University of Nottingham

Railway Track Asset Management Modelling

Thesis submitted to The University of Nottingham for the degree of Doctor of Philosophy

Author:
Steve Clarke

Supervisors:
Darren Prescott
John Andrews

March 16, 2021

Acknowledgements

I would like to thank my supervisors, Darren Prescott and John Andrews. The extensive knowledge, guidance and patience you have shown me has enabled me to grow and complete this thesis. I have learnt so much on this endeavour.

I would never have made it through this task if not for the amazing people at NTEC, in particular my fellow PhD students Panos Yianni and Paul Kilsby. I enjoyed our deep discussions about the world and all the fun we had (while learning).

A special acknowledgement to the University of Nottingham for all the opportunities I have been given over the past 8 years. Choosing Nottingham University has to be one of the best decisions I have ever made and I will never forget my time spend there.

My research would not have been possible without the technical expertise of Network Rail as well as the financial assistance. Special thank you to Andy Kirwin, Sam Chew and Julian Williams for all you guidance.

I would also like to thank my family for their support. Special thank you to my grandparents Joan and Alan Clarke and Jack Davidson, you helped make me who I am today and gave me the drive to do this PhD, I really wish you were able to see me finish it.

Extra special thank you to my partner, Sarah. You have been with me through the ups and many downs always supporting and encouraging me to do my best. I would never have finished this thesis without you.

Publications

Conference Publications

Clarke S, Prescott DR (2017). Identifying the factors influencing track degradation on the UK rail network. The Stephenson Conference 2017: Research for Railways. London, UK, 25-27 Apr 2017.

Abstract

Railways are an important type of transport infrastructure but can be expensive to run with the UK railway costing £1.525 billion to maintain in 2018/19. To reduce the cost without reducing the quality of the infrastructure improved asset management is required. To enable the impact of possible decisions to be understood, and the optimum ones chosen, a railway track asset management model, such as the one developed in this thesis, is required.

This thesis first studies the degradation and maintenance of railway track. The UK railways network actual geometry and maintenance data was used to understand track degradation. This data covered 8 years of the whole UK rail network, a much greater length of time and track than previous research has considered. Many aspects were shown to have a significant impact on the rate such as track speed, sleeper type and maintenance history. The methodology of singling out factors showed that rail types have less of an impact on track geometry than previous research had shown. Weibull distributions were then used to characterise the rates of degradation of separate combinations of these significant aspects. The improvement in geometry from maintenance is also explored, with the effectiveness reducing with each further maintenance action. The improvement in geometry from maintenance is modelled using a linear fit with a stochastic element added to model the effectiveness variability. Maintenance output rate, which has previously not been considered in literature, has also been analysed and modelled, utilising Weibull distributions, allowing working window lengths for maintenance to be incorporated within models.

The likelihood of different rail faults occurring was also explored using data from the UK railway. The analysis showed that the rail type, joint and age were not only linked to the rate of faults but also the track geometry. This link has been mentioned in literature but has previously not been proved or quantified. The link between rail faults and track geometry shows how the railway tracks assets are interlinked and hence need to be modelled as such. Saving money by reducing the amount of track geometry maintenance will increase the quantity and hence the cost of rail faults. Rail faults have been modelled using probabilities within this thesis, with the fault rates of each fault type related to the track geometry.

The second part of this thesis develops a Coloured Petri Net model which incorporates

the analysis and models developed. A Coloured Petri Net has been used in a novel way of acting as a decision framework, joining the separate degradation and maintenance models together and allowing them to interact. This removes some shortcomings of state based modelling techniques such as requiring discrete states of degradation. The model predicts, over any given line of track and time period, the number of inspections, maintenance actions, track quality and number of speed restrictions. Utilising the model the user can assess the impact of decisions such as maintenance thresholds, asset upgrades and traffic changes. Additional aspects such as opportunistic maintenance, as well as maintenance productivity and work windows lengths are considered. This allows aspects like opportunistic maintenance thresholds and varying the maintenance window length to be analysed, which previous models in literature have not. Different scenarios can be run through the model and the outputs compared to enable evidence based asset management decisions to be made.

