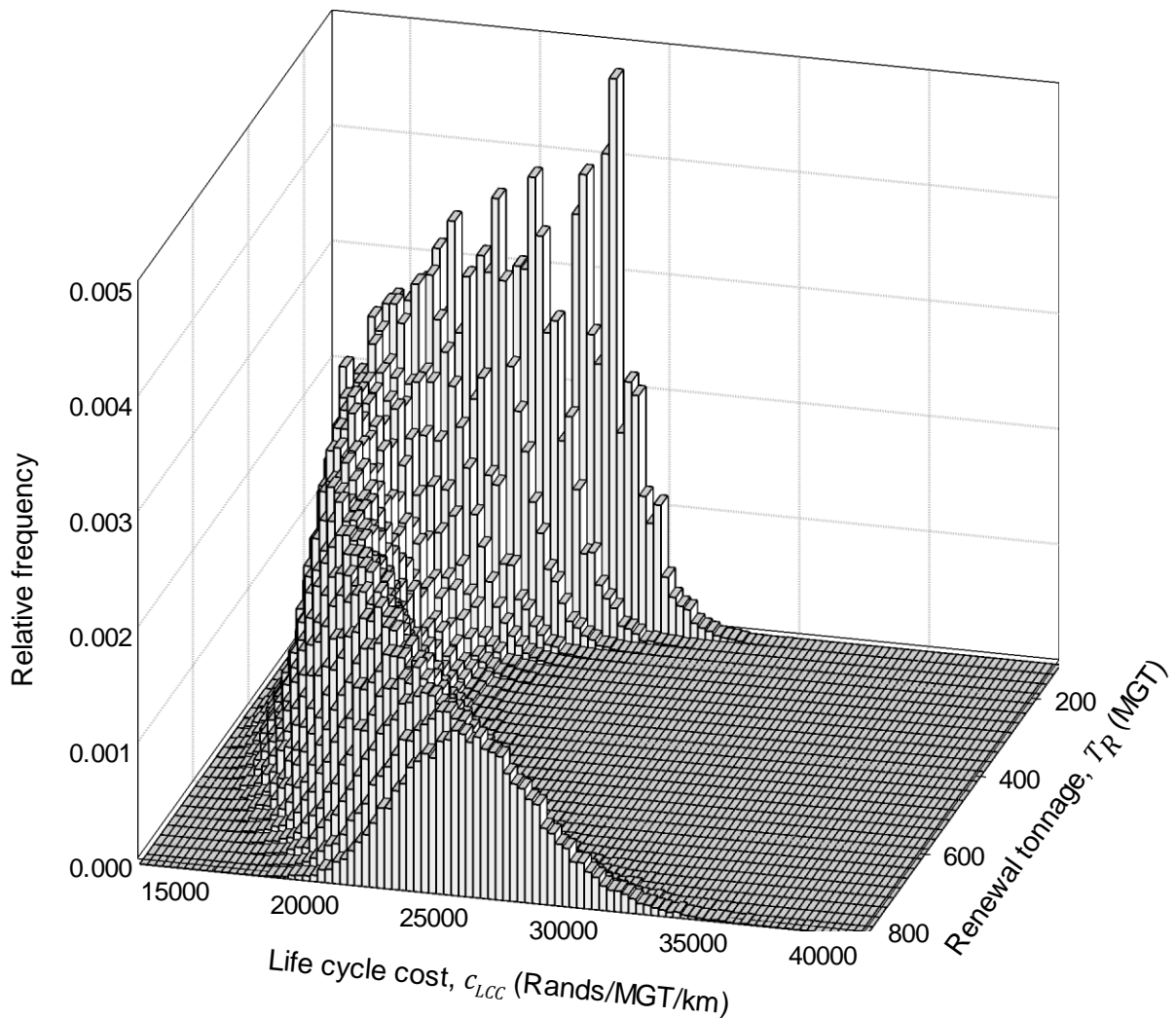


Figure 4.3: Distribution of LCC versus renewal tonnage in three-dimensional space

4.3 POSTERIOR DISTRIBUTION ANALYSIS

The behaviour of the posterior distributions produced by the model's Monte Carlo simulation is addressed in this section. The effect of the number of simulations N_{sim} and the renewal tonnage T_R on the adequacy of the posterior distributions is demonstrated. The posterior distributions of concern for each defect type are:

- The tonnage-to-defect initiation t_d
- The P-F interval length t_{p-f}
- Probability of detection η

The parameters used for these distribution tests are as stated in Table 4.1 through Table 4.3 for the reference case analysis. The adequacy of the generated distributions are tested using the Pearson Chi-

square goodness-of-fit test where applicable. It is known that the result of the Pearson Chi-square test is subject to the method of binning used. According to Moore (1986), accurate and repeatable results are obtained for the Pearson Chi-square goodness-of-fit test when using equiprobable bins and a number of bins as specified in Equation 4.1. The $\lfloor \cdot \rfloor$ in Equation 4.1 indicates that the value calculated should be rounded down to the nearest integer.

$$n_{bins} = \left\lfloor 2(n_{data})^{\frac{2}{5}} \right\rfloor \quad (4.1)$$

With:

n_{bins} = the number of equiprobable bins to use in the Chi-Square goodness-of-fit test

n_{data} = the number of data points in the sample to be tested

Bin edges are thus calculated for the n_{bins} ensuring that each bin is equiprobable under the hypothesised distribution. Category A defects modelled using minimal maintenance and all Category B defects cannot be tested using the Pearson Chi-square goodness-of-fit test. This is because the value of a in Equation 3.5 is not equal to zero except for the special case where $T = 0$ MGT. This implies that the sampling distribution for each new defect (for $T > 0$) is dependent on the value of a . Therefore, the resulting defect inter-arrival tonnages t_d do not follow a single defined distribution against which the Pearson Chi-square goodness-of-fit test can be applied. Conducting a goodness-of-fit test for a non-homogeneous Poisson process (NHPP) under these circumstances is not menial and is discussed by Lindqvist & Rannestad (2011). Thus, proving that the inter-arrival tonnages t_d follow the correct distribution for Category A defects modelled using minimal maintenance and all Category B defects falls outside the scope of this study. However, the behaviour of these distributions is demonstrated in Section 4.3.1 in order to provide evidence of their correctness.

The defect inter-arrival tonnage t_d for Category A defects which are modelled using perfect maintenance are derived from a NHPP with $a = 0$ in Equation 3.5. Therefore, the distribution of inter-arrival tonnages t_d for Category A defects modelled using perfect maintenance can be tested using the Pearson Chi-square goodness-of-fit test.

The empirical cumulative distribution function (ECDF) and the desired cumulative distribution function (CDF) for the defect inter-arrival tonnages t_d and P-F interval t_{P-F} for all defect types are compared for $N_{sim} = 5000, 10\ 000, 20\ 000$ and $40\ 000$ virtual life cycles at renewal tonnages T_R ranging from 100 MGT to 800 MGT in 100 MGT increments where applicable as discussed above. The respective P-value is also calculated for each case. The P-value is the probability that the test statistic (in this case

the Pearson Chi-square statistic) will take on a value that is at least as extreme as the observed value of the statistic when the null hypothesis is true (Montgomery & Runger, 2011). The null hypothesis is that the distribution derived through Monte Carlo simulation follows the desired distribution (specified using the input distribution parameter values). High P-values suggest strong evidence towards the null hypothesis and are thus desirable. P-values larger than 0.1 are deemed satisfactory for this study.

4.3.1 Distribution of Defect Inter-Arrival Tonnages

Figure 4.4 (a) shows the P-values for the Pearson Chi-square goodness-of-fit tests conducted on the distribution of inter-arrival tonnages t_d of ATW defects for the reference case analysis. It can be seen that the P-value does not decrease below 0.3 regardless of the number of virtual life cycles simulated or the renewal tonnage considered. This indicates that there is insignificant evidence to reject the null hypothesis and it is therefore assumed that the distribution of inter-arrival tonnages t_d for ATWs modelled using perfect maintenance does indeed follow the desired distribution. The number of inter-arrival tonnage data points used for each calculation is indicated in Figure 4.4 (b). The number of welds appears to increase exponentially with an increase in tonnage during a simulated life cycle and therefore the number of data points appears to increase linearly with tonnage when plotted on a logarithmic scale.

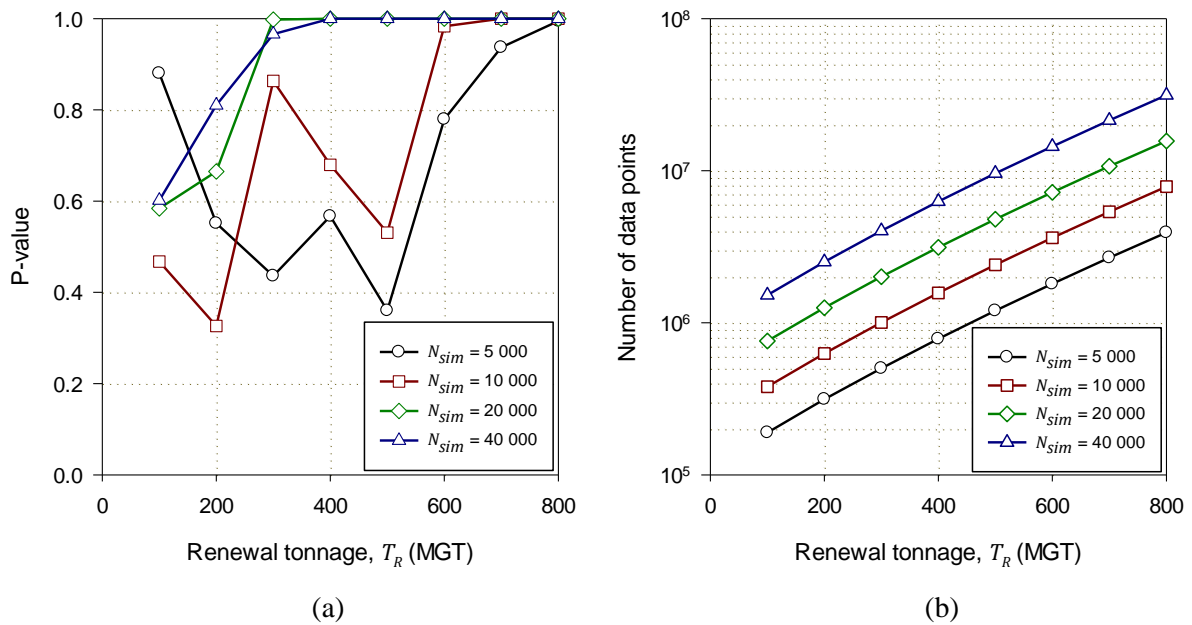


Figure 4.4: (a) P-value and (b) number of data points used in the calculation thereof at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of inter-arrival tonnages t_d for ATWs assuming perfect maintenance

The behaviour of the CDF used to sample the inter-arrival tonnages t_d of Category A defects modelled using minimal maintenance and all Category B defects is demonstrated in Figure 4.5. Equation 3.5 is used during the Inverse Transform Method to sample the defect inter-arrival tonnages. Equation 3.5 gives rise to the CDFs plotted in Figure 4.5 for values of $a = 0, 100, 200, 300, 400$ and 500 MGT respectively. The CDFs in Figure 4.5 were determined using the Weibull parameters for squat defects. It can be seen that the CDF of defect inter-arrival tonnages t_d tends to rise more steeply with larger values of a . This behaviour is expected because an increasing value of a indicates that the system is getting older and thus the defect inter-arrival tonnage t_d would be expected to become less as the system ages with increasing tonnage. The rate of decrease of defect inter-arrival tonnage t_d is governed by the Weibull shape parameter α .

An infinite number of distributions exists corresponding to values of a in the interval $[0, T_R)$. Thus, it is not possible to test the goodness-of-fit of the defect inter-arrival tonnages t_d sampled through Monte Carlo simulation using the Pearson Chi-square test under these conditions. However, the distribution of P-F interval lengths t_{P-F} for the different defect types can be tested using the Pearson Chi-square test in a similar manner as presented in Figure 4.4.

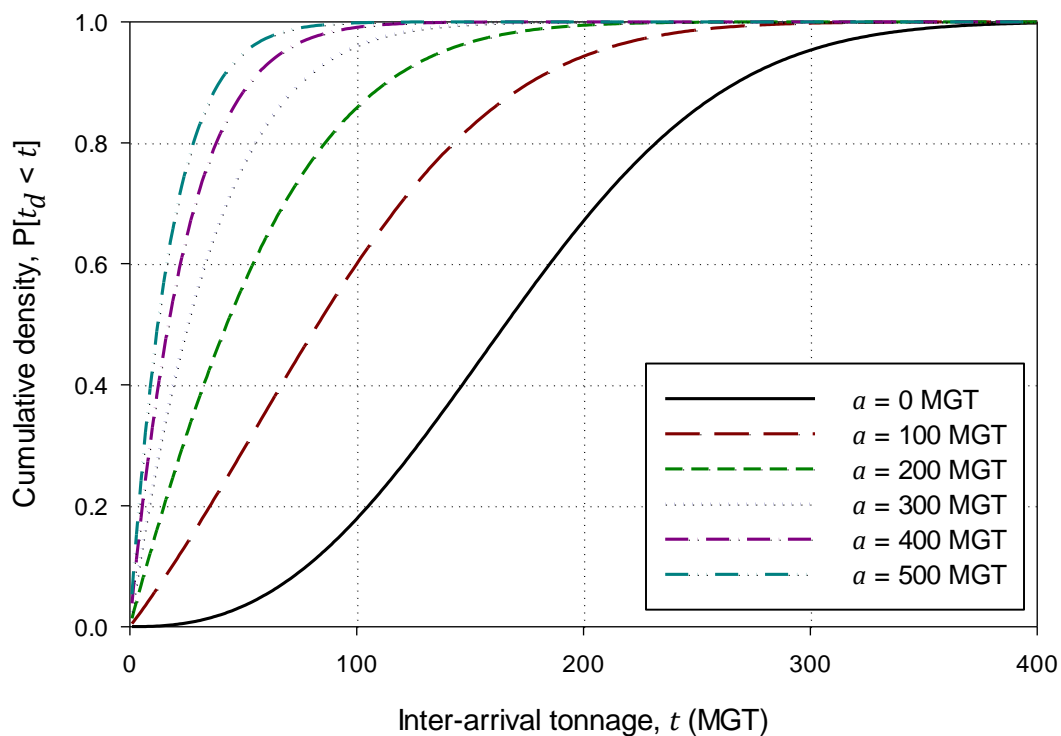
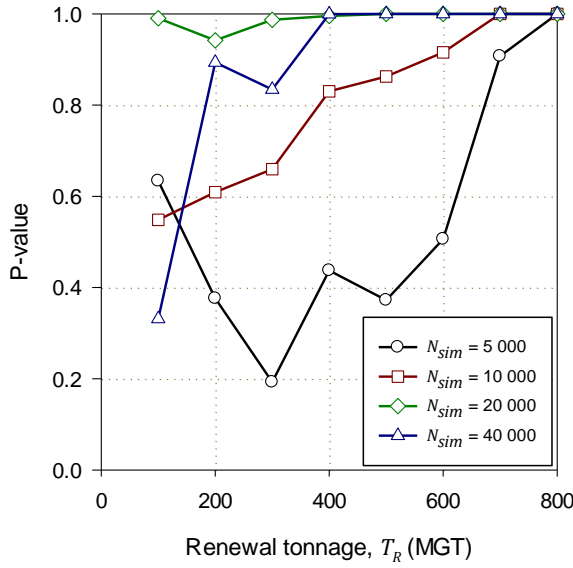


Figure 4.5: Illustrative behaviour of the CDF used to sample the defect inter-arrival tonnages t_d of Category A defects modelled using minimal maintenance and all Category B defects

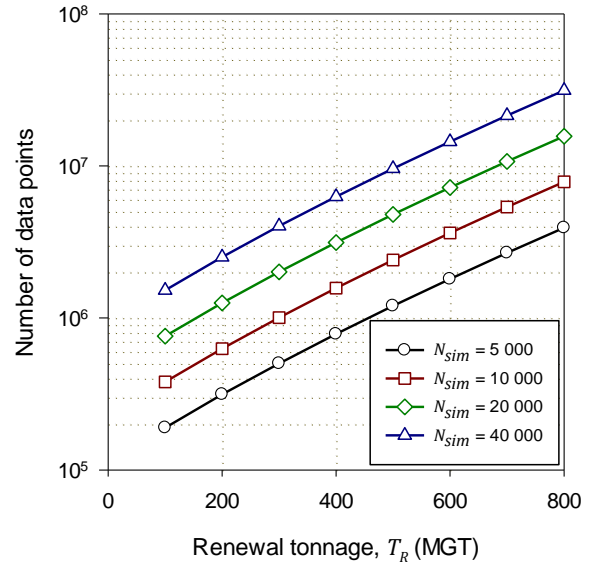
4.3.2 Distribution of P-F Interval Lengths

The P-F interval lengths are modelled using a stationary exponential distribution as was described in Section 3.2.2. Pearson Chi-square goodness-of-fit tests were conducted to test the hypothesis that the distribution of values produced through Monte Carlo simulation follows the desired distribution specified by the mean value parameter μ for a specific defect type (see Equation 3.6). The P-value for each test was calculated at 100 MGT increments and for $N_{sim} = 5\ 000, 10\ 000, 20\ 000$ and $40\ 000$ virtual life cycles respectively. The number of data points available for each test at each 100 MGT increment and each value of N_{sim} was also determined in order to visualise the size of each sample tested. The resulting P-values are shown in subplot (a) of Figure 4.6 through Figure 4.8 for ATW defects, squat defects and tache ovale defects respectively. The sample size for each test is shown in subplot (b) of the corresponding figure.

The P-value does not drop below a value of 0.1 across all tests in Figure 4.6 through Figure 4.8. This is assumed to be sufficient evidence that the desired distribution of P-F interval lengths t_{P-F} for all defect types have been achieved by the Monte Carlo simulation. Note that the plots of renewal tonnage T_R versus number of data points for the Category B defects (Figure 4.7 (b) and Figure 4.8 (b)) does not follow a linear trend when plotted on a logarithmic scale as is the case for Category A defects (Figure 4.6 (b)). The reason therefore is due to the slower increase in the number of Category B defects over time when compared to Category A defects. This is not a consequence of the Weibull parameters used for the defects but rather a consequence of the fact that each Category B defect is replaced by two new welds which then in turn produce their own Category A defects in a repetitive cycle. Therefore, the potential to generate Category A defects increases more rapidly than Category B defects.