Contents

List of Figures

List of Tables

Acronyms

ANFIS	Adaptive Neural-based Fuzzy Inference System.
ANN	Artificial Neural Network.
BCF	Ballast Condition Factor.
BFI	Ballast Fouling Index.
BH	Bullhead.
CDF	Cumulative Distribution Function.
CEMGT	Cumulative Equivalent Million Gross Tonnage.
COV	Coefficient of Variation.
CPN	Coloured Petri Net.
CTR	Combined Track Record.
CWR	Continuous Welded Rail.
ECDF	Empirical Cumulative Distribution Function.
ELR	Engineer Line Reference.
EMGT	Equivalent Million Gross Tonnage.
ESR	Early Speed Restriction.
FB	Flatbottom.
FIS	Fuzzy Interface System.
GOF	Goodness Of Fit.
GPR	Ground Penetrating Radar.
IRJ	Insulated Rail Joint.
LCC	Life Cycle Cost.
LSCT	Large-Scale Cyclic Triaxial.

LTSF Local Track Section Factor.

M&R Maintenance and Renewal.

MF Membership Function.

MGT Million Gross Tonnage.

MSA Million Standard Axles.

MTTF Mean Time To Failure.

NR Network Rail.

OLE Overhead Line Equipment.

ORE Office of Research and Experiments.

PDF Probability Density Function.

PL Plain Line.

PN Petri Net.

RAMS Reliability, Avaliability, Maintainability and Safety.

RCF Rolling Contact Fatique.

RMSE Root Mean Squared Error.

S&C Switches and Crossing.

SD Standard Deviation.

SRS Strategic Route Section.

TGP Track Geometry Parameters.

TQI Track Quality Index.

T-SPA Track Strategic Planning Application.

UIC International Union of Railways.

VTISM Vehicle Track Interaction Strategic Model.

Chapter 1

Introduction

Railways are an important type of transport infrastructure deployed in most countries around the world. They have many advantages allowing quick and efficient transportation of freight and passengers across an area with lower emissions than road traffic. The railway industry also employs large quantities of people from research, production, installation and then use, with Oxera (2014) estimating that the UK railway industry employs around 212,000 people and contributes £9.3 billion to the economy. Due to the benefits of a country having high quality railway infrastructure it is important to constantly maintain and upgrade it. However, a railway system is expensive to run, with Network Rail (2019) reporting a yearly expenditure of £12.639 billion for 2018/19. This consists of £3.082 billion on renewals, £3.164 billion on enhancement projects and £1.525 billion on maintenance for the 32,186 km of track and other assets such as stations and bridges. Due to these high costs it is imperative to be able to run a railway network efficiently, which is where asset management techniques are employed. One of the many tools deployed as part of asset management are Life Cycle Cost (LCC) models, which are based on mathematical and/or statistical relationships and can be used to estimate asset's future costs. These costs, depending on the model, can include inspection, maintenance, renewal and upgrades costs of an asset. The results obtained are used to estimate future expenditure, plan major renewals and upgrades and optimise asset policies such as maintenance schedules. As part of this thesis an LCC model has been developed to forecast the future costs of running sections of railway track including inspections, maintenance, renewals and upgrades.

1.1 Problem Definition

Managing railway infrastructure is complex with one of the main tasks consisting of maintaining the track geometry for safety purposes and passenger comfort. The track geometry worsens with time as trains traverse the track, and once the track quality reaches an undesirable state, maintenance actions are undertaken. The geometry degra-

dation is a seemingly random process which is related to many factors, such as weather, traffic loads and speed, so is especially hard to estimate accurately within a model. On top of that the alternative maintenance activities each have varying effects on improving the track geometry with each use. Other aspects of the track also degrade such as the rails where faults can occur. Within the railway track the degradation rates of these different assets are interconnected but these relationships are very hard to pinpoint and quantify. Due to the complexity and size of a railway network the logistics of performing work is troublesome with some work having to be scheduled years in advance. There is also a finite amount of resources available so certain work has to be prioritised. This includes physical resources like the number of maintenance equipment and people, as well as aspects like the availability of the track to perform maintenance on. All this makes producing a useful LCC model that accurately models the dynamic system of the track a difficult task with complex degradation mechanisms and interactions, varying maintenance effectiveness and highly dependent work scheduling. Despite this it is a problem that is important to solve. Due to the complexity of the system to model and the high cost of maintaining the railway infrastructure, there are large optimisation opportunities. These could be cost savings whilst maintaining the same track quality or improved quality for the same expenditure, improving passenger comfort and safety.

1.2 Research Aims and Objectives

The primary research aim is to develop an LCC model which is capable of estimating the future costs of managing a section of track. The model will also allow the possibility of optimising the maintenance and inspection strategies to sustain a high standard of track quality whilst reducing costs. To accomplish this, multiple research objectives have been identified:

1. Understand the benefits of asset management and the uses of LCC models .
2. Review of research into the types of track assets, their degradation mechanisms and maintenance actions.
3. Review of current modelling techniques and previous railway track models.
4. Analyse the factors which affect the track geometry, maintenance effectiveness and rail faults using data supplied by Network Rail (NR).
5. Establish a modelling method which will be suitable to model the railway track.
6. Construct a model which takes into account track degradation, inspections, maintenance, upgrades and renewals.
7. Demonstrate the model's uses using existing track sections within the UK railway network.

1.3 Thesis Layout

Chapter 2 includes background information on asset management, railway track assets, their degradation processes and maintenance actions as well as an overview of previous research undertaken on railway track degradation and other LCC models.

In Chapter 3 data obtained from the UK railway network has been used to quantify the track degradation and perform an analysis to uncover the main factors that affect the rate of geometry degradation.

Within chapter 4 the effectiveness of the primary maintenance methods used to correct the track geometry has been analysed and the main factors identified. As part of this, a stochastic model is introduced to predict the improvement from maintenance, as well as the output of mechanical maintenance machines.

Chapter 5 introduces rail fault data which is analysed to identify the main factors which impact the rate of faults occurring. This includes quantifying the link between the track geometry quality and the rail fault rate, which has been used to produce a probabilistic model.

A stochastic CPN asset management model that simulates the track geometry degrading and rail faults occurring has been developed in chapter 6. The model allows for asset management decisions such as inspection intervals, maintenance thresholds and access times to be modelled and the impact of decisions known.

Chapter 7 gives conclusions from the work undertaken and recommendations for future work.

The Appendix contains summary statistics of the data used and backing results from the analysis in Chapters 3. Additional tables related to the proposed rail fault model are also included.

Chapter 2

Literature Review

2.1 Introduction

To develop a Life Cycle Cost (LCC) model there are three things which need to be understood. First are the uses of an LCC model and how it fits into asset management, with these understood it is possible to decide on required outputs of the developed models and design variables. Secondly, it is important to understand the system that is being modelled, in this case a railway track. As part of this, previous research into the assets involved and how they degrade has been reviewed. This will serve as the basis to an analysis of the factors which change the degradation rate of the track, such as track speed. Additionally, the types of maintenance activities have also been researched. Thirdly, a review of previous track models and general modelling techniques has been undertaken. This will allow for an informed decision on the modelling techniques used as the basis of the LCC model.

2.2 Asset Management

Asset management is an important part of most industries and applies to both tangible assets, such as machinery and buildings, as well as intangible assets, which includes patents and brand recognition. Within this thesis railway infrastructure assets will be focused on. Asset management is a complex idea that encompasses multiple uses and possible outcomes with many inputs. A formal definition of asset management, which is also one of the most widely used definitions, was outlined by the Institute of Asset Management (2008a) as being:

> *Systematic and coordinated activities and practices through which an organization optimally and sustainably manages its assets and asset systems, their associated performance, risks and expenditures over their life cycles for the purpose of achieving its organizational strategic plan.*

There are a multitude of principles and attributes which are used to optimise asset management with some primary ones being outlined by the Institute of Asset Management (2008b) as:

- Holistic; look at the whole picture.
- Systematic; methodical approach.
- Systemic; consider individual assets as part of a system and optimising the system instead of just the individual assets.
- Risk-based; decisions of expenditure and resources affected by their associated risks.
- Optimal; best value comparing factors such as performance, cost and risk over an asset's life cycle.
- Sustainable; considering long-term effects of short-term activities enabling future long-term objectives such as economic, environmental sustainability and the system performance to be achieved.