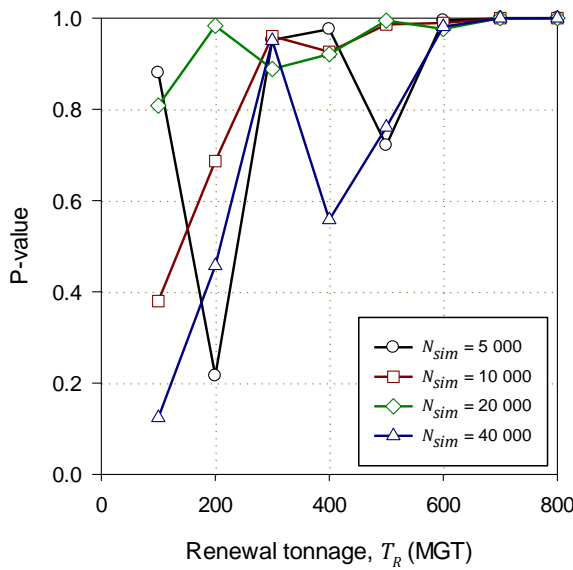


(a)

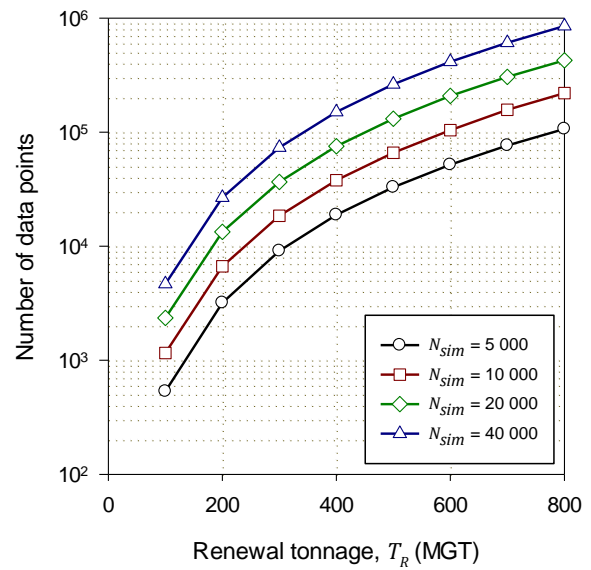


(b)

Figure 4.6: (a) P-value and (b) number of data points used in the calculation thereof at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of P-F interval lengths t_{p-F} for ATWs



(a)



(b)

Figure 4.7: (a) P-value and (b) number of data points used in the calculation thereof at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of P-F interval lengths t_{p-F} for squat defects

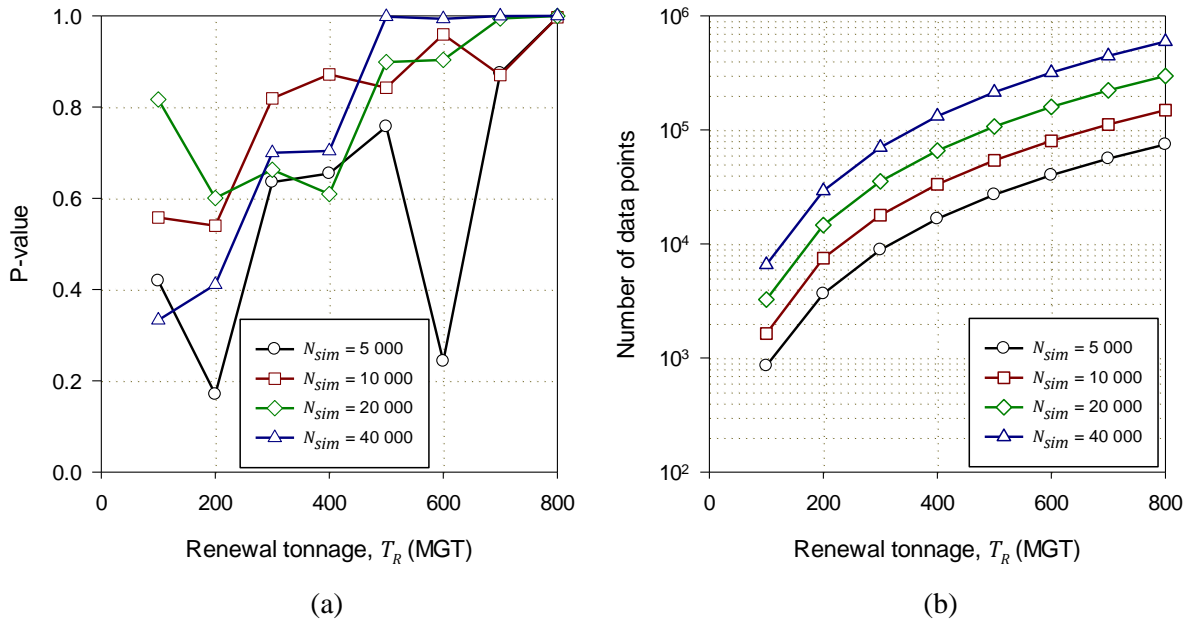


Figure 4.8: (a) P-value and (b) number of data points used in the calculation thereof at the respective renewal tonnages for varying number of simulated life cycles N_{sim} for the distribution of P-F interval lengths t_{P-F} for tache ovale defects

4.3.3 Distribution of Detectability of Defects

The probability of detecting a defect by ultrasonic inspection η is specified for each defect individually. The ratio of the number of times defects were detected against the total number of times defects were passed (and either detected or not detected) by an inspection was investigated for renewal tonnages T_R varying from 100 MGT to 800 MGT in 100 MGT increments. This was done for $N_{sim} = 5\,000$, 10 000, 20 000 and 40 000 virtual life cycles respectively. This is to ensure that the correct probability of detection η is achieved by the simulations.

Subplot (a) of Figure 4.9 through Figure 4.11 shows the proportion of successful inspections as a ratio of the total number of inspections across all defects in the model for ATW defects, squat defects and tache ovale defects respectively. Subplot (b) of Figure 4.9 through Figure 4.11 shows the corresponding total number of defect inspections modelled for each renewal tonnage T_R and number of virtual life cycles N_{sim} considered. No strict mathematical test was conducted on the accuracy of the achieved probability of detection η for each defect. However, through visual inspection it can be seen that the probability of detection η for each defect type rapidly approaches the desired value (see Table 4.1) as T_R and N_{sim} increases. However, significant deviation from the desired values occurs at $T_R = 100$ MGT for $N_{sim} = 5\,000$ and 10 000 virtual life cycles for squat defects (Figure 4.10(a)). This is due to non-

uniform pseudorandom number generation for these cases. The deviation from the desired value is no longer present at $N_{sim} = 20\ 000$ virtual life cycles.

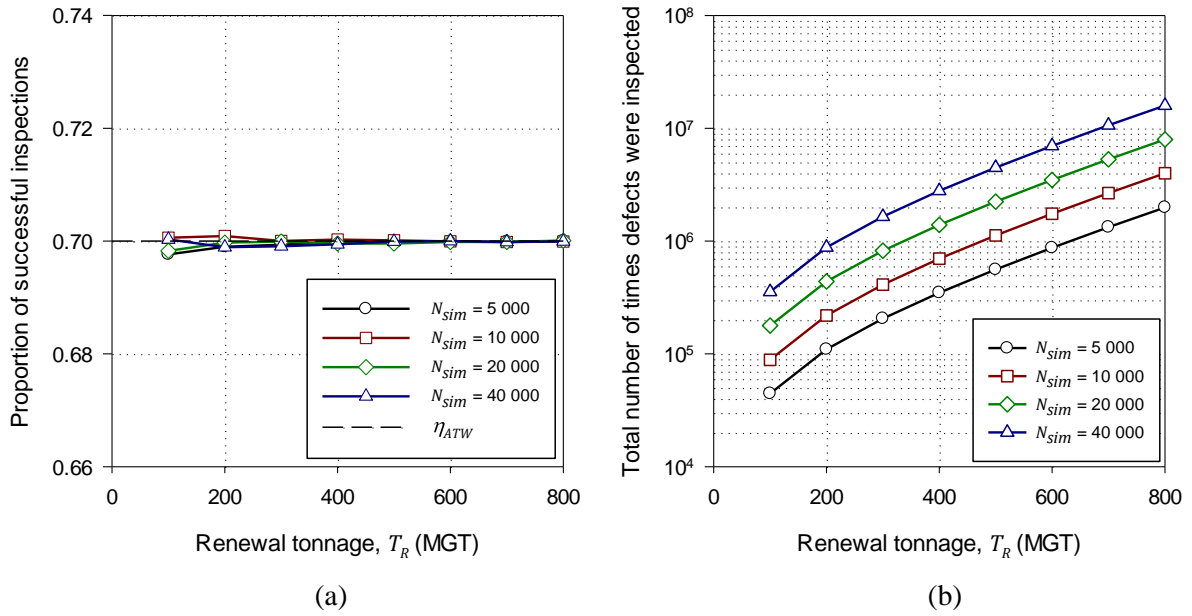


Figure 4.9: (a) Proportion of successful inspections as well as (b) the total number of times defects were inspected at the respective renewal tonnages for varying number of simulated life cycle N_{sim} for ATW defects

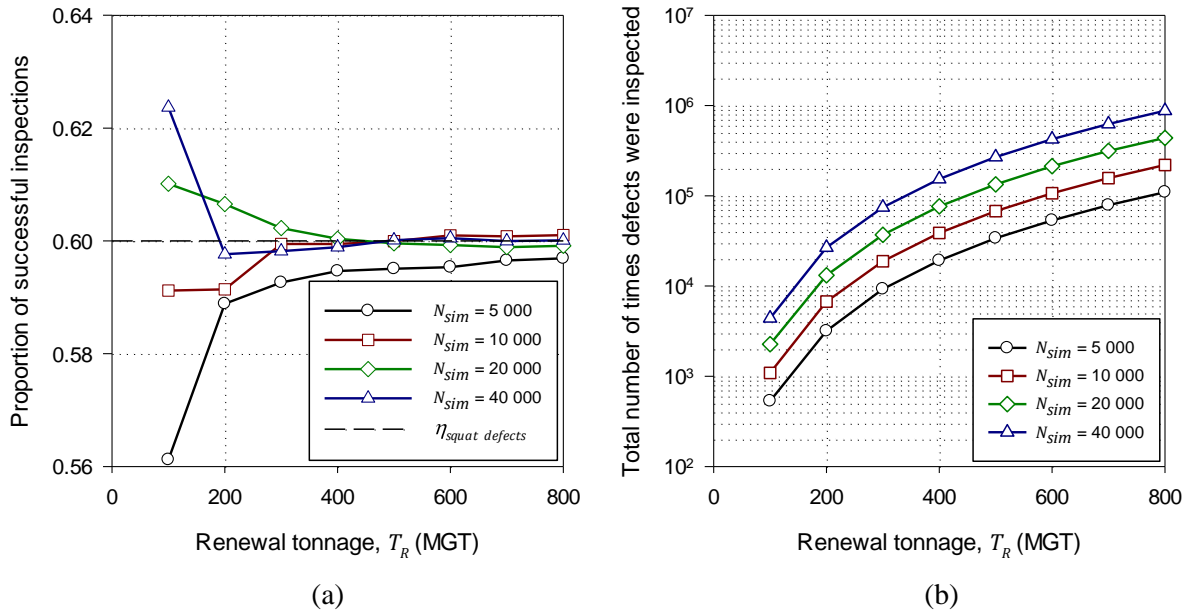


Figure 4.10: (a) Proportion of successful inspections as well as (b) the total number of times defects were inspected (b) at the respective renewal tonnages for varying number of simulated life cycle N_{sim} for squat defects

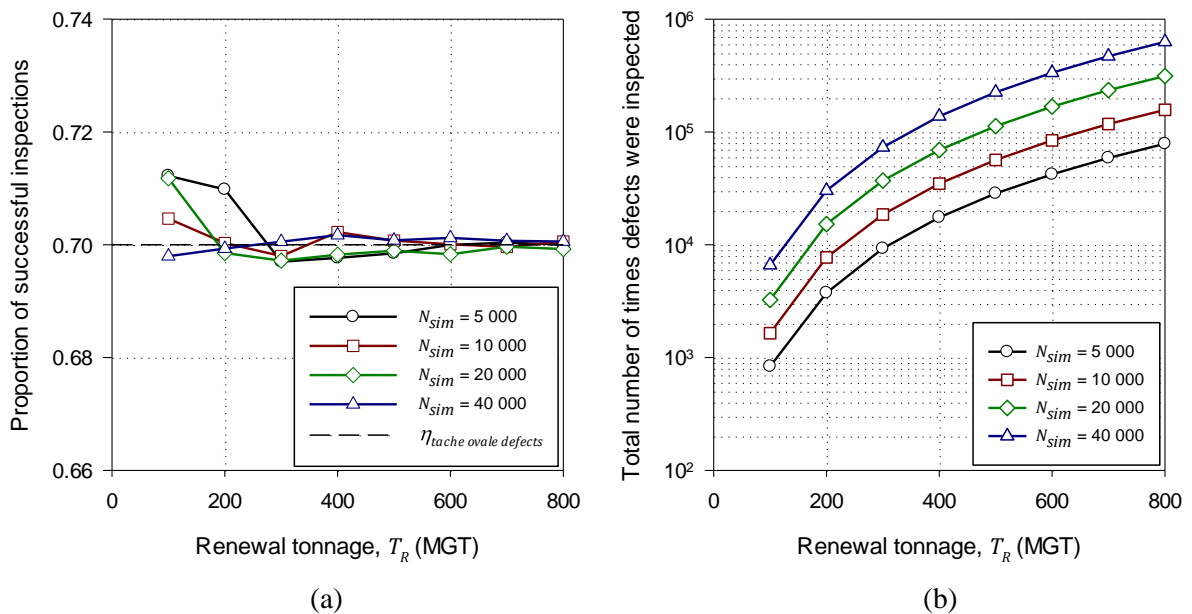


Figure 4.11: Proportion of successful inspections (a) as well as the total number of times defects were inspected (b) at the respective renewal tonnages for varying number of simulated life cycle N_{sim} for tache ovale defects

4.3.4 Accuracy, Number of Simulations and Computational Time

It is noticeable that the adequacy of the posterior distributions wavers at renewal tonnages T_R in the interval (0,200] MGT across all N_{sim} considered in the figures in Sections 4.3.1 through 4.3.3. A single analysis with $N_{sim} = 20\ 000$ virtual life cycles took approximately 35 minutes to run, whereas a single analysis with $N_{sim} = 40\ 000$ virtual life cycles took approximately 1 hour and 15 minutes to run. Therefore, a trade-off had to be made between accuracy and computational time. All subsequent analyses were conducted using $N_{sim} = 20\ 000$ virtual life cycles. This is based on the evidence in the figures presented in Sections 4.3.1 through 4.3.3. Furthermore, the quantity of data produced poses potential problems with regard to the required quantity of random access memory required to analyse and process the data.

4.4 VALIDATION AGAINST AN EXISTING STOCHASTIC MODEL

The results of the developed model were compared with the results of a model developed under similar assumptions by Zhao et al. (2006). The model developed by Zhao et al. (2006) uses the expected value of all input distributions to determine an expected LCC. Thus, the mean LCC is calculated from the results of the model developed for this study and compared to the results obtained using the Zhao et al. (2006) model. The assumptions and limitations of the Zhao et al. (2006) model were thoroughly discussed in Section 2.7.2.

The Zhao et al. (2006) model was programmed into MATLAB. An analysis was run using the parameters for the reference case model as shown in Table 4.1 through Table 4.3. However, the initial number of FBWs $n_{0_{FBW}}$ was set to zero and the initial number of ATWs $n_{0_{ATW}}$ was set to 22. This is due to a modelling limitation of the Zhao et al. (2006) model. Furthermore, the Zhao et al. (2006) model takes into account a derailment probability and a cost of derailment for undetected defects whereas the model developed for this study does not. Thus, the cost of derailment and the probability of derailment were both set to zero in order to make the results comparable. No FBWs were modelled for this analysis as FBWs are treated differently between the two models.

Figure 4.12 shows a comparison of the results obtained for the two models. The mean LCC curves for the two models agree well for renewal tonnages T_R less than 200 MGT. However, the two curves diverge with an increase in renewal tonnage T_R . Upon closer inspection it was found that the maintenance costs (sum of both planned and unplanned maintenance costs) were responsible for this behaviour. This can be seen from the maintenance cost curves (dashed lines) in Figure 4.12 which demonstrate the same divergent behaviour as the LCC curves.

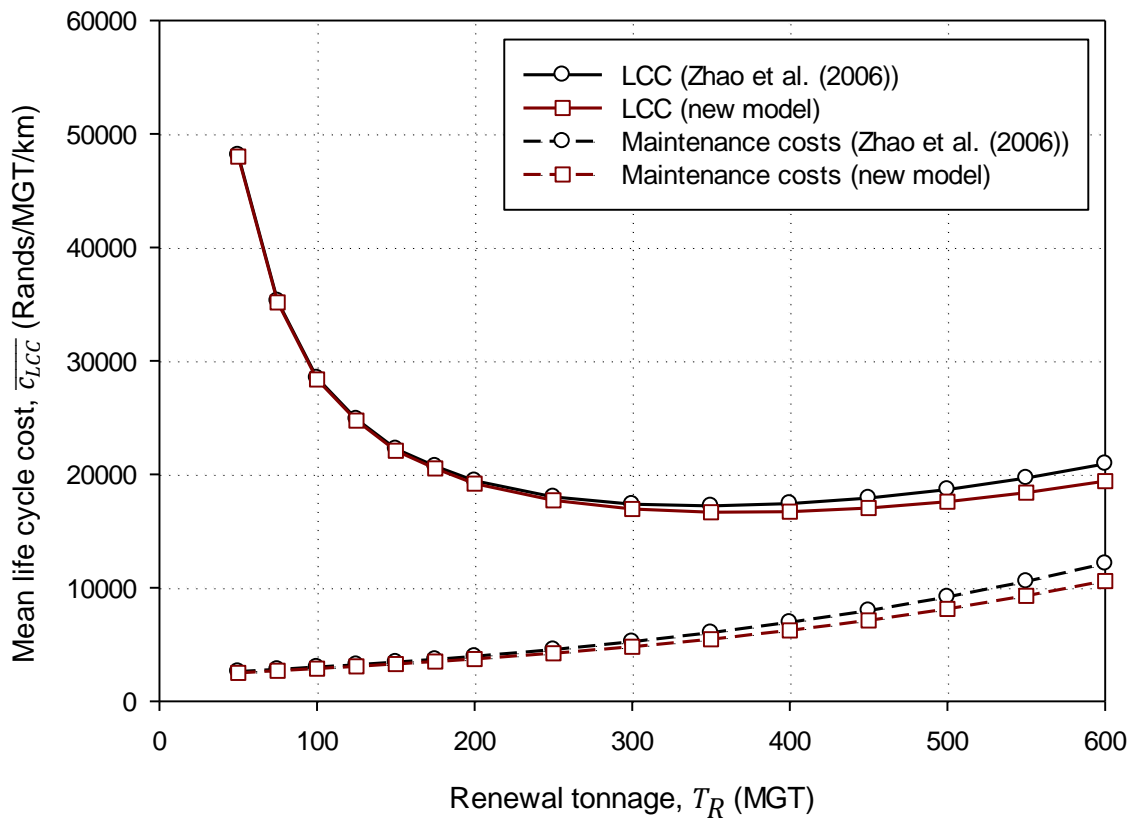


Figure 4.12: Comparison of LCC and maintenance costs calculated from the model developed by Zhao et al. (2006) and the model developed for this study

It is suspected that this illustrated difference between the two models results from an implicit assumption within the Zhao et al. (2006) model. The equations derived by Zhao et al. (2006) assume that the hazard rate of welds (Category A defects) increases in relation to the hazard rate of Category B defects. Whilst this makes logical sense, it is the exact nature of the relationship between these two hazard rates which results in the differences illustrated in Figure 4.12. The derivation by Zhao et al. (2006) is such that the hazard rate of Category A defects increases with *initiation* of Category B defects. In the derivation for this model the hazard rate of Category A defects only increases when the Category B defect is actually maintained and the new welds are installed. It is suspected that it is this slight offset in the hazard rate functions which cumulates with increasing renewal tonnage T_R and causes the diverging behaviour of the curves illustrated in Figure 4.12. Despite this difference, the results of the two models tend to agree reasonably well with each other.

4.5 ASSESSING LIFE CYCLE COST UNCERTAINTY

In Section 4.2 it was shown that an empirical bivariate LCC-renewal tonnage distribution can be calculated in three-dimensional space for a given set of input parameters. Furthermore, it was shown in

Section 4.2 that empirical univariate LCC distributions can be obtained in two-dimensional space if sections of constant renewal tonnage are taken in the LCC-relative frequency plane of the corresponding bivariate distribution. Figure 4.13 shows relative frequency plots for the reference case analysis of the distribution of LCC c_{LCC} for renewal tonnages $T_R = 400, 600$ and 800 MGT respectively. A total of 40 bins were used to group the data. It appears as if the distributions shown in Figure 4.13 belong to a common family of distributions.