With the implementation of good asset management techniques there are many benefits achievable which have a positive impact on all areas of a business. Some of these benefits are discussed by the Institute of Asset Management (2008a), which describes how the inclusion of successful management to optimise asset's life cycles can result in:

- Improved health, safety and environmental performance;
- Optimised investment return and growth;
- Improved long term planning and confidence in future plans;
- Enhanced customer satisfaction;
- Demonstrate the best value for money, with the prediction of future costs being used to obtain funding. A comprehensive estimate of future costs reduces uncertainty and increases the likelihood of obtaining funding;
- Improved corporate reputation.

The railway industry is a substantial business that includes a large network of many types of assets. Due to this the railway industry is well suited to see many large benefits from implementing effective asset management, including cost savings, safety improvements, increased customer satisfaction and usage growth. As part of this, the industry has to deliver the outputs desired by the customers, funders and other key shareholders sustainably with the lowest whole life cost (UIC, 2010). This is done by focusing maintenance, renewals and enhancements that are key to the customers as opposed to just prioritising work according to asset condition. It is also important to have great emphasis on evidence-based decision-making by using knowledge on how assets degrade and fail to optimise the timings and types of interventions. With the use of extensive knowledge it is possible to maintain the same level of sustainable performance whilst reducing the volume of work. This knowledge is also desired to demonstrate to external shareholders the value of works and to prove the plan with the best value whole life cost is implemented.

Poor asset management can lead assets to degrade to the point where they are unfit for purpose. At best, this can lead to system down time and at worst, potential safety risks. The railway infrastructure contains a large range of assets that must be maintained in such a way that they are kept in a serviceable condition. The consequences of failure are often high, due to the potentially high monetary expense of correction and associated down time and also the risk of fatalities and injuries to the public, users and workforce. There are numerous examples of cases where poor maintenance has allowed catastrophic failures to occur, such as the infamous Potters Bar train derailment, which occurred on the 10th May 2002. This accident was caused by the failure of a set of points where flawed maintenance procedures had allowed a number of components to degrade to a poor condition. This caused a passenger train travelling at 97 mph to derail on its approach to Potters Bar railway station causing a large amount of damage to the track, a bridge and the station (HSE Potters Bar Investigation Board, 2003). The crash caused 7 fatalities and 76 major and minor injuries. Due to the incident Network Rail was fined £3,150,000 (Bright, 2011). This incident highlights the importance of effective asset management and the requirement to consider safety and condition, as well as costs.

2.2.1 Framework

The desired outcomes of asset management are achievable through the implementation of an asset management framework which should identify the key components of an asset management system, like the one seen in Figure 2.1 which is related to the railway industry. This framework describes the important mechanisms involved, as well as the core decisions and activities required, starting from the top and working down the diagram to lower level strategies and decisions. This thesis will concentrate on the enabling mechanism, LCC tools.

The enabling mechanisms on the left of Figure 2.1 outline aspects that are required to optimise asset management to get the greatest return for investment leading to a higher level of performance and safety or cost savings. This includes detailed information about the assets in the railway network and their locations. It is important to have information of the assets types, ages and conditions along with maintenance and renewal histories, failure histories and their impacts on performance and safety, as well as unit costs of previous works and failures, to predict future costs. Decisions within the railway network should be supported and optimised using mechanisms such as LCC tools. These tools can be used for decisions on maintenance activities and renewals on the railway infrastructure assets and should also be used for forecasts of work volumes and the associated costs. To make this possible the tools should include an understanding of the different assets' degradation and failure mechanisms, and the impact of these on the railway service and safety. Different maintenance and renewal options should be included within the model along with accurate cost estimates to enable comparisons of the possible options and to assist in decision-making, such as inspection intervals and maintenance thresholds.

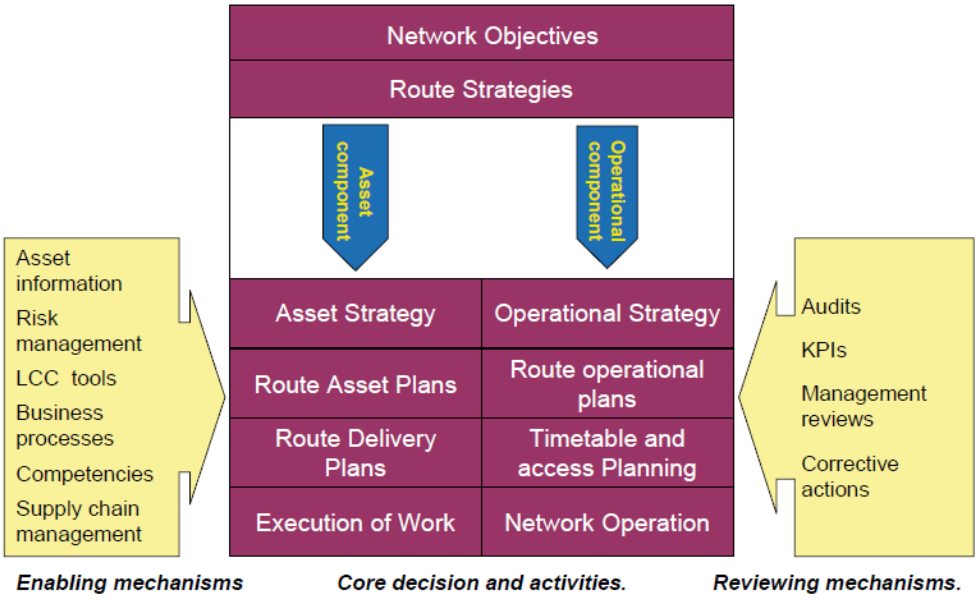


Figure 2.1: Asset management framework (UIC, 2010)

2.2.2 Life Cycle

The British Standard Institution (2004) identifies the important phases of an asset's life cycle as inception, design, manufacturing, installation, operation and maintenance, and disposal. Operation and maintenance are considered to be the most important aspects in the rail industry due to the long operational life of the assets, for example rails can last 10-15 years and concrete sleepers have a typical life of 30-40 years (Lichtberger, 2005). This results in the majority of costs occurring over the operational part of the life cycle. During system design, it is important to look not only at the cost of an installation but also to take account of the whole system life cycle and the associated costs of each part. This means considering factors such as Reliability, Avaliability, Maintainability and Safety (RAMS) and their impact on the LCC, taking account of technological advancements and changes if possible. Historically many new technological advancements have helped to reduce operational and maintenance costs, such as, Continuous Welded Rail (CWR) (introduced in the 1950's) and steel sleepers (introduced in the 1980's). These factors are considered for every asset in the system and used to inform decision-making both during the design phase and during operation when the optimal asset management strategy can be selected.

2.2.3 Summary of Asset Management

It is important for the railway industry to utilise effective asset management to increase return of investment which will in turn improve safety and performance or reduce costs. This can be accomplished by setting up a framework that has clear objectives and poli-

cies, which are chosen by considering risk, as well as short, and long-term effects on the railway network's performance and safety. As part of this the management of assets such as track need to optimised in order to identify the best decisions. These decisions can be supported by mechanisms such as LCC tools, which require detailed information of the assets in the railway network, how they degrade, the effectiveness of maintenance interventions, the associated costs and the availability of work access and resources. These tools can also be used to optimise maintenance and inspection regimes to achieve higher service quality and safety levels or reduced costs by increasing the return of investment.

2.3 Railway Track Assets

2.3.1 Introduction

Railway track mainly consists of Plain Line (PL) track and Switches and Crossings (S&Cs). These are made of many components with different degradation mechanisms, maintenance procedures and inspection approaches. Successful asset management requires an understanding of the separate components such as ballast and sleepers, their degradation mechanisms, interactions and any associated risks. It is also important to understand the maintenance actions and inspection choices that are available so these can also be embedded into the model. In the UK there are about 29,800 km of in use plain line track and 18,400 S&C units of which 99.8% use ballast to spread the load and hold the sleepers in place. The components that make up a plain line track are:

- Rails; provide a smooth running surface and guidance.
- Sleepers; support the rails at the correct inclination and spacing (gauge), transmitting vertical, lateral and longitudinal forces into the ballast.
- Rail Pads; sit between the rails and the sleepers reducing forces on the sleepers and providing electrical insulation.
- Fastenings; secure the rail to the sleepers.
- Ballast; consists of crushed rock, which supports the sleepers at the correct level and alignment. It spreads the forces into the formation, and also enables surface water to drain away.
- Formation; supports the ballast and collects water into the drainage system.
- Subgrade; this is the natural geological layer that all other parts of the track system are built on.
- Drainage; conveys water away from the track.