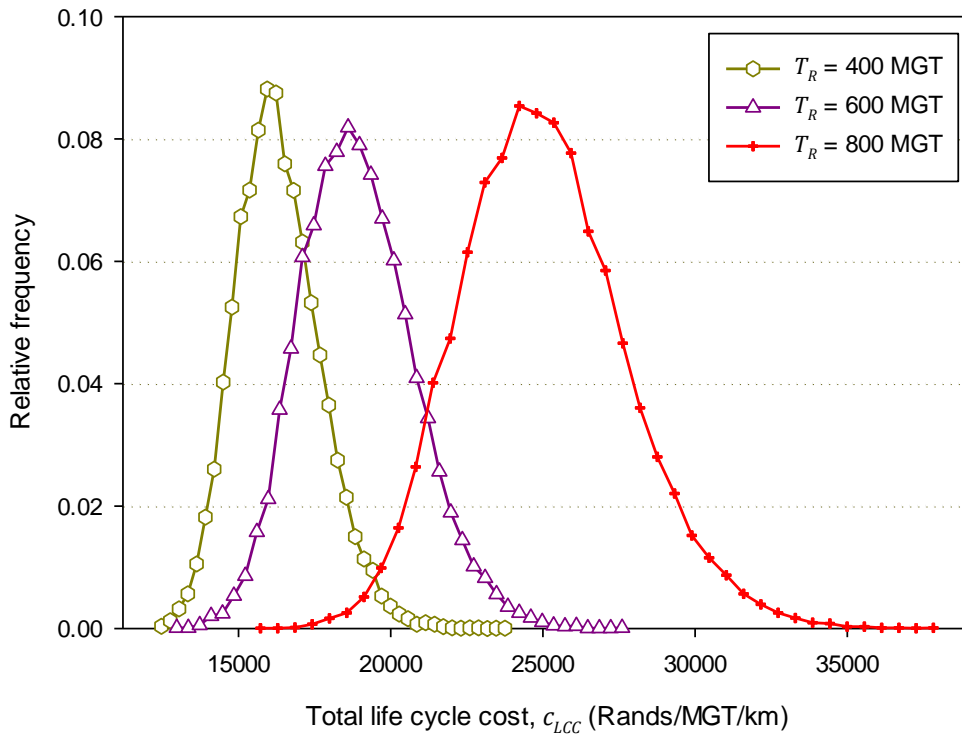


Figure 4.13: Relative frequency plot of observed total LCC c_{LCC} at $T_R = 400, 600$ and 800 MGT respectively

4.5.1 Fitting of a Defined Probability Distribution

The two-parameter normal, lognormal and Weibull probability distributions were fitted to the data using the method of maximum likelihood in an attempt to determine whether the univariate distribution of LCC at a specific renewal tonnage follows a defined family of probability distributions. Thereafter, the Pearson Chi-square test was used to determine the relative goodness-of-fit of the LCC data to the fitted distributions. From the analysis of the posterior distributions conducted in Section 4.3 it can be seen that the posterior distributions most closely followed the desired distributions at larger renewal tonnages. This is expected because the number of data points increases which naturally ensures a more uniform

distribution of sampled pseudorandom numbers. Thus, a renewal tonnage of $T_R = 800$ MGT was selected for testing the fit of the normal, lognormal and Weibull distributions to the data sets.

The probability distributions were fitted to the total LCC data c_{LCC} for different inspection intervals. All other parameters were as specified for the reference case analysis. The P-values of the tests for each inspection interval length s_I and probability distribution type are shown in Table 4.4. The P-values in Table 4.4 indicate that the lognormal probability distribution provides the best fit to the distribution of total LCC c_{LCC} values.

Table 4.4: P-value for fitted distributions at varying inspection intervals lengths for $T_R = 800$ MGT

Inspection interval length, s_I (MGT)	<i>Normal</i>	<i>Lognormal</i>	<i>Weibull</i>
0.1	4.68×10^{-40}	7.48×10^{-15}	0
0.5	1.21×10^{-29}	0.064	0
1.0	4.99×10^{-28}	0.951	0
2.5	2.22×10^{-41}	0.116	0
5.0	4.13×10^{-35}	0.488	0
10.0	1.03×10^{-30}	0.068	0
20.0	3.11×10^{-27}	0.521	0

Figure 4.14 shows the CDFs of the fitted normal, lognormal and Weibull distributions alongside the ECDF for the total LCC c_{LCC} for $T_R = 800$ MGT and $s_I = 1$ MGT. It is clear that the lognormal distribution provides the best fit to the ECDF as indicated by the respective P-value in Table 4.4. Figure 4.15 (a) through (d) show the fitted lognormal distribution and the ECDF for $T_R = 200, 400, 600$ and 800 MGT respectively at $s_I = 1.0$ MGT. The lognormal distributions fit the observed data well and the goodness-of-fit increases with an increase in renewal tonnage T_R .

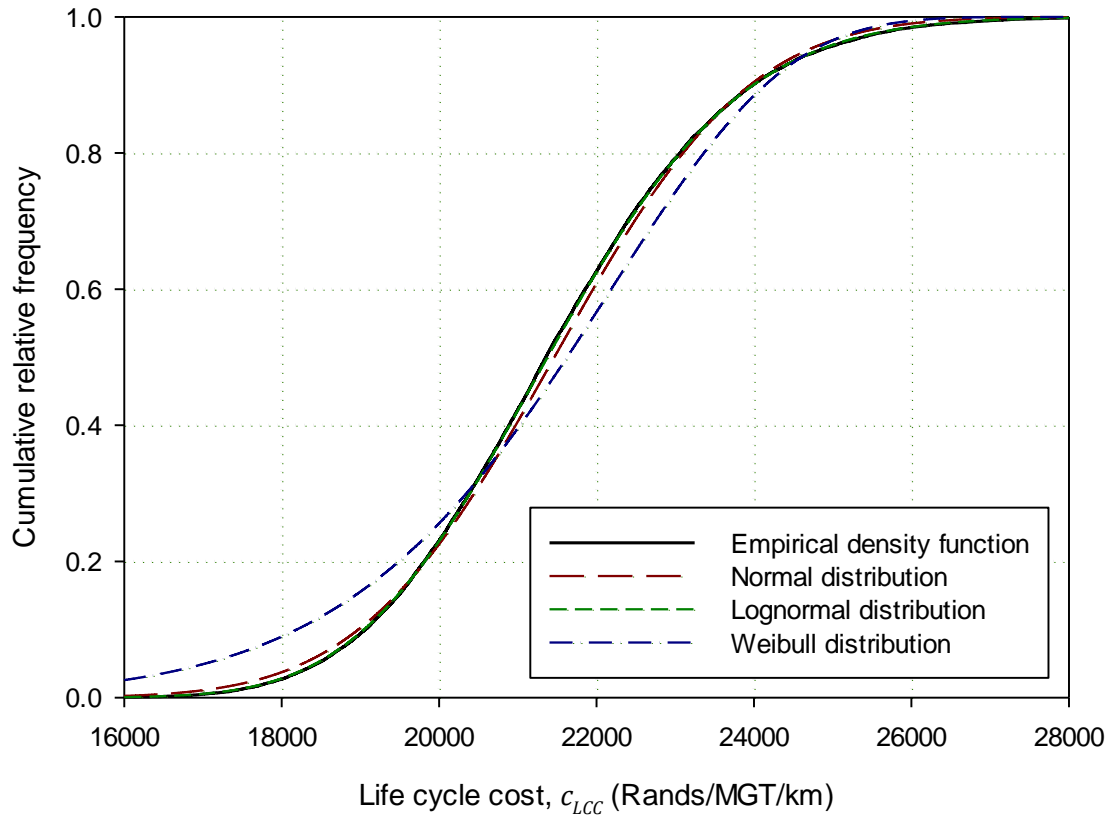
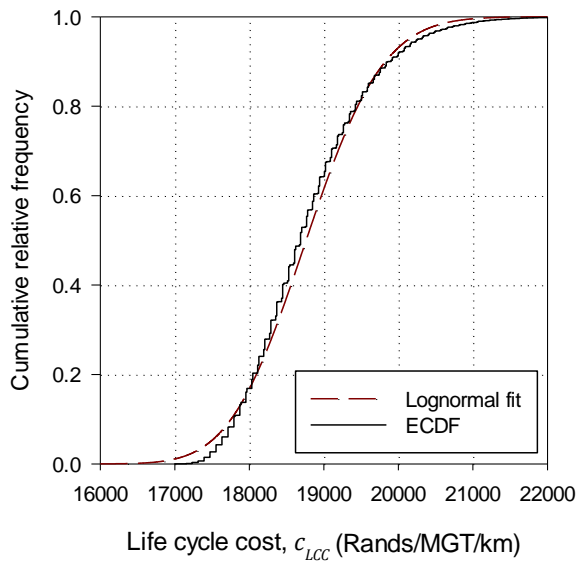
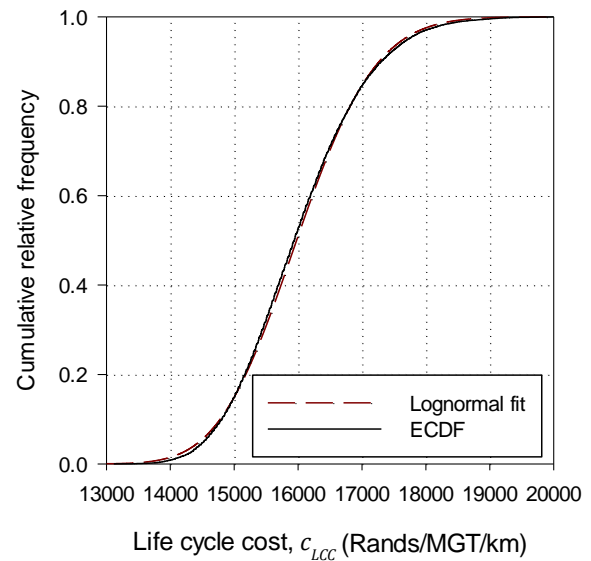


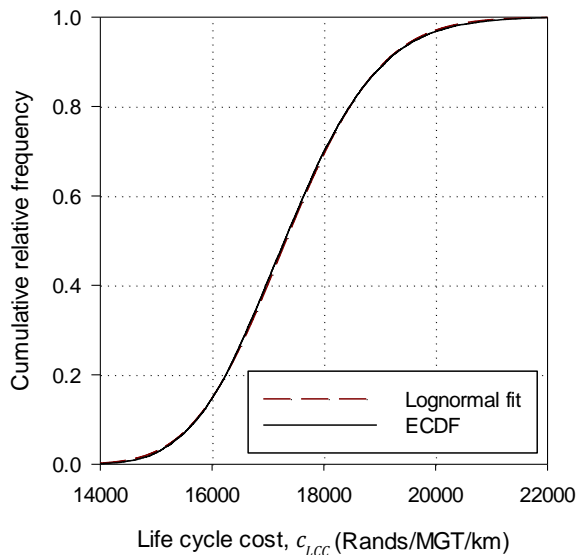
Figure 4.14: Illustration of the goodness-of-fit of different probability distributions to the ECDF of the LCC c_{LCC} for the reference case analysis with $s_I = 1$ MGT and $T_R = 800$ MGT



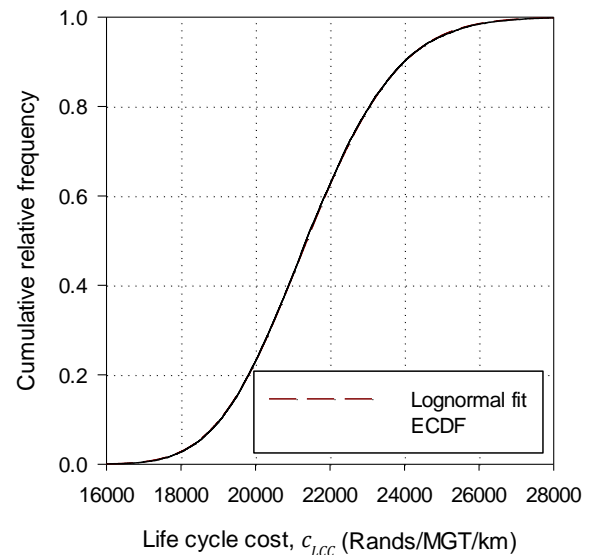
(a)



(b)



(c)



(d)

Figure 4.15: Fitted CDF and the ECDF for the univariate distribution of total LCC c_{LCC} for the reference case analysis with $s_I = 1$ MGT at $T_R =$ (a) 200 MGT, (b) 400 MGT, (c) 600 MGT and (d) 800 MGT

Based on the evidence presented, the lognormal distribution is used to describe the distribution of the LCC data for the remainder of this report. There are no zero values present within the total LCC data c_{LCC} . However, when considering the maintenance component of the LCC data, zeros may be present at small renewal tonnages T_R corresponding to instances where no maintenance occurred. It then becomes impossible to fit a lognormal distribution to the data because the natural logarithm of zero is undefined. Thus, a special case of a Box-Cox transformation is used as shown in Equation 4.2 (Box & Cox, 1964):

$$c' = \ln(c + 1) \quad (4.2)$$

With:

c = the cost data in the natural coordinate system containing zeros

c' = the transformed data which is normally distributed if c is lognormally distributed

The key characteristic of this transformation is that it maps a value of zero in the natural coordinate system to a value of zero in the transformed coordinate system. A normal distribution is then fitted to the transformed data and the mean and standard deviation determined using the method of maximum likelihood. The transformed mean and standard deviation determined from the transformed data c' are subsequently transformed to their natural values using the transformations shown in Equations 4.3 and 4.4 respectively:

$$\mu_{LCC} = e^{\mu'_{LCC} + 0.5\sigma'^2_{LCC}} \quad (4.3)$$

$$\sigma_{LCC} = \mu_{LCC} \cdot (e^{\sigma'^2_{LCC}} - 1) \quad (4.4)$$

With:

μ_{LCC} = the mean expressed using the natural coordinate system

σ_{LCC} = the standard deviation expressed using the natural coordinate system

μ'_{LCC} = the mean expressed using the transformed coordinate system

σ'_{LCC} = the standard deviation expressed using the transformed coordinate system

The remainder of the analyses conducted in this section are concerned with the lognormal distributions fitted to the data. The standard deviation σ_{LCC} of the distribution in the natural coordinate system is assumed to be proportional to the uncertainty present. Lognormal distributions were fitted to the observed distribution of LCC c_{LCC} for renewal tonnages T_R in 100 MGT increments. The resulting probability distributions are plotted in three-dimensional space in Figure 4.16 over the bivariate histogram of the observed LCC data c_{LCC} . The lognormal PDFs were re-scaled in order to align them with the relative frequency values of the bivariate histogram.

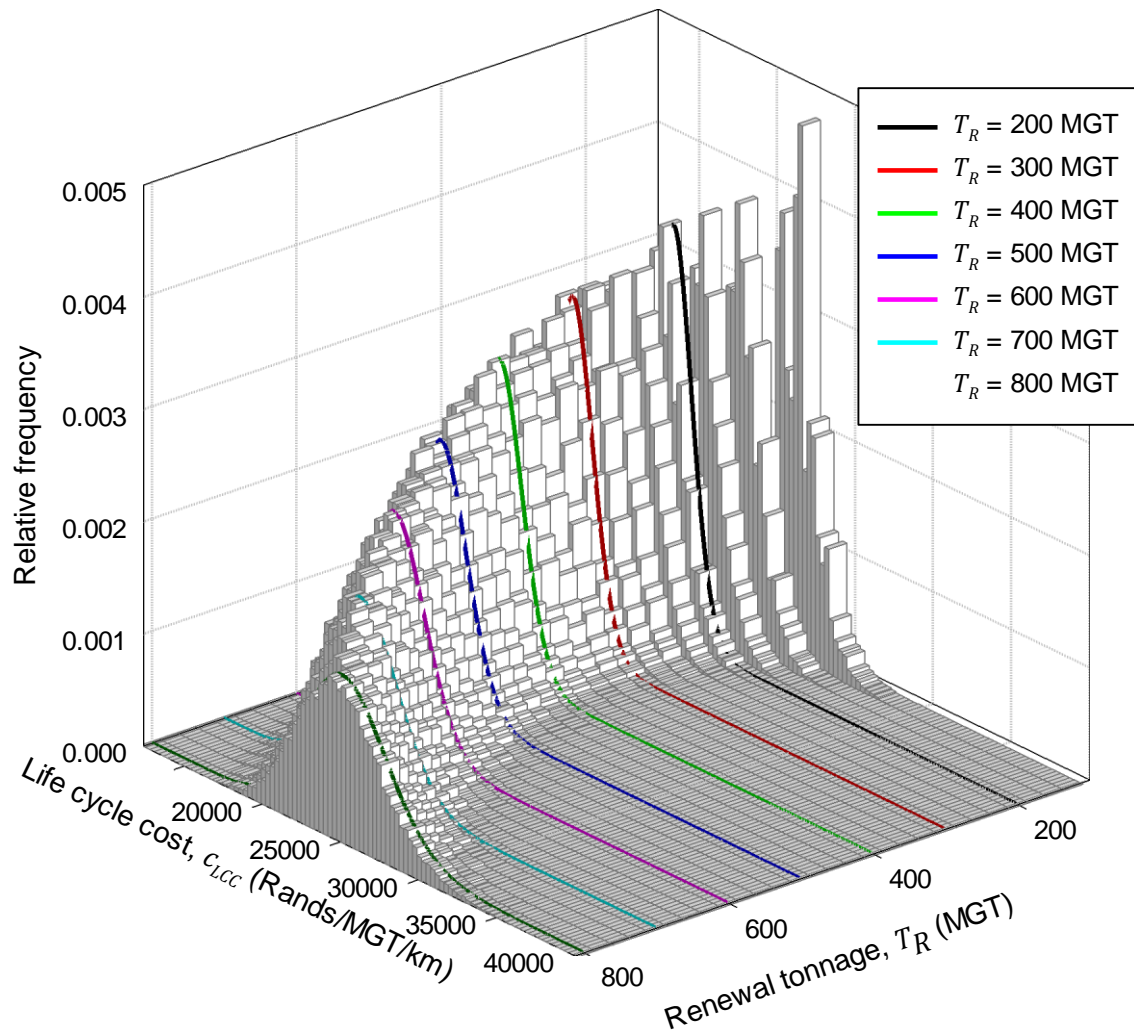


Figure 4.16: Lognormal probability distributions at fixed renewal tonnages fitted over the bivariate histogram of the original data from which they were fitted

Figure 4.17 shows the lognormal distributions in Figure 4.16 as viewed in the LCC-relative frequency plane. It can be seen that the mean μ_{LCC} and standard deviation σ_{LCC} of the distributions do vary.

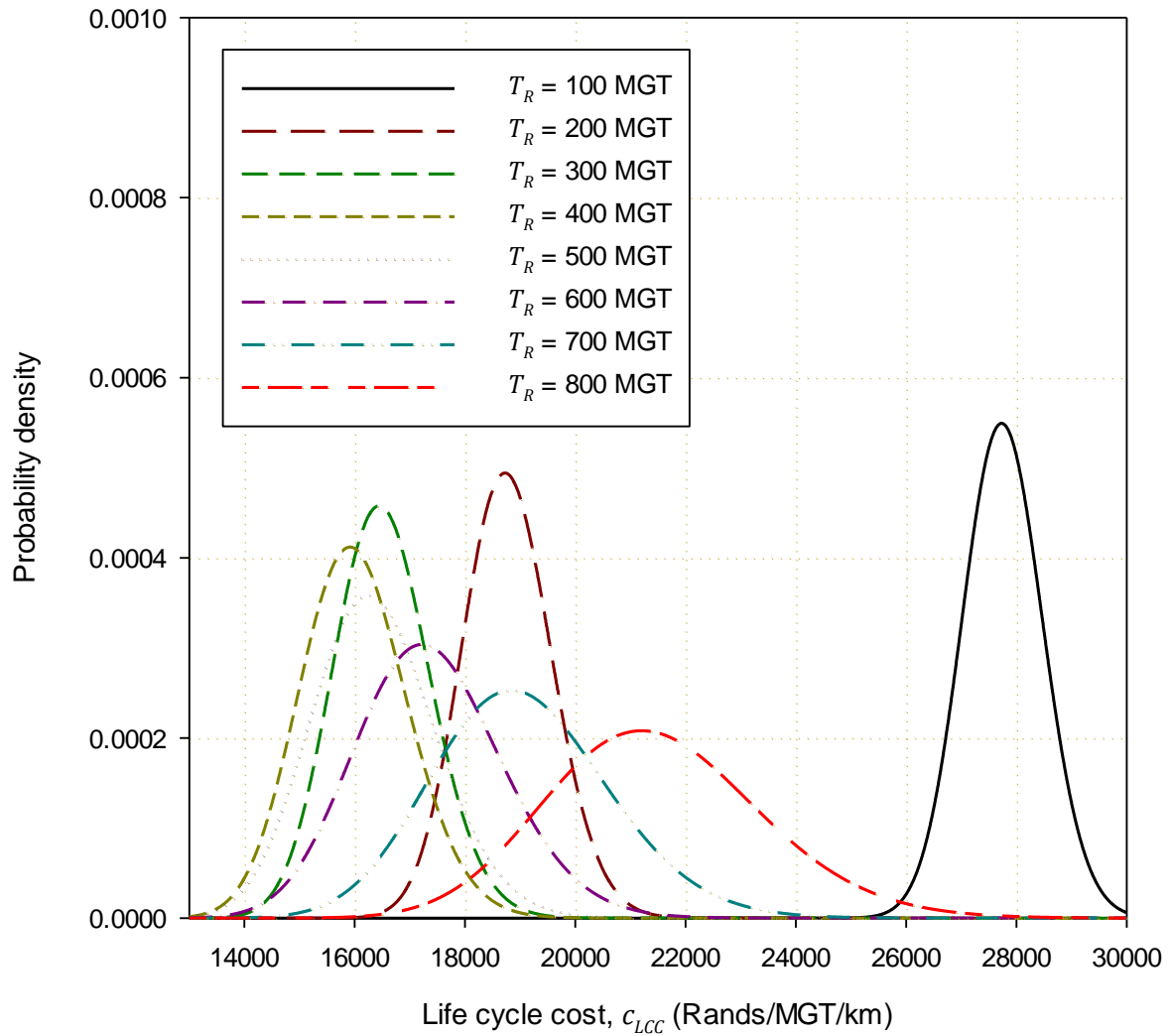


Figure 4.17: Lognormal PDFs at varying renewal tonnages for the reference case analysis with $s_I = 1.0$ MGT

From Figure 4.17 it appears as if the standard deviation of the life cycle cost σ_{LCC} increases with an increase in renewal tonnage. In order to aid in visualising the difference in spread between the distributions, the distributions are normalised against their mean such that their modes all coincide with a value of 1. The result is illustrated in Figure 4.18. From Figure 4.18 it is clear that the uncertainty in the total LCC σ_{LCC} increases with an increase in renewal tonnage T_R . Trends between uncertainty as measured using the standard deviation σ_{LCC} as a metric, renewal tonnage T_R and inspection interval length s_I are explored in Section 4.6.

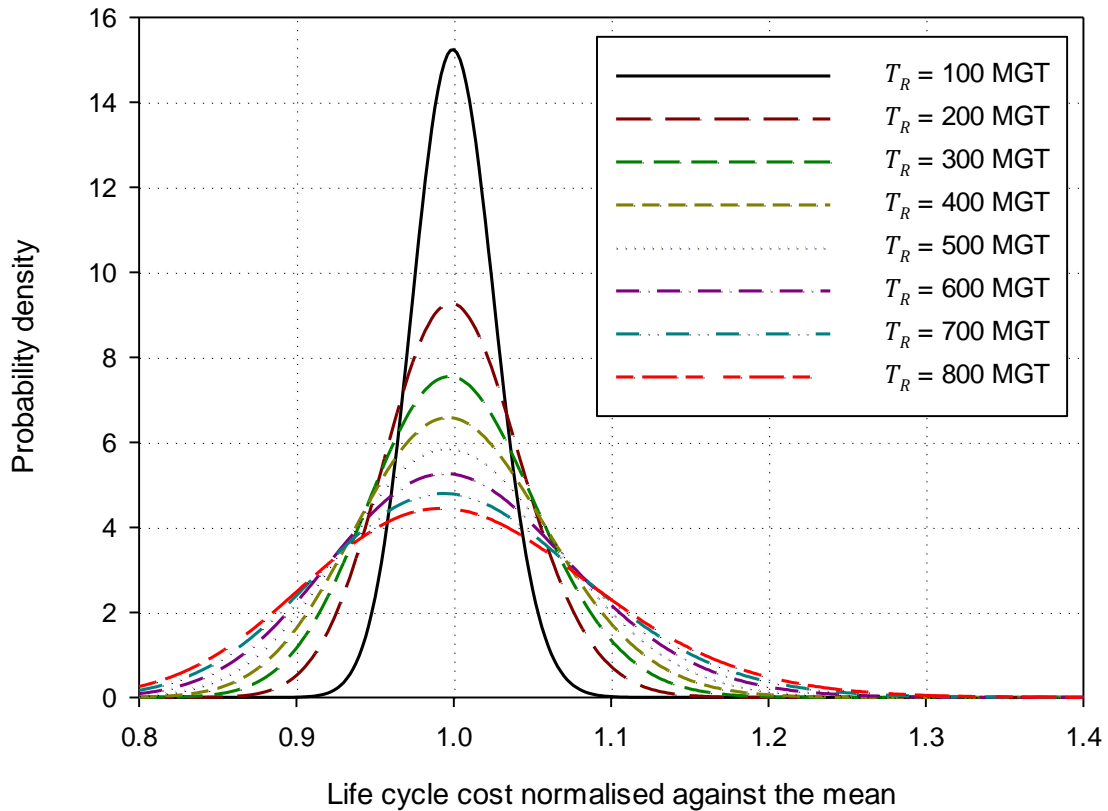


Figure 4.18: Normalised lognormal PDFs at varying renewal tonnages for the reference case analysis with $s_I = 1.0$ MGT

4.6 INFLUENCE OF INSPECTION INTERVALS

This section is divided into two subsections, focusing on the influence of inspection intervals. Section 4.6.1 is concerned with the mean LCC and Section 4.6.2 is concerned with the distribution of LCC.

4.6.1 Influence of Inspection Intervals on the Mean Life Cycle Cost

The influence of the inspection interval length s_I on the resulting mean LCC curve was investigated. The range of inspection intervals investigated was $s_I = 0.1, 0.5, 5, 10$ and 20 MGT. All other parameters were as specified for the reference case analysis. The resulting mean LCC curves are shown in Figure 4.19. It can be seen that an optimal inspection interval is attainable which would minimise the mean LCC. Inspection interval lengths larger or smaller than optimum raise the minimum attainable mean LCC $\overline{c_{LCC_{min}}}$. The minimum attainable mean LCC is extremely sensitive to short inspection intervals as can be seen for the case where $s_I = 0.1$ and 0.2 MGT. However, shorter inspection intervals produce a flatter mean LCC curve allowing a longer window period for renewal operations whilst maintaining a

near optimum mean LCC $\overline{c_{LCC}}$. This observation was also noted by Zhao et al. (2006). For larger inspection intervals, the LCC curve has a higher curvature and a shorter optimal window period for renewal operations. Furthermore, shorter inspection intervals shift the optimum renewal tonnage to a larger value and visa versa.

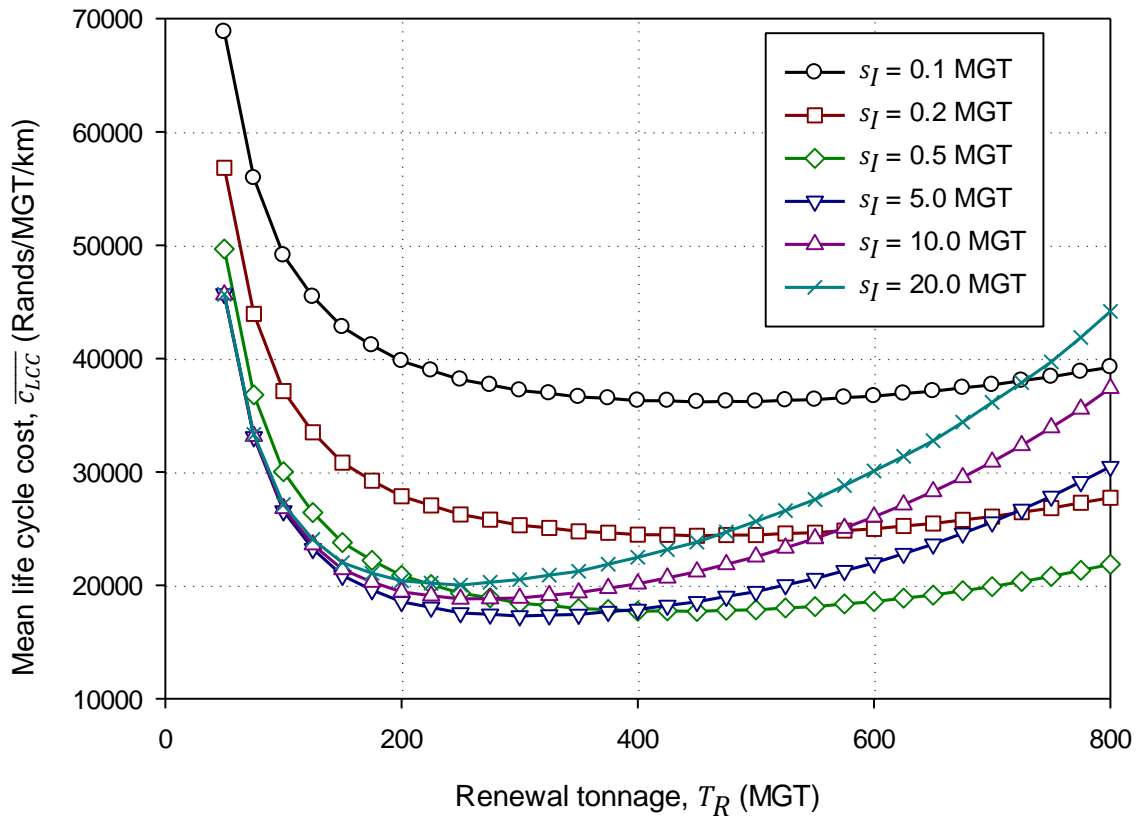


Figure 4.19: The effect of the inspection interval s_I on the resulting mean LCC curve

The mean LCC shown in Figure 4.19 results from a combination of a decreasing renewal cost and an increasing maintenance cost with an increasing renewal tonnage T_R . The maintenance costs themselves can be divided into a planned maintenance cost originating from planned maintenance activities and an unplanned maintenance cost originating from unplanned maintenance activities. The combined maintenance cost is then the sum of the planned and unplanned maintenance cost. The planned, unplanned and combined maintenance cost curves are shown in Figure 4.20 for an inspection interval length $s_I = 0.5$ and 10.0 MGT. The combined maintenance cost at an inspection interval length $s_I = 10.0$ MGT is higher than the combined maintenance cost for an inspection interval length $s_I = 0.5$ MGT. However, it also appears as if the proportion of unplanned maintenance cost in relation to the combined maintenance cost is higher for the larger inspection interval length.

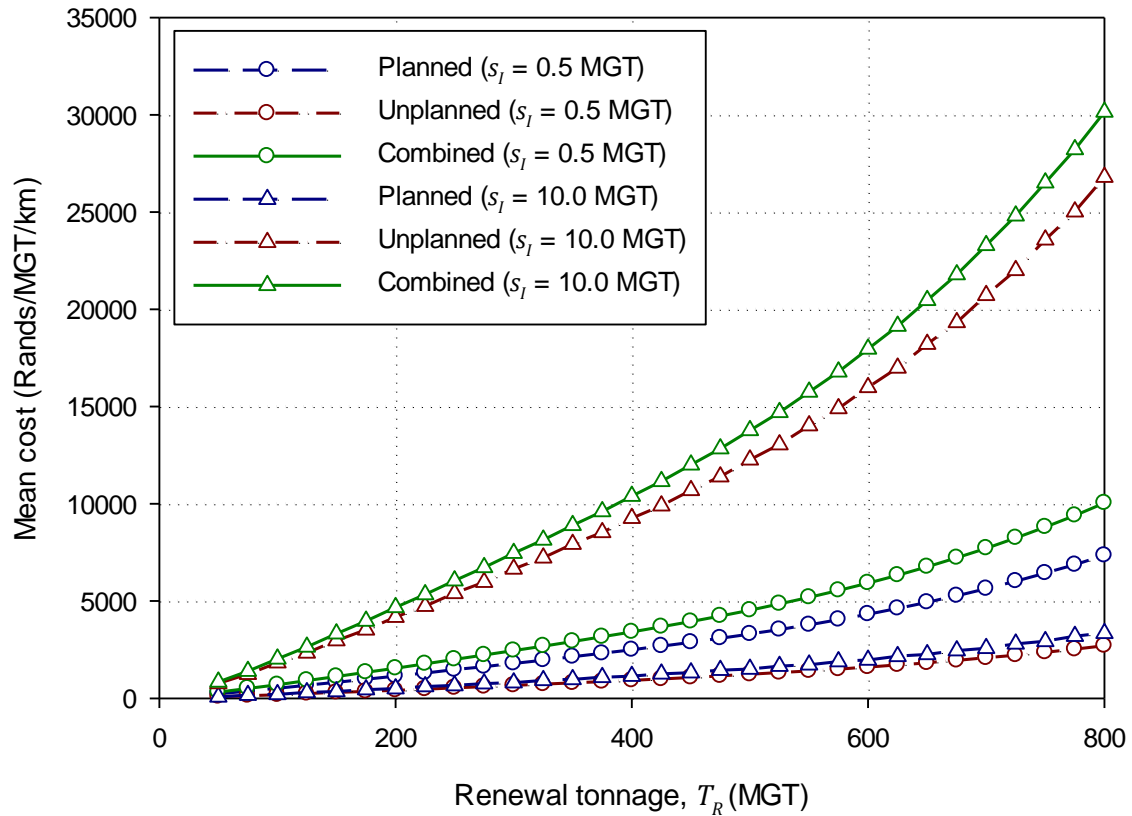


Figure 4.20: Planned, unplanned and combined maintenance costs for inspection interval lengths $s_I = 0.5$ and 10.0 MGT

To investigate the relative contribution towards the total maintenance cost of the planned and unplanned components respectively, the planned maintenance cost and unplanned maintenance cost were expressed as a percentage of the combined maintenance cost for each specific renewal tonnage T_R . It was found that this ratio was constant except for insignificant random fluctuation. Figure 4.21 shows the average contribution of the planned and unplanned maintenance costs towards the combined maintenance cost for inspection interval lengths of $s_I = 0.5$ and 10.0 MGT. It is clear that the larger combined maintenance cost associated with a larger inspection interval length $s_I = 10.0$ MGT is caused by a significant increase in the relative proportion of unplanned maintenance activities.

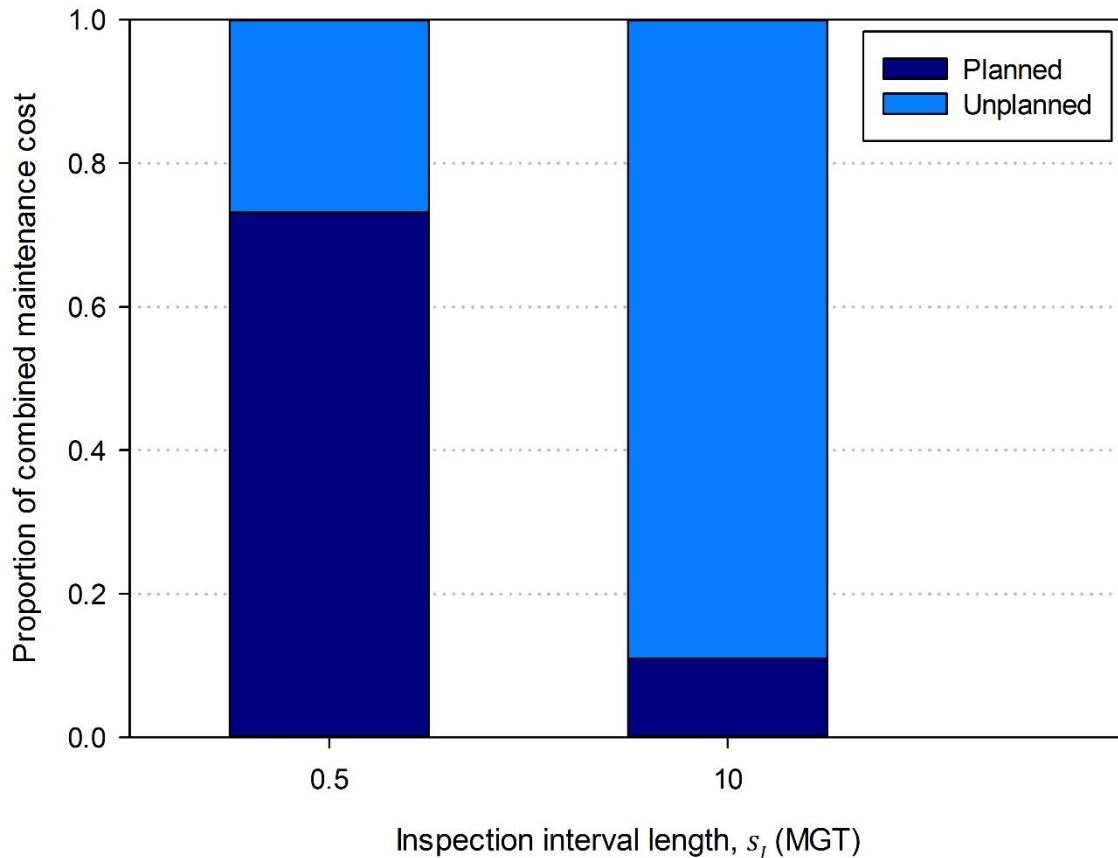


Figure 4.21: Proportion of combined maintenance cost originating from planned and unplanned maintenance activities for inspection interval lengths $s_I = 0.5$ and 10.0 MGT respectively

4.6.2 Influence of Inspection Intervals on Life Cycle Cost Uncertainty

Figure 4.22 shows the same LCC curves shown in Figure 4.19 but only for $s_I = 0.1$ and 5.0 MGT. Also shown in the figure are the 1st and 99th percentiles. It is evident from the diverging nature of the 1st and 99th percentile plots that the variability in the LCC c_{LCC} increases with renewal tonnage T_R as was demonstrated in Figure 4.18. Furthermore, the interquartile distance at a given renewal tonnage is larger for $s_I = 5.0$ MGT than for $s_I = 0.1$ MGT.

In order to investigate this phenomenon further, the fitted normalised lognormal distributions at $T_R = 800$ MGT were compared for $s_I = 0.1, 0.5, 1.0, 2.5, 5.0, 10.0$ and 20.0 MGT respectively. These plots are shown in Figure 4.23. It is clear that the standard deviation and hence the uncertainty increases with an increase in the inspection interval length s_I .