Within this thesis many of the asset types are considered due to the the railway track assets being linked. The primary focus is the degradation of the ballast and rails.

2.3.2 Rails

The rail is made out of steel with a cross section consisting of the head, web and foot with the train wheels running on the head. In the UK there are about 4,050 km of rail which is jointed where the rail lengths are held together using bolted fish-plated joints as seen in Figure 2.3a. The more modern technique of CWR, as seen in Figure 2.3b, is used on about 25,750 km of track in the UK with more jointed track being converted to CWR every year.

Some of the most common rail sizes in the UK are CEN56 and CEN60 flat bottom rails, as well as S95 bullhead. These are named after their weight in kg/m for flat bottom and lb/yd for the older bullhead designs. The difference in shape between bullhead and flat bottom rail can be seen in Figure 2.2, with dimensions for the three most common UK rails given in Table 2.1.

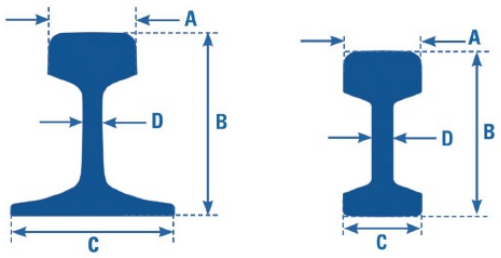


Figure 2.2: Flat bottom rail (left) and bullhead rail (right) (Balfour Beatty Rail, 2010)

Table 2.1: Typical rail sizes (Balfour Beatty Rail, 2010)

Rail Type	Weight [kg/m]	A [mm]	B [mm]	C [mm]	D [mm]
CEN56	56.27	96.85	158.75	140.00	20.00
CEN60	60.03	72.00	172.00	150.00	16.50
S95	47.07	69.85	145.26	96.85	19.05

Rail Joints

Jointed rail has many disadvantages when compared to CWR due to the additional forces at the gap between the joined rail ends, which is incorporated to allow for thermal rail expansion, and the strength of the fishplates and bolts that hold the joints together. A comparison of the forces on a flat, continuous rail head to those in jointed rail can be seen in Figure 2.4, which illustrates the increased force at the joint, due to the decreased contact area, as the wheel hits the edge. The discontinuity leads to severe, early, localised damage either to the rail head or to the fishplated joint, resulting in a shorter service life. CWR has a much longer service life than jointed rail as discussed by Zong and Dhanasekar (2013), who reported that Insulated Rail Joints (IRJs), which are joints that

include epoxy resin to enable track circuits for signalling, have an average life of about 20% compared to CWR. They also discuss the increased forces due to the discontinuity of rail joints increasing ballast breakup and formation compaction and hence can lead to faster track geometry degradation. Many factors affect the service life of bolted rail joints. The bending strength of a pair of fishplates, one either side of the rail, in practice does not exceed 30% of that of the rail. The fishplates can also be damaged by the passing of each wheel and the degradation is accelerated by loose joint sleepers, loose or over-tight bolts and excessive expansion gaps, (Mundrey, 2010). It is estimated that 60% of rail failures occur at the rail joints with many of these failures being due to yielding around the bolt holes (Akhtar et al., 2010). Bolt hole cracking is dangerous since failure often leads to a piece of rail becoming detached (Reid, 1993). The cracks are caused by shear stresses created from the dynamic wheel/rail forces generated by the discontinuity in the running surface. Failures of rail joints can be revealed either by eddy current and ultrasonic inspection or by visual inspection. Track geometry recordings (vertical profile of the track) can also detect rail joint faults since the faults can induce large spikes in the readings (sudden dips in the rail top).