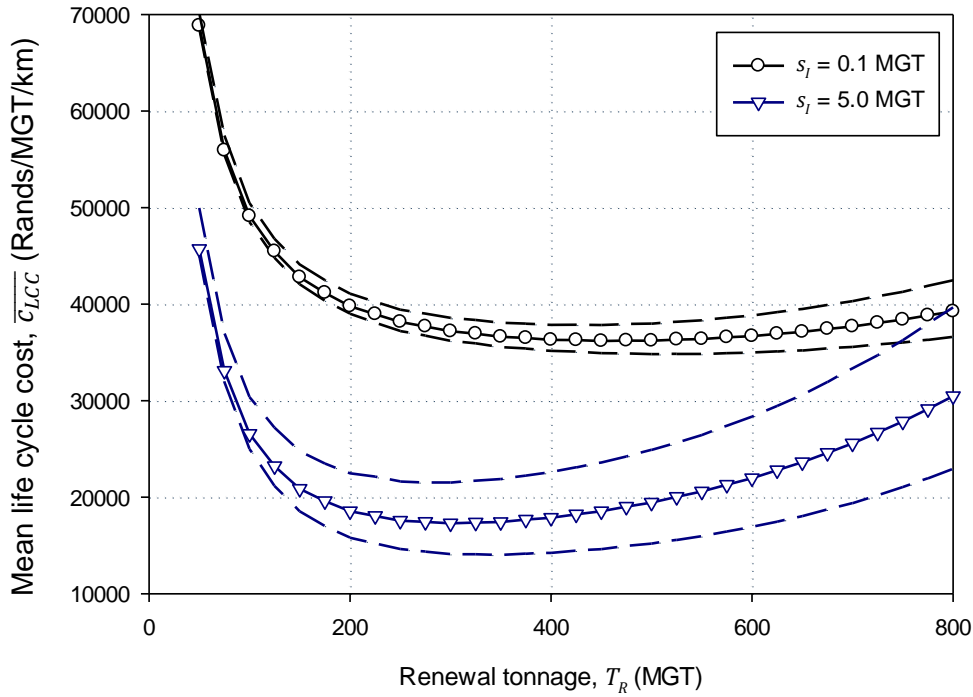


Figure 4.22: Mean LCC curve showing the 1st and 99th percentiles for $s_I = 0.1$ and 5.0 MGT

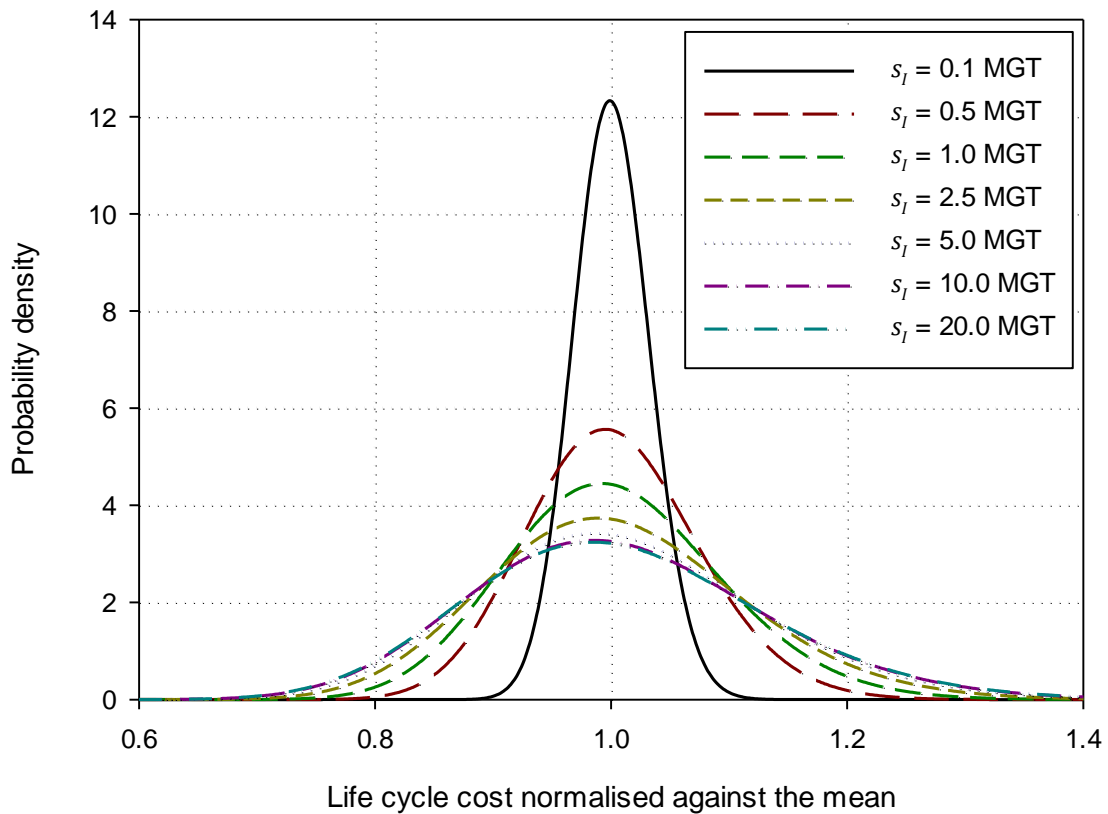


Figure 4.23: Uncertainty in the form of standard deviation and its relationship with varying inspection interval lengths s_I as measured using a standard lognormal PDF normalised against the mean LCC at $T_R = 800$ MGT

If the standard deviation σ_{LCC} is plotted versus renewal tonnage T_R at varying inspection intervals, the plots in Figure 4.24 are produced. Figure 4.24 clearly indicates that the standard deviation of the LCC σ_{LCC} increases with an increasing renewal tonnage across all inspection interval lengths considered in this study. The plots in Figure 4.24 all demonstrate an inflection point after which the rate of increase in σ_{LCC} becomes positive.

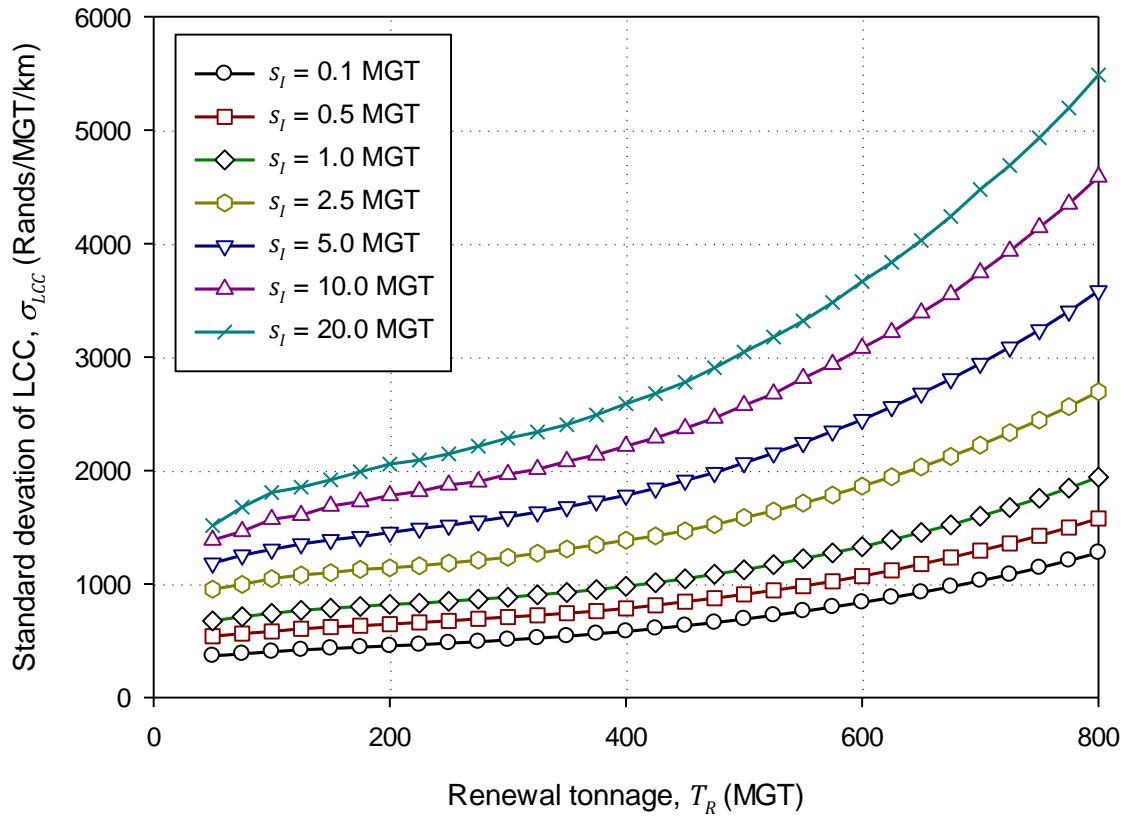


Figure 4.24: Influence of renewal tonnage T_R on total LCC uncertainty σ_{LCC} at varying inspection intervals s_I

Figure 4.25 illustrates the same data as Figure 4.24 but on a set of axes of standard deviation σ_{LCC} versus inspection interval length s_I . Plotting the data in such a manner clearly demonstrates the increasing uncertainty in the total LCC σ_{LCC} with an increase in inspection interval length s_I .

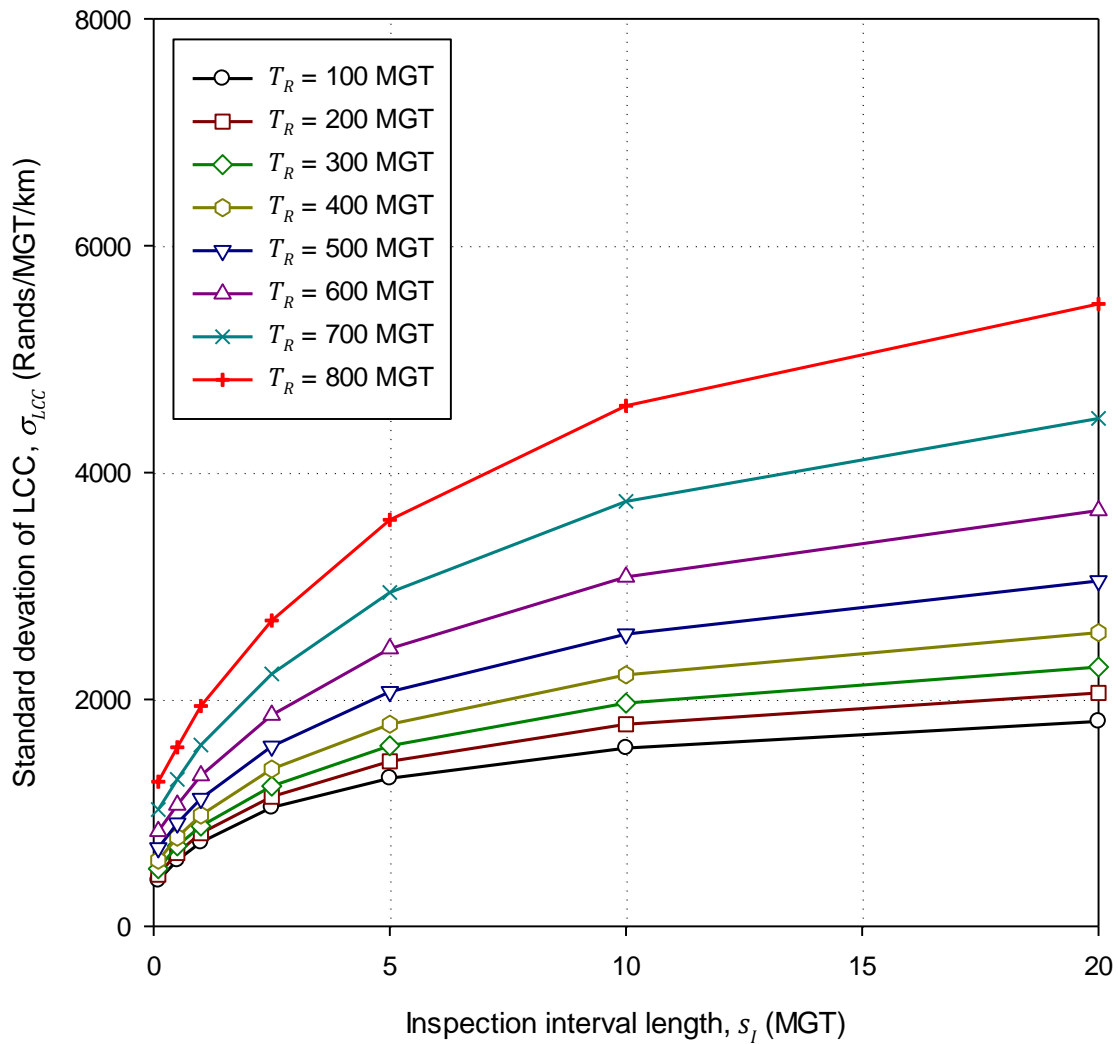


Figure 4.25: Influence of inspection interval length s_I on LCC uncertainty σ_{LCC} at varying renewal tonnages T_R

The total LCC c_{LCC} has both deterministic and stochastic components. The stochastic components consist of the planned c_{LCC_p} and unplanned c_{LCC_f} maintenance costs both with their associated uncertainty represented by the standard deviation σ_{LCC_p} and σ_{LCC_f} respectively. Figure 4.26 illustrates the development of the uncertainty of planned maintenance costs σ_{LCC_p} with an increase in renewal tonnage T_R . The uncertainty σ_{LCC_p} initially increases rapidly before dropping to a minimum after which a steady increasing trend is observed.

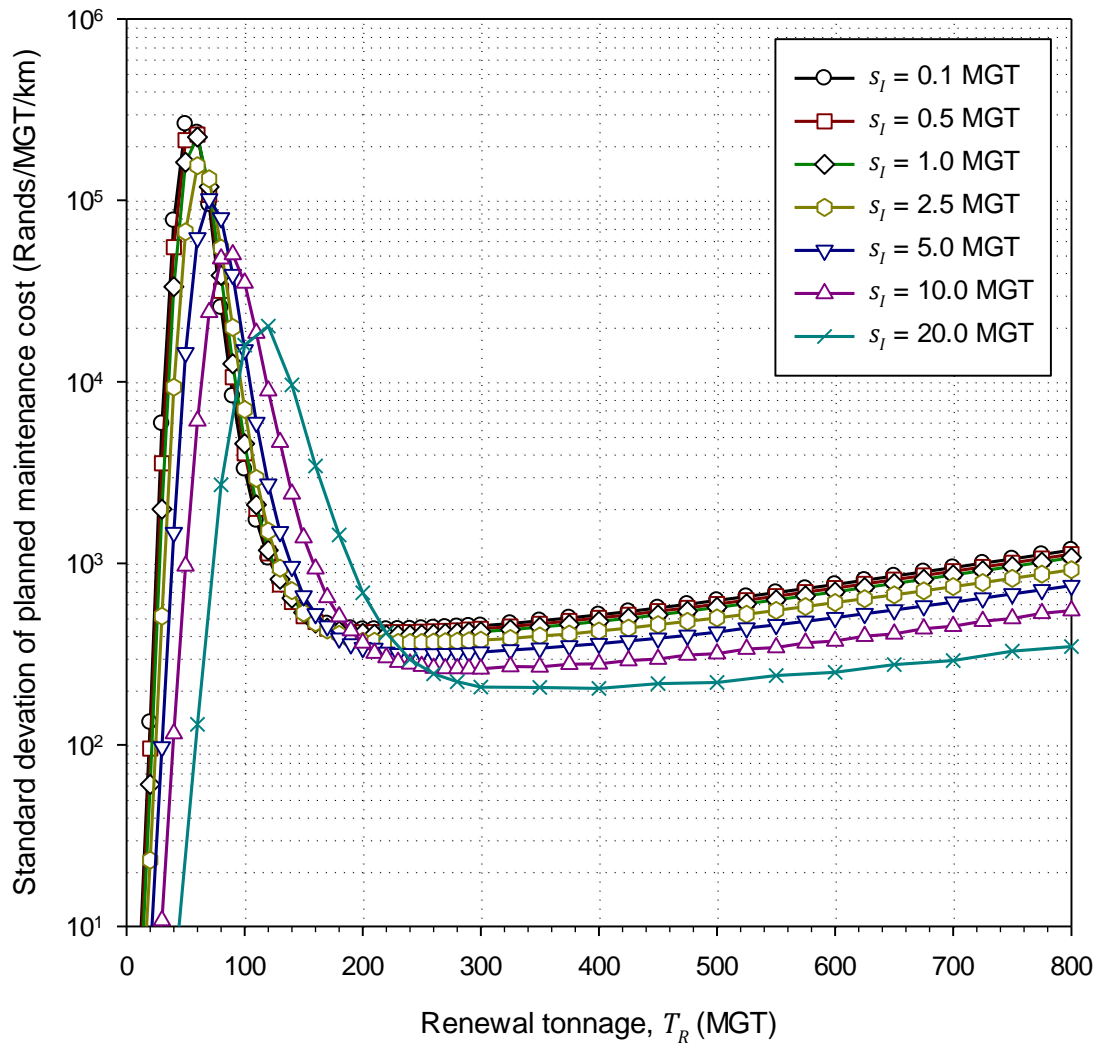


Figure 4.26: Influence of renewal tonnage T_R on planned maintenance cost uncertainty at varying inspection intervals s_I

The behaviour of the curves in Figure 4.26 for a renewal tonnage interval (0,200) is better observed on the re-scaled set of axes in Figure 4.27. The initial rapid increase in σ_{LCC_p} is attributed to an increase in the number of maintenance activities conducted. The maximum uncertainty in planned maintenance costs increases for a respective decrease in inspection interval length s_I . This is because the number of planned maintenance activities increases with a decrease in inspection interval length s_I . Figure 4.26 demonstrates that the planned maintenance cost uncertainty σ_{LCC_p} in the steady region of the curves is less for a larger inspection interval length s_I .

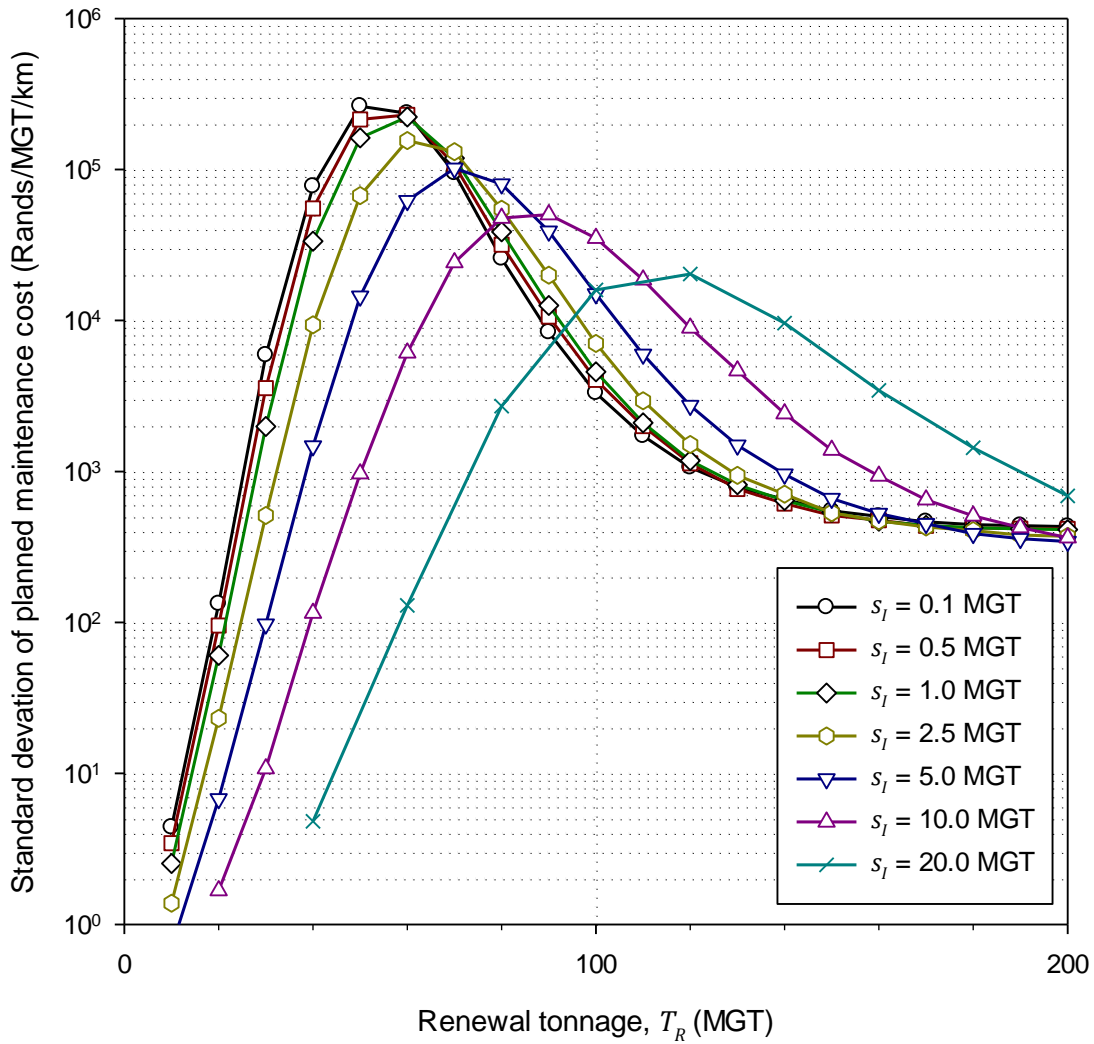


Figure 4.27: Influence of renewal tonnage T_R on planned maintenance cost uncertainty at varying inspection intervals s_I on the renewal tonnage interval (0, 200) to illustrate the curve behaviour

Figure 4.28 illustrates the development of the uncertainty of unplanned maintenance costs σ_{LCC_f} with an increase in renewal tonnage T_R . The trend for a given inspection interval length s_I in relation to an increasing renewal tonnage T_R is similar to that displayed by the uncertainty in planned maintenance costs σ_{LCC_p} in Figure 4.26. The difference between the planned maintenance cost uncertainty σ_{LCC_p} and unplanned maintenance cost uncertainty σ_{LCC_f} curves shown in Figure 4.26 and Figure 4.28 respectively is that in the steady region of the curves the uncertainty of unplanned maintenance costs σ_{LCC_f} is less for a shorter inspection interval length s_I . This is opposite to that of the planned maintenance costs. This behaviour is a direct result of the proportion of planned to unplanned maintenance activities for a given inspection interval length s_I as was shown in Figure 4.21.

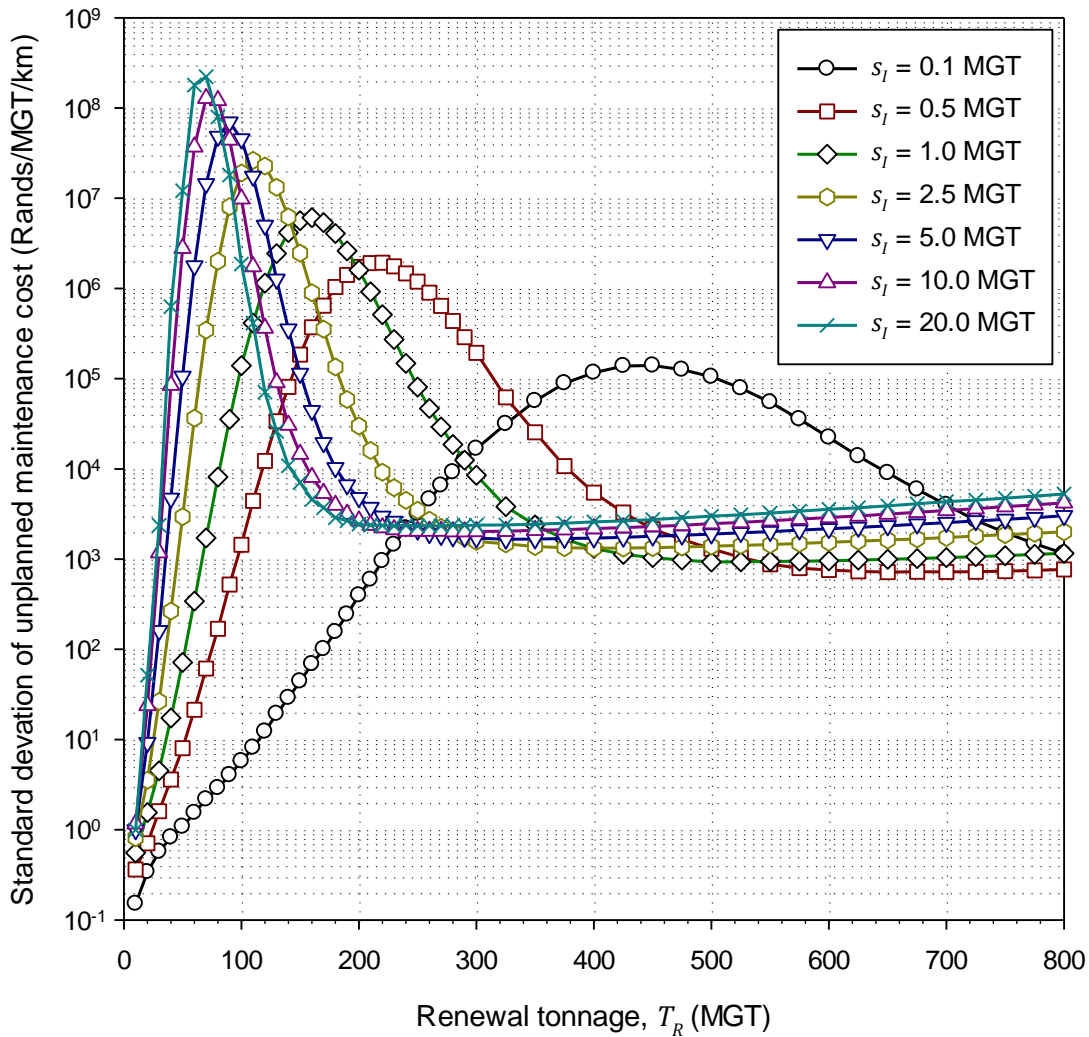


Figure 4.28: Influence of renewal tonnage T_R on unplanned maintenance cost uncertainty at varying inspection intervals s_I

The combined stochastic maintenance cost component $c_{LCC_{combined}}$ comprises the sum of planned c_{LCC_p} and unplanned c_{LCC_f} maintenance costs. These two costs interact to form a combined maintenance cost uncertainty $\sigma_{LCC_{combined}}$. The combined maintenance cost uncertainty is not a straightforward summation of the two underlying uncertainties. It is influenced by the ratio of planned to unplanned maintenance actions as well as the ratio of planned to unplanned maintenance costs. Figure 4.29 (a) and (b) shows the planned σ_{LCC_p} , unplanned σ_{LCC_f} and combined $\sigma_{LCC_{combined}}$ maintenance cost uncertainty for inspection intervals $s_I = 0.1$ and 10.0 MGT respectively. The figure illustrates how the uncertainty of the combined maintenance cost $\sigma_{LCC_{combined}}$ for the smaller inspection interval $s_I = 0.5$ MGT is more strongly influenced by the uncertainty of the planned maintenance costs σ_{LCC_p} . The combined maintenance cost uncertainty $\sigma_{LCC_{combined}}$ tends more strongly towards the planned maintenance cost

uncertainty σ_{LCC_p} in Figure 4.29 (a) and more strongly towards the unplanned maintenance cost uncertainty σ_{LCC_f} in Figure 4.29 (b).

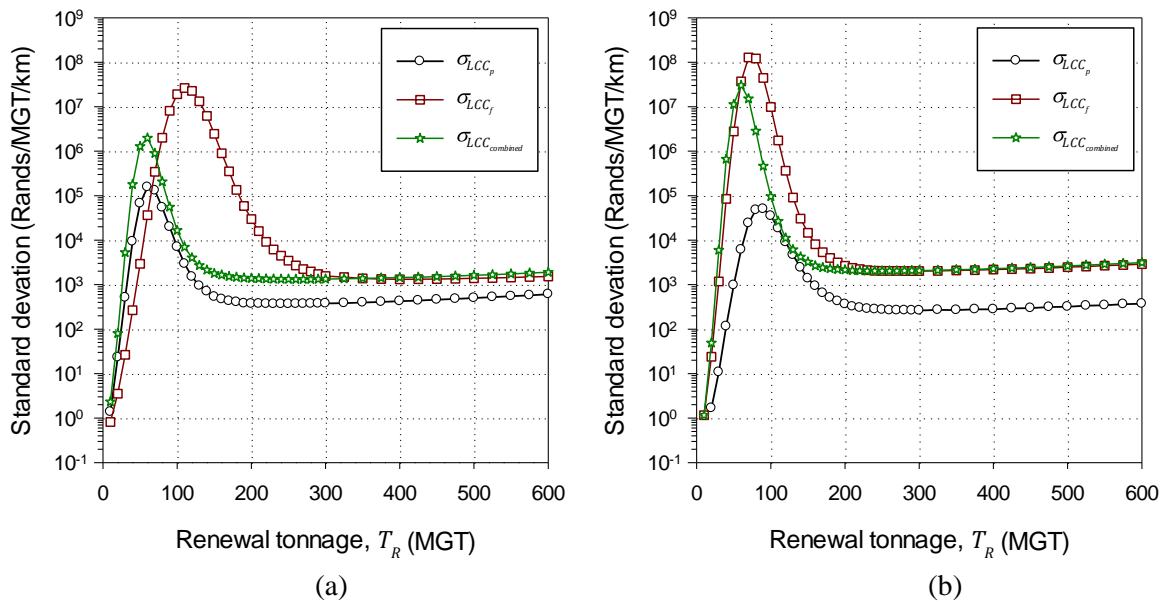


Figure 4.29: Planned, unplanned and combined maintenance cost standard deviation for (a) $s_I = 0.5$ MGT and (b) $s_I = 10$ MGT

The combined maintenance cost uncertainty $\sigma_{LCC_{combined}}$ for all inspection intervals considered is plotted in Figure 4.30. The general trend in the curves remain the same as for the planned and unplanned maintenance cost uncertainties. In the steady region ($T_R > 250$ MGT) the standard deviation of the combined maintenance cost $\sigma_{LCC_{combined}}$ is higher for a larger inspection interval length s_I . This relationship carries over to the total LCC uncertainty σ_{LCC} as shown in Figure 4.25.

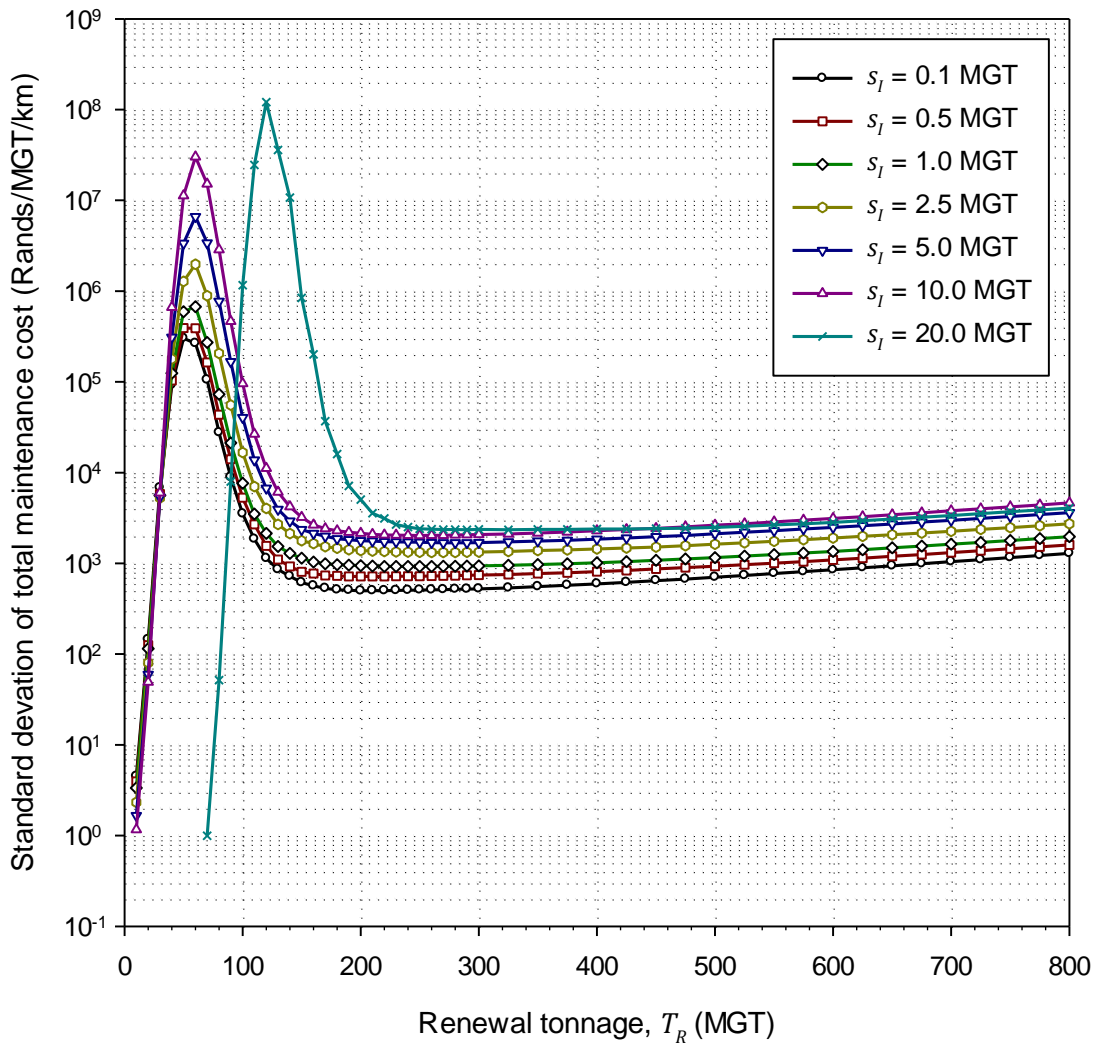


Figure 4.30: Influence of renewal tonnage T_R on total maintenance cost uncertainty at varying inspection intervals s_I

4.7 SELECTING A WELD TYPE TO USE FOR RAIL MAINTENANCE

The viability of using mobile flash butt welding versus alumino-thermic welding to conduct maintenance on the rail was investigated. Mobile flash butt welding is more costly to conduct due to the large and expensive equipment required. However, the integrity of FBWs is superior to that of ATWs. The superiority of FBWs over ATWs is taken into account in the model using the Weibull parameters which model the hazard function of the welds. Figure 4.31 (a) and (b) shows the PDF and CDF respectively for the Weibull probability distribution with the parameters as used to model the hazard function of ATWs and FBWs. From the PDF it is noticeable that a higher probability exists for early failures in ATWs. However, upon inspection of the CDF it can be seen that at $T_R = \pm 250$ MGT the probability of a FBW defect occurring becomes greater than the probability of an ATW defect occurring.

This phenomenon is due to the high concentration of probability density of the FBW PDF at $T_R = \pm 200$ MGT. In order to more accurately model the superior integrity of FBWs throughout their life, a dual PDF approach may be used as suggested by Zhao et al. (2007). The one PDF models infant mortality of the ATWs and the other models the fatigue defect life of the ATWs. However, this approach has not been considered in this study.

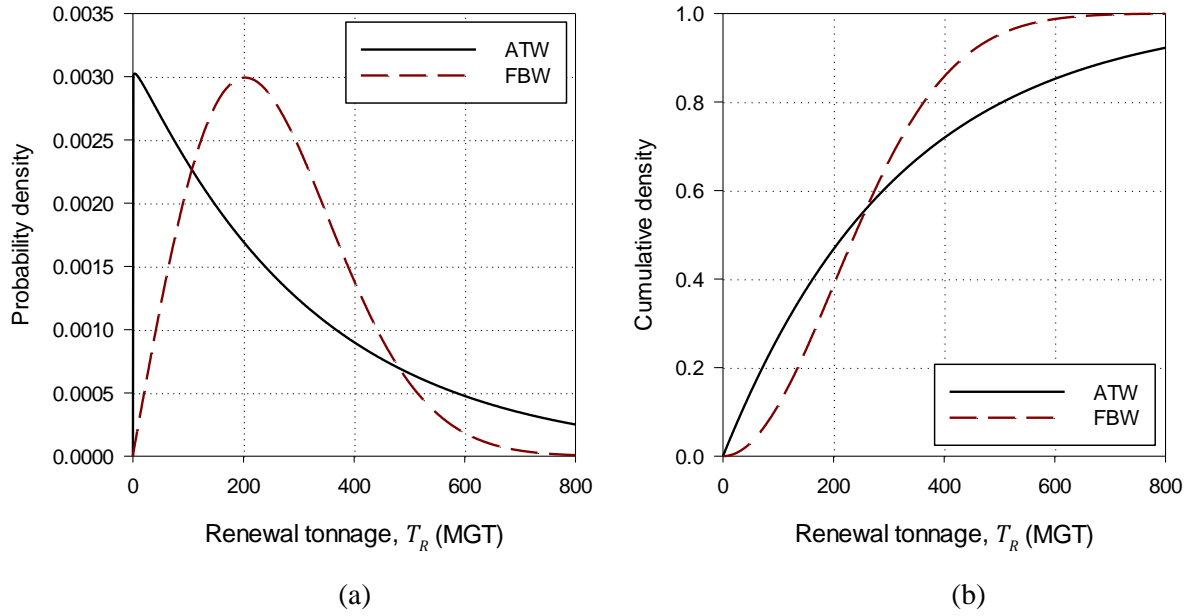


Figure 4.31: (a) PDF and (b) CDF for ATWs and FBWs

Due to the higher probability of ATW defects forming earlier in the ATW life as compared to FBWs, it can be expected that using FBWs could reduce the minimum attainable LCC c_{LCCmin} . However, the fact that flash butt welding is more costly than alumino-thermic welding must be taken into consideration. Therefore, an approach is used whereby the cost of planned maintenance using flash butt welding c_{pFBW} is specified as a ratio of the cost of planned maintenance using alumino-thermic welding c_{pATW} . This cost ratio c_{pFBW}/c_{pATW} is used to determine the point at which it is no longer financially viable to use FBWs as opposed to ATWs for maintenance. The cost of unplanned maintenance using FBWs c_{fFBWs} is calculated using the cost ratio c_{pFBW}/c_{pATW} and the cost of unplanned maintenance using ATWs c_{fATW} as shown in Equation 4.5:

$$c_{fFBW} = \left(\frac{c_{pFBW}}{c_{pATW}} \right) \cdot c_{pATW} + (c_{fATW} - c_{pATW}) \tag{4.5}$$

With:

$$\left(\frac{c_{p_{FBW}}}{c_{p_{ATW}}}\right) = \text{the cost ratio}$$

Equation 4.5 ensures that the proportion of the delay cost within the unplanned maintenance costs $c_{f_{ATW}}$ and $c_{f_{FBW}}$ remains the same. Thence, it is assumed that no additional delays are experienced as a result of conducting FBW maintenance over ATW maintenance and that any additional costs resulting from FBW maintenance is only as a result of equipment and materials used. The cost ratio $c_{p_{FBW}}/c_{p_{ATW}}$ was varied from 1 ($c_{p_{FBW}} = c_{p_{ATW}}$) to 4 ($c_{p_{FBW}} = 4 \cdot c_{p_{ATW}}$) and the resulting average LCC $\overline{c_{LCC}}$ calculated. The results are shown in Figure 4.32. Also plotted in Figure 4.32 is the result of the reference case analysis which used ATWs for maintenance. The shape of the curves vary according to the weld type used for maintenance. This is a result of the underlying shape difference in hazard functions used for ATWs and FBWs. The LCC curves for the analyses conducted using FBWs for maintenance show that the minimum attainable mean LCC $\overline{c_{LCC}}_{min}$ rises as the cost ratio increases. Furthermore, the renewal tonnage coinciding with the minimum attainable LCC $T_{R_{cr}}$ decreases as the cost ratio increases.