Despite giving a stronger connection than fishplated joints, welded rail joints are still a source of weakness in the rail. This is due to the difficulty of perfectly matching up the cross-sections. Small errors can lead to increased stresses in the joint. The process of welding can also add impurities to the steel and the heat can negatively affect the material properties, causing metallurgical defects to develop, which can lead to sudden failures (Steenbergen and Esveld, 2006; Ministry of Railways, 2005). A primary drawback of CWR lies in its ability to deal with thermal expansion and contraction forces in the rail caused by changes in the environmental temperature (Sung et al., 2005). CWR lacks expansion/contraction joints and therefore requires the ballast to hold the sleepers and rail in place. In hot weather, as the rail expands, if the latitudinal or longitudinal forces from the rail expansion become larger than the confining forces of the ballast the rail could buckle, as seen in Figure 2.5a. In cold weather contraction can cause rail breaks to occur, as seen in Figure 2.5b. One advantage of CWR over jointed rail is that it gives a smoother running surface, resulting in less vibration and noise, reduced wear on the bogie and lower maintenance costs due to the longer life span (Lei and Feng, 2004).

(a) Typical fishplated rail joint
(Interflon, 2011)

(b) Typical welded rail joint
(Zwolski, 2012)

Figure 2.3: Rail joints

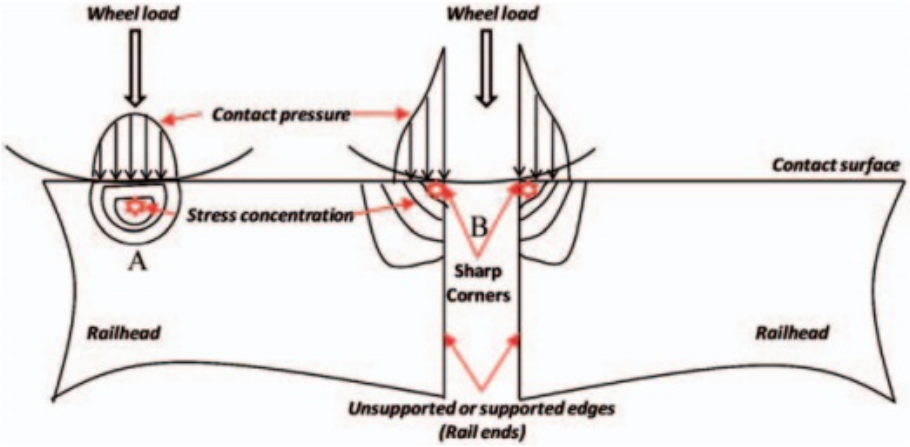


Figure 2.4: Wheel contact loading at a gap-jointed rail (Zong and Dhanasekar, 2013)

(a) Rail buckling caused by thermal expansion
(Zwolski, 2012)

(b) Rail break caused by thermal contraction
(Zwolski, 2012)

Figure 2.5: Consequences of thermal expansion/contraction on CWR

Rail Failure Mechanisms

Rail has many failing mechanisms, most of which result from fatigue caused by large forces. These larger forces on the rails normally occur around curve sections of track and hence rail installed on curved track tends to develop more faults. This is shown by Table 2.2, where Sawley (2001) estimated the rail life in Million Gross Tonnage (MGT) for different radii curves on the US railway network. Table 2.2 shows the dramatic effect that the increased loading that occurs at curved sections of track has on the rails serviceable life. Rail degrade by a combination of fatigue and wear, leading to many types of faults. There is also an environmental impact as rails rust, weakening them. Within this section a brief outline of the primary failure mechanisms are discussed.

Table 2.2: Estimated life of rail on track sections of varying curvature (Sawley, 2001)

Curve Radius [m]	Estimated Rail Life [MGT]
Straight	1460
1,750	1050
875	640
580	540
440	510
350	440
290	390
250	380
220	370
190	350
175	330

Rail Corrugations

Rail corrugations are cyclic, commonly vertical, irregularities on the rail surface. They can be short pitch with wavelengths between 30 and 90 mm (0.2-0.3 mm deep) or long pitch with wavelengths above 300 mm (0.1-2.0 mm deep), (Wilson, 2012). An example of short pitch corrugations can be seen in Figure 2.6.

Short pitch corrugations are caused by repetitive longitudinal oscillations of the wheel sliding on the rail and are commonly made by lighter axle loads (<20 tonnes). Long pitch corrugations develop due to plastic flow of the steel rail caused by high wheel/rail contact stresses. Heavier and faster trains, higher track and bogie stiffness, smaller wheel/rail contact interface and softer rails all increase the amount of long pitch corrugations.

Rail corrugations are discovered by visual inspections and act to increase dynamic wheel loads and vibration, thereby escalating the rate of degradation and failure of many track

and vehicle