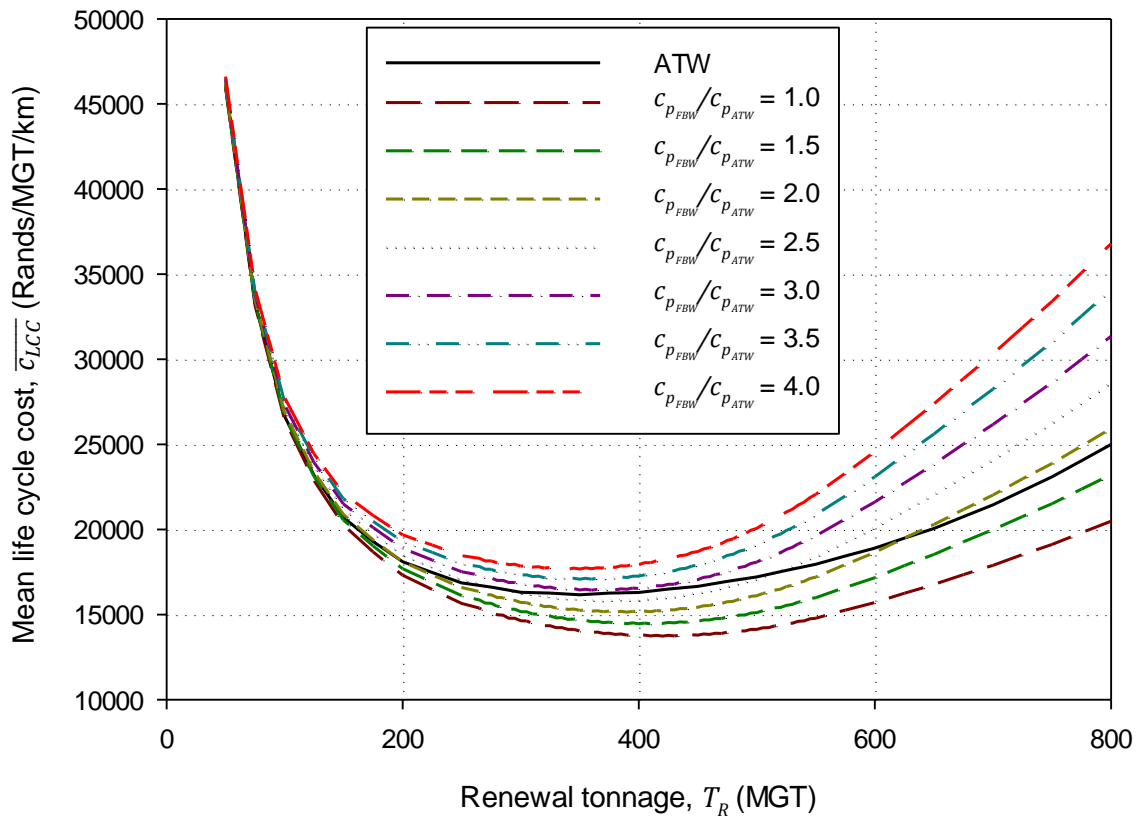


Figure 4.32: Average LCC $\overline{c_{LCC}}$ versus renewal tonnage T_R for varying $c_{p_{FBW}}/c_{p_{ATW}}$ cost ratios

Figure 4.33 shows a plot of the cost ratio $c_{p_{FBW}}/c_{p_{ATW}}$ versus the minimum attainable mean LCC $\overline{c_{LCC}}_{min}$ for cost ratios varying from 1 through 6. The minimum attainable mean LCC $\overline{c_{LCC}}_{min}$ using

ATWs for maintenance is also indicated in the figure. The cost ratio coinciding with the minimum attainable LCC for the ATW case shall be referred to as the critical cost ratio. Any cost ratios higher than the critical ratio would indicate that it would be more economical to use alumino-thermic welding than flash butt welding. The critical cost ratio is dependent on the hazard functions for ATW and FBW defects. Therefore, accurate estimates of both the hazard functions and the costs of ATWs versus FBWs would have to be determined before an informed decision could be made on which welds to use for maintenance. For the case considered it would be more economical to use FBWs if they cost less than approximately 2.75 that of ATWs.

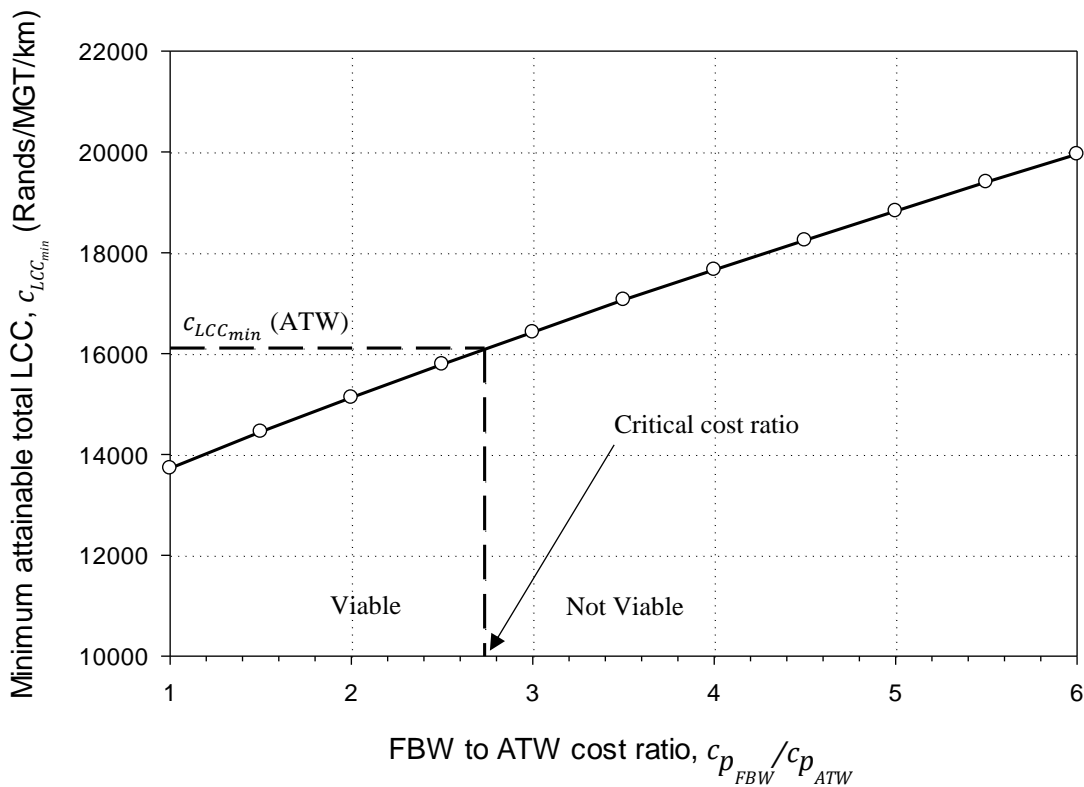


Figure 4.33: Minimum attainable LCC $c_{LCC_{min}}$ versus cost ratio $c_{p_{FBW}}/c_{p_{ATW}}$

4.8 EFFECT OF DERAILMENTS

The cost of derailments have not been considered in any previous analyses. Derailments are extreme events and are extremely costly. Due to the large cost associated with such an event and the low probability of such an event occurring, it is considered separately in this study. In a previous model to the one presented in this study, derailments were included as chance events which could occur with every unplanned maintenance action. However, a much larger number of simulations had to be conducted in these analyses in order to adequately represent the derailment probability due to the small

probability of a derailment occurring. On assessment of the resulting LCC distributions it was seen that unique distributions could be fitted to the LCC c_{LCC} at a given renewal tonnage by grouping the life cycles according to the number of derailments that occurred within them. This approach was discarded due to the large number of simulations required.

The effect of derailments on the mean LCC curve is taken into account by adding the result of Equation 4.6 to the mean LCC from the reference case analysis for a given renewal tonnage T_R :

$$c_{d_{norm}} = n_d \frac{c_d}{T_R} \quad (4.6)$$

With:

$c_{d_{norm}}$ = the normalised cost of derailments during the life cycle

n_d = the number of derailments which occurred during the life cycle

c_d = the cost of a single derailment incident

Figure 4.34 illustrates the influence of derailments on the mean LCC curve for varying numbers of derailment incidents. The cost of a derailment was set to $c_d = R\ 60$ million. The modelling of derailments in this manner is not stochastic. Rather, it is assumed that a specified number of derailments has occurred during the life cycle. From Figure 4.34 it can be seen that derailments significantly influence the mean LCC curve. The mean life cycle cost curves are translated upwards indicating an overall higher mean LCC as well as being stretched in the renewal tonnage axis. This means that if a minimum mean LCC is to be attained, it will only occur at a much larger renewal tonnage T_R . However, this minimum mean LCC is still significantly larger than if no derailments had occurred.

It is foreseeable that the probability that further derailments occur in a single life cycle after a previous derailment has occurred is increased. This is due to the large increase in the renewal tonnage at which the minimum mean LCC now occurs. The increased service life of the rail required to reach this new minimum LCC allows further opportunity for derailments to occur. It may therefore become advisable to renew the rail before $T_{R_{cr}}$ is reached if this effect were taken into account. However, this has not been considered in this study.

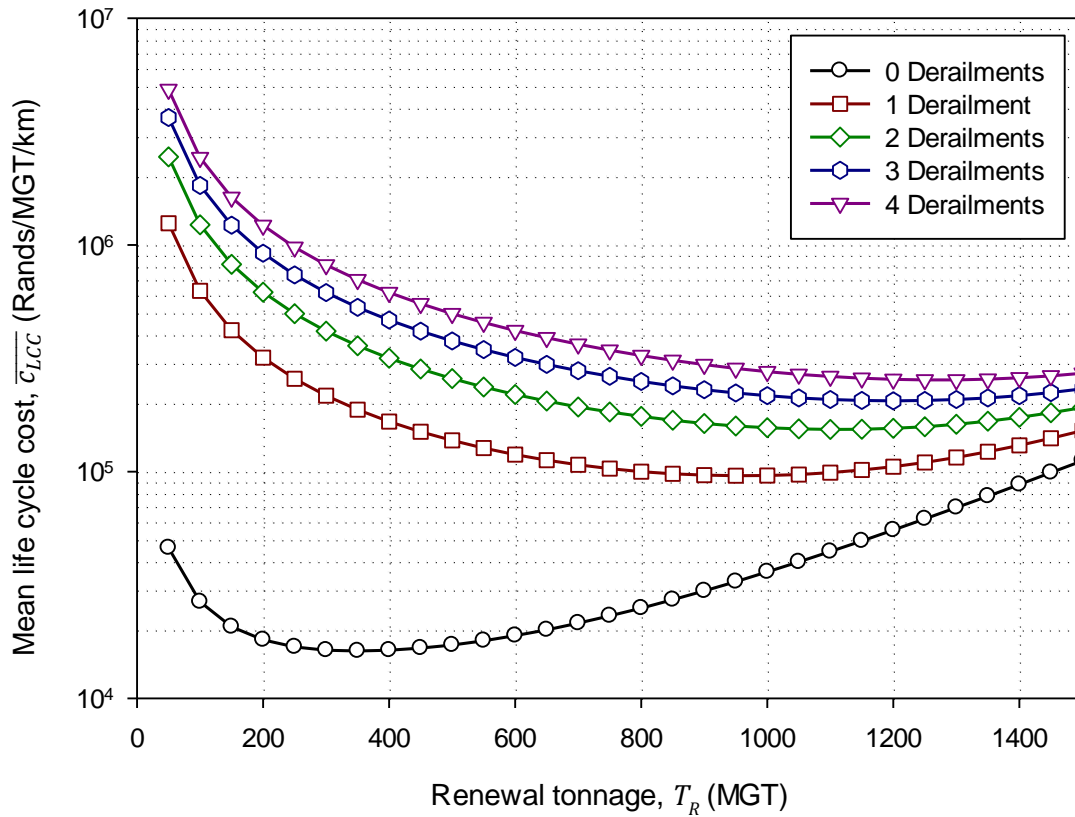


Figure 4.34: Influence of derailments on the mean LCC $\overline{c_{LCC}}$ versus renewal tonnage T_R curves

Figure 4.35 (a) illustrates the relationship between the number of derailments that occur during a life cycle and the minimum attainable mean LCC $\overline{c_{LCC_{min}}}$. The relationship is not linear and the minimum attainable mean LCC $\overline{c_{LCC_{min}}}$ is not simply offset by the cost of derailments $n_d c_d$. This is because normalised cost is considered and not the absolute cost. The consecutive differences between the minimum attainable mean LCC $\overline{c_{LCC_{min}}}$ for scenarios representing 1, 2, 3 and 4 derailment incidents is R 57 598.00, R 51 739.00 and R 48 590.00 respectively indicating that the increase in $\overline{c_{LCC_{min}}}$ in relation to the number of derailments is not linear but rather hyperbolic or inversely exponential. The exact nature of this relationship is not investigated further. The value of $T_{R_{cr}}$ in relation to the number of derailment incidents is shown in Figure 4.35 (b). This relationship appears to be logarithmic because when the same data was plotted on a logarithmic scale a linear plot resulted. The logarithmic plot has not been shown due to the small range of the values on the number of derailments axis.

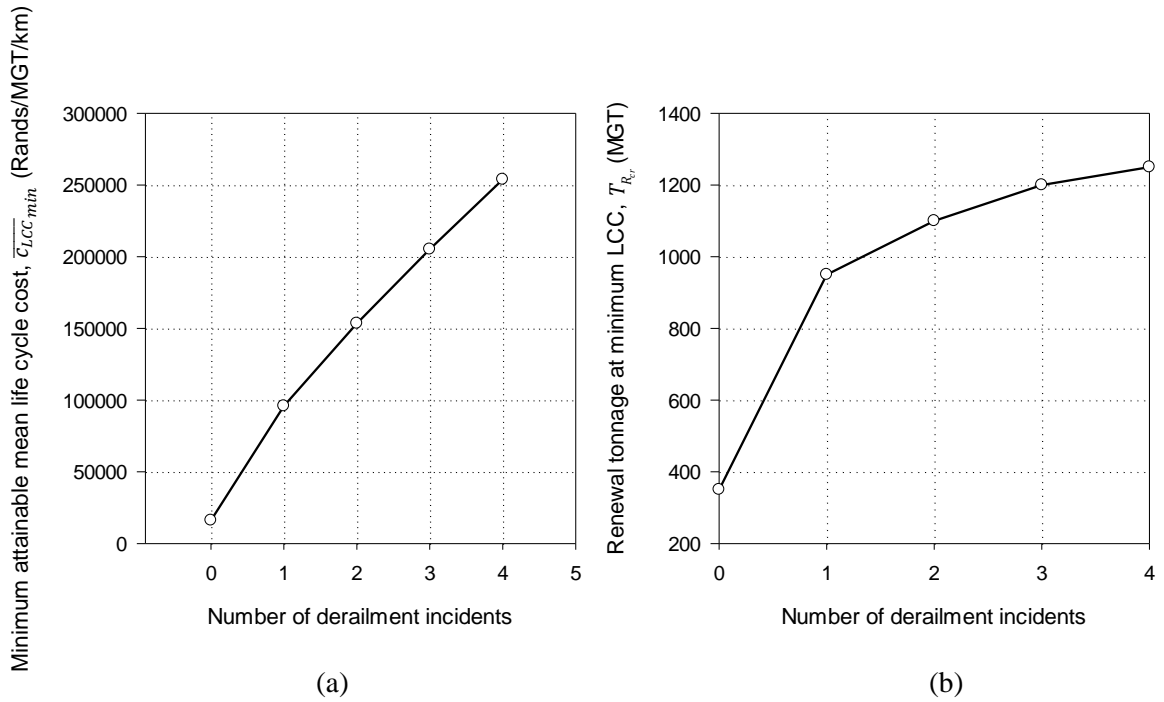


Figure 4.35: The influence of derailments on (a) the minimum attainable mean LCC $\overline{c_{LCC_{min}}}$ and (b) the renewal tonnage at which it occurs $T_{R_{cr}}$

5 CONCLUSIONS AND RECOMMENDATIONS

The conclusions and recommendations of this study are presented in Sections 5.1 and 5.2 respectively.

5.1 CONCLUSIONS

The conclusions made with regard to this study are summarised with respect to the modelling procedure, the behaviour of the mean life cycle cost (LCC), the distribution of the LCC, the uncertainty in the LCC, the use of flash butt welds (FBWs) over alumino-thermic welds (ATWs) and the effect of derailments on the LCC in the respective subsections which follow.

5.1.1 Modelling Procedure

The following conclusions are made with regard to the general modelling procedure used for the study:

- Monte Carlo simulation allows for flexible modelling and relaxation on assumptions otherwise required for simpler closed form solutions to LCC estimation.
- Monte Carlo simulation allows quantification of the uncertainty in the LCC associated with the input parameters of the model. This allows properly informed decisions to be made with regard to the level of risk the infrastructure manager (IM) is willing to accept.
- The accuracy of the results of the model are influenced by the repeatability of the pseudorandom numbers generated. Thus, the results of two simulations run with exactly the same parameters will vary. However, this did not have a significant influence on the data produced and interpreted for this study.
- The distribution of defect inter-arrival times t_d and P-F interval lengths t_{p-f} as produced by Monte Carlo sampling were tested for goodness-of-fit against the desired distribution using the Pearson Chi-square test where applicable. The P-value for all defects and all renewal tonnages considered was always larger than 0.1 indicating an adequate fit.

5.1.2 Behaviour of the Mean Life Cycle Cost and its Cost Components

The following conclusions are made with regard to the behaviour of the mean LCC $\overline{c_{LCC}}$ as a function of renewal tonnage T_R :

- The mean LCC curve was found to strongly agree with the results of a similar model developed by Zhao et al. (2006) when the same input parameters were used. The minor differences are suspected to be as a result of underlying implicit assumptions in the Zhao et al. (2006) model

which influence the manner in which the hazard rate increases when a defect initiates and then subsequently fails or is maintained.

- For a given set of conditions an optimal renewal tonnage $T_{R_{cr}}$ exists at which the rail should be renewed in order to achieve a minimum normalised mean life cycle cost $\overline{c_{LCC}}_{min}$.
- The minimum attainable mean LCC $\overline{c_{LCC}}_{min}$ and the renewal tonnage at which it occurs $T_{R_{cr}}$ are dependent on numerous factors such as inspection interval length, the weld type used for maintenance as well as the cost of maintenance, ultrasonic inspection and grinding.
- The mean LCC $\overline{c_{LCC}}$ versus renewal tonnage T_R curve becomes flatter with a reduction in the length of the inspection interval s_I . This allows a larger window for rail renewal at a near minimal mean LCC $\overline{c_{LCC}}_{min}$. However, the minimum attainable mean LCC $\overline{c_{LCC}}_{min}$ becomes larger for inspection interval lengths shorter than optimum.
- A decrease in the inspection interval length s_I will cause the optimum renewal tonnage $T_{R_{cr}}$ to increase.
- The proportional contribution of unplanned maintenance cost towards the total maintenance cost increases with an increase in the inspection interval length s_I . This indicates that the total increase in maintenance cost associated with a larger inspection interval length is as a result of a higher proportion of unplanned maintenance.

5.1.3 Life Cycle Cost Distribution

The following conclusions are made with regard to the distribution of the total LCC produced by the model developed:

- The LCC for a given set of input parameters can be expressed as a bivariate LCC-renewal tonnage probability distribution.
- The LCC can be expressed as an infinite number of univariate distributions corresponding to a given renewal tonnage.
- The hypothesis that the univariate distribution of LCC at a renewal tonnage $T_R = 800$ MGT follows a lognormal distribution was not rejected at a significance level of 1 % for all inspection intervals considered excepting $s_I = 0.1$ MGT as determined using the Pearson Chi-square goodness-of-fit test.
- Using the mean LCC does not provide a true indication of the risks associated with renewing the rail at a specific renewal tonnage T_R . Monte Carlo simulation provides a distribution for LCC which is more transparent and allows quantification of the uncertainty and risk involved with renewing the rail at a specific renewal tonnage T_R under a given set of conditions.

5.1.4 Uncertainty Trends in the Life Cycle Cost Distribution

The following conclusions are made with regard to the uncertainty of the LCC and its respective components using the standard deviation of the fitted lognormal distribution as a metric:

- The uncertainty of the total LCC σ_{LCC} for a fixed renewal tonnage T_R increases with an increase in inspection interval length s_I .
- The uncertainty of the LCC σ_{LCC} for a fixed inspection interval length s_I increases with an increase in renewal tonnage T_R .
- The trend of the uncertainty of the planned σ_{LCC_p} , unplanned σ_{LCC_f} and combined $\sigma_{LCC_{combined}}$ maintenance costs with an increase in renewal tonnage T_R can be divided into two distinct stages. The first stage occurs at renewal tonnages which are low relative to the inspection interval length. In this stage a rapid increase in uncertainty is experienced as a result of the increasing number of maintenance actions on the rail. The uncertainty then peaks and drops to a local minimum. The second stage occurs after this local minimum and may be described as a steady state of increase in uncertainty. The renewal tonnage at which this steady state occurs is dependent on the inspection interval length s_I .
- It was observed that the uncertainty with regard to planned maintenance cost σ_{LCC_p} increases with a decrease in inspection interval length s_I during the steady region of the curve for a given renewal tonnage. This is because there is an increase in the number of planned maintenance activities associated with a decrease in inspection interval length s_I .
- Furthermore, it was observed that the uncertainty with regard to unplanned maintenance cost σ_{LCC_f} decreases with a decrease in inspection interval length s_I during the steady region of the curve for a given renewal tonnage. This is because there is a decrease in the number of unplanned maintenance activities associated with a decrease in inspection interval length s_I .
- The combined maintenance cost uncertainty curve tended more towards the planned maintenance uncertainty curve for smaller inspection interval lengths and more towards the unplanned maintenance uncertainty curve for larger inspection interval lengths.

5.1.5 Using Flash Butt Welds over Alumino-thermic Welds for Maintenance

An approach was proposed which can be used to determine whether it would be more economical to use flash butt welding or alumino-thermic welding to conduct rail maintenance. The following was concluded:

- For the reference case analysis, it was more economical to use flash butt welding over alumino-thermic welding if the cost of flash butt welding did not exceed more than 2.75 times the cost of alumino-thermic welding.

- The critical cost ratio at which the cost of alumino-thermic welding and flash butt welding would produce the same minimum attainable mean LCC $\overline{c}_{LCC_{min}}$ was found to dependant on the Weibull distribution used to model alumino-thermic weld and flash butt weld defects.

5.1.6 The Effect of Derailments

The following conclusions are made with regard to the modelling and effect of derailments on the mean LCC:

- Due to their low probability of occurrence, derailments are better modelled separately from the Monte Carlo simulation process.
- The effect of derailments during a life cycle was to increase the minimum attainable mean LCC $\overline{c}_{LCC_{min}}$ in proportion to the normalised derailment cost.
- Furthermore, derailments tend to shift the optimum renewal tonnage $T_{R_{cr}}$ to much larger values than for the case where no derailments occur. This is due to the high cost of a derailment which needs to be reduced through normalisation against the tonnage borne and counter balanced by an increasing maintenance cost.

5.2 RECOMMENDATIONS

The following recommendations are made for future research in the field:

- Field or laboratory studies could be conducted to establish whether any correlation between the defect inter-arrival tonnage t_d and the P-F interval length t_{P-F} exists in reality. This could provide influential information with regard to how these two important parameters should be modelled.
- The correlation (if any) between the defect inter-arrival tonnage t_d and the P-F interval length t_{P-F} could be investigated for each weld defect in each virtual life cycle and related to the tonnage at which the weld was installed. The effect of this cross correlation between virtual life cycles could be investigated with regard to the resulting LCC.
- Industry should promote effective data collection with regard to data which will aid in accurately determining the defect inter-arrival tonnage t_d and P-F interval length t_{P-F} of rail defects experienced on the IM's line.
- The hazard rate of ATWs can be expressed using the superposition of two probability density functions as done by Zhao et al. (2007), one to represent infant mortality of the welds and the second to represent the fatigue life of the welds. This could aid in better estimation of the critical cost ratio $c_{p_{FBW}}/c_{p_{ATW}}$ at which it becomes more economical to use flash butt welding over alumino-thermic welding to conduct rail maintenance.

- A sensitivity analysis could be conducted on the stochastic input parameters of the model in order to gauge the effect of the input distribution parameters on the resulting uncertainty of the LCC σ_{LCC} .
- Inspection intervals could be modelled to decrease with cumulative tonnage. The resulting effect on the development of uncertainty in the LCC could be investigated. It is suspected that a shortening inspection interval will reduce the uncertainty at larger renewal tonnages and lead to a reduced minimum obtainable mean LCC $\overline{c_{LCC_{min}}}$.

6 REFERENCES

- Andrade, A.R., 2008. *Renewal decisions from a Life-cycle Cost (LCC) Perspective in Railway Infrastructure: An integrative approach using separate LCC models for rail and ballast components*. MSc Thesis. Universidade Técnica de Lisboa.
- Besuner, P.K., Stone, D., Schoeneberg, K.W. & De Herrera, M.A., 1978. Probability analysis of rail defect data. In *Proceedings of the 1978 Heavy Haul Railways Conference*. Perth, Australia, 1978.
- Bevilacqua, M. & Braglia, M., 2000. The analytic hierarchy process applied to maintenance strategy selection. *Reliability Engineering & System Safety*, 70, pp.71-83.
- Block, H.W., Borges, W.S. & Savits, T.H., 1985. Age dependent minimal repair. *Journal of Applied Probability*, 22, pp.370-85.
- Box, G.E.P. & Cox, D.R., 1964. An Analysis of Transformations. *Journal of the Royal Statistical Society*, 26(2), pp.211-52.
- Cannon, D.F., Edel, K.O., Grassie, S.L. & Sawley, K., 2003. Rail defects: an overview. *Fatigue & Fracture of Engineering Materials & Structures*, 26(10), pp.865-86.
- Cannon, D. & Pradier, H., 1996. Rail rolling contact fatigue research by the European Rail Research Institute. *Wear*, 191, pp.1-13.
- Chan, P.K.W. & Shaw, L., 1993. Modeling repairable systems with failure rates that depend on age and maintenance. *IEEE Transactions on Reliability*, 42, pp.566-70.
- Chattopadhyay, G. & Reddy, V., 2007. Development of Integrated Model for Assessment of Operational Risks in Rail Track. In *IEEE International Conference on Industrial Engineering and Engineering Management*. Singapore, 2007. IEEE.
- Chattopadhyay, G., Reddy, V. & Larsson-Kräik, P., 2005. Decision on economic rail grinding interval for controlling rolling contact fatigue. *International Transactions in Operations Research*, 12, pp.545-58.

Dekker, R., 1996. Applications of maintenance optimization models: a review and analysis. *Reliability Engineering and System Safety*, 51, pp.229-40.

Department of Defense, 1981. *Definitions of Terms for Reliability and Maintainability (MIL-STD-721C)*. Military Standard. Washington.

Endrenyi, J. et al., 2001. The Present Status of Maintenance Strategies and the Impact of Maintenance on Reliability. *IEEE Transactions on Power Systems*, 16(4), pp.638-46.

Esveld, C., 2001. *Modern Railway Track*. 2nd ed. Delft University of Technology.

Fröhling, R.D., 2007. Wheel/rail interface management in heavy haul railway operations - applying science and technology. *Vehicle System Dynamics*, 45(7-8), pp.649-77.

Grassie, S.L., 2012. Squats and squat-type defects in rails: the understanding to date. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 226(3), pp.235-42.

IHHA, 2001. Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues. Virginia, USA, 2001. International Heavy Haul Association.

Jablonski, D., Tang, Y.H. & Pelloux, R.M., 1990. Simulation of railroad crack growth life using laboratory specimens. *Theoretical and Applied Fracture Mechanics*, 14, pp.27-36.

Jeong, D.Y., Tang, Y.H. & Orringer, O., 1997. Damage tolerance analysis of detail fractures in rail. *Theoretical and Applied Fracture Mechanics*, 28, pp.109-15.

Kalbfleisch, J.D. & Prentice, R.L., 2002. *The Statistical Analysis of Failure Time Data*. 2nd ed. New Jersey: John Wiley & Sons, Inc.

Khouy, I.A. et al., 2014. Geometrical degradation of railway turnouts: A case study from a Swedish heavy haul railroad. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 228(6), pp.611-19.

Kijima, M., 1989. Some results for repairable systems with general repair. *Journal of Applied Probability*, 26(1), pp.89-102.

Kijima, M. & Nakagawa, T., 1991. Accumulative damage shock model with imperfect preventive maintenance. *Naval Research Logistics*, 38, pp.145-56.

Kim, H., Kwon, Y. & Park, D., 2007. Adaptive sequential preventive maintenance policy and Bayesian consideration. *Communications in Statistics - Theory and Methods*, 36(6), pp.1251-69.

Kobayashi, H., Mark, B.L. & Turin, W., 2012. *Probability, Random Processes, and Statistical Analysis*. 1st ed. New York: Cambridge University Press.

Kumar, S., 2006. *A Study of the Rail Degradation Process to Predict Rail Breaks*. Licentiate Thesis. Luleå: Luleå University of Technology.

Kumar, S., Espling, U. & Kumar, U., 2008. Holistic procedure for rail maintenance in Sweden. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 222(4), pp.331-44.

Lekule, S., 2015. *Flash Welding Process*. [Online] Available at: <http://sosteneslekule.blogspot.co.za/2015/12/flash-welding-process-flash-butt.html> [Accessed 02 May 2016].

Lie, C.H. & Chun, Y.H., 1989. An algorithm for preventive maintenance. *Journal of Applied Probability*, 26(1), pp.89-102.

Lindqvist, B.H. & Rannestad, B., 2011. Monte Carlo exact goodness-of-fit tests for nonhomogeneous Poisson processes. *Applied Stochastic Models in Business and Industry*, 27, pp.329-41.

Lui, Y., Huang, H. & Zhang, X., 2012. A data-driven approach to selecting imperfect maintenance models. *IEEE Transactions on Reliability*, 61(1), pp.101-12.

Magel, E., Roney, M., Kalousek, J. & Sroba, P., 2003. The blending of theory and practice in modern rail grinding. *Fatigue and Fracture of Engineering Materials and Structures*, 26(10), pp.921-29.

Magel, E.E., Sawley, K.J., Sroba, P.S. & Kalousek, J., 2005. A Practical Approach to Controlling Rolling Contact Fatigue in Railways. In *Proceedings of the 8th International Heavy Haul Conference*. Rio de Janeiro, Brazil, 2005.

Malik, M.A.K., 1979. Reliable preventive maintenance policy. *AIIE Transactions*, 11(3), pp.221-28.

Marais, J.J. & Mistry, K.C., 2003. Rail integrity management by means of ultrasonic testing. *Fatigue & Fracture of Engineering Materials & Structures*, 26(10), pp.931-38.

Montgomery, D.C. & Runger, G.C., 2011. *Applied Statistics and Probability for Engineers*. 5th ed. Asia: John Wiley & Sons, Ltd.

Moore, D.S., 1986. Tests of Chi-Square Type. In R.B. D'Agostino & M.A. Stephens, eds. *Goodness-of-fit Techniques*. 1st ed. New York: Marcel Dekker, Inc. pp.63-95.

Morant, A., Galar, D. & Tamarit, J., 2012. Cloud computing for maintenance of railway signalling systems. In *Proceedings of the Ninth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies*. London, 2012. British Institute of Non-Destructive Testing.

Moubray, J., 1997. *Reliability-centered Maintenance*. 2nd ed. New York: Industrial Press Inc.

Muster, H., Schmedders, H., Wick, K. & Pradier, H., 1996. Rail rolling contact fatigue. The performance of naturally hard and head-hardened rails in track. *Wear*, 191, pp.54-64.

Nakagawa, T., 1979. Optimum policies when preventive maintenance is imperfect. *IEEE Transactions on Reliability*, 28(4), pp.331-32.

Nakagawa, T. & Yasui, K., 1987. Optimum policies for a system with imperfect maintenance. *IEEE Transactions on Reliability*, 36(5), pp.631-33.

Nguyen, D.G. & Murthy, D.N.P., 1981. Optimal preventive maintenance policies for repairable systems. *Operations Research*, 29(6), pp.1181-94.

Olofsson, U. & Nilsson, R., 2002. Surface cracks and wear of rail: a full-scale test on a commuter train track. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 216(4), pp.249-64.

Orringer, O., Morris, J. & Steele, R., 1984. Applied research on rail fatigue and fracture in the United States. *Theoretical and Applied Fracture Mechanics*, 1, pp.23-49.

Orringer, O., Tang, Y., Jeong, D. & Perlman, A., 1999. *Risk/Benefit Assessment of Delayed Action Concept for Rail Inspection*. Report DOT/FRA/ORD-99/03. US Department of Transportation.

Patra, A.P., 2007. *RAMS and LCC in Rail Track Maintenance*. PhD Thesis. Luleå University of Technology.

Patra, A.P., Söderholm, P. & Kumar, U., 2008. Uncertainty in Life Cycle Cost of Railway Track. In *Proceedings of the Annual Reliability and Maintainability Symposium (RAMS)*. Las Vegas, 2008.

Patra, A.P., Söderholm, P. & Kumar, U., 2009. Uncertainty estimation in railway track life-cycle cost: a case study from Swedish National Rail Administration. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 223(3), pp.285-93.

Pham, H. & Wang, H., 1996. Imperfect maintenance. *European Journal of Operational Research*, 94(3), pp.425-38.

Plasser South Africa, 2015. *Flash Butt Welding Machines*. [Online] Available at: <http://www.plasser.co.za/en/machines-systems/flash-butt-welding.html> [Accessed 02 May 2016].

Qian, C.H., Ito, K. & Nakagawa, T., 2005. Optimal preventive maintenance policies for a shock model with given damage level. *Journal of Quality in Maintenance Engineering*, 11(3), pp.216-27.

Reinschmidt, A. et al., 2015. Chapter 5: Wheel and Rail Damage Mechanisms. In J. Leeper & R. Allen, eds. *Guidelines to Best Practices For Heavy Haul Railway Operations: Management of the Wheel and Rail Interface*. Omaha: Simmons-Boardman Books, Inc.

Rigdon, S.E. & Basu, A.P., 2000. *Statistical Methods for the Reliability of Repairable Systems*. 1st ed. New York: John Wiley & Sons, Inc.

Robert, C.P. & Casella, G., 2004. *Monte Carlo Statistical Methods*. 2nd ed. New York: Springer.

Roth, C., 2008. Use of the Weibull Distribution to Predict Rail Defect Occurrence Rates. In *International Federation of Operational Research Societies 18th Triennial Conference*. Sandton, South Africa, 2008.

Rubinstein, R.Y. & Kroese, D.P., 2008. *Simulation and the Monte Carlo Method*. 2nd ed. New Jersey: John Wiley & Sons, Inc.

Sawley, K. & Reiff, R., 2000. *Rail Failure Assessment for the Office of the Rail Regulator*. Pueblo, Colorado: Transportation Technology Centre, Inc. A Subsidiary of the Association of American Railroads.

Schlake, B.W., Barkan, C.P.L. & Edwards, J.R., 2011. Train Delay and Economic Impact of In-Service Failures of Railroad Rolling Stock. *Transportation Research Record: Journal of the Transportation Research Board*, 2261, pp.124-33.

Selig, E.T. & Waters, J.M., 1994. *Track Geotechnology and Substructure Management*. London: Thomas Telford Publications.

Shafiee, M., Patriksson, M. & Chukova, S., 2016. An optimal age-usage maintenance strategy containing a failure penalty for application to railway tracks. *Proceedings of the Institution of Mechanical Engineers Part F: Journal of Rail and Rapid Transit*, 230(2), pp.407-17.

Smith, R.A., 2003. The wheel-rail interface - some recent accidents. *Fatigue & Fracture of Engineering Materials & Structures*, 26(10), pp.901-07.

The Guardian, 2006. *The Guardian*. [Online] Available at:
<http://www.theguardian.com/uk/2006/jul/25/hatfield.transport> [Accessed 05 January 2016].

Van der Westhuizen, N.J. & Gräbe, P.J., 2013. The integration of railway asset management information to ensure maintenance effectiveness. *Journal of the South African Institution of Civil Engineering*, 55(3), pp.18-29.

Vatn, J., Podofillini, L., Zio, E. & Svee, H., 2003. A risk based approach to determine type of ultrasonic inspection and frequencies in railway applications. In *Proceedings of the World Congress on Railway Research*. Edinburgh, UK, 2003.

Wang, H., 2002. A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3), pp.469-89.

Wang, H. & Pham, H., 1996. A quasi renewal process and its applications in imperfect maintenance. *International Journal of Systems Science*, 27(10), pp.1055-62.

Witt, D., 2007. The 19th Jenkin Lecture, 23 September 2006: Railways: The Technical Challenges of their Renaissance. *Society of Oxford Engineers News*, Issue 6. pp.27-32. Available at: <http://www.soue.org.uk/files/souenews6.pdf> [Accessed 17 May 2016].

Wu, H., Woody, S.S. & Blank, R.W., 2005. Optimization of Rail Grinding on a North American Heavy Haul Railroad Line. In *Proceedings of the 8th International Heavy Haul Conference*. Rio de Janeiro, 2005.

Zarembski, A.M., 1991. Forecasting of Track Component Lives and Its Use in Track Maintenance Planning. In *International Heavy Haul Railway Conference*. Vancouver, Canada, 1991.

Zhao, J., Chan, A.H.C. & Burrow, M.P.N., 2007. Probabilistic model for predicting rail breaks and controlling risk of derailment. *Transportation Research Record: Journal of Transportation Research Board*, 1995, pp.76-83.

Zhao, J., Chan, A.H.C., Roberts, C. & Stirling, A.B., 2006. Assessing the economic life of rail using a stochastic analysis of failures. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 220(2), pp.103-11.

Zoeteman, A., 2001. Life cycle cost analysis for managing rail infrastructure: concept of a decision support system for railway design and maintenance. *European Journal of Transport & Infrastructure Research*, 1(4), pp.391-413.

Zoeteman, A., Dollevoet, R. & Li, Z., 2014. Dutch research results on wheel/rail interface management: 2001-2013 and beyond. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 228(6), pp.642-51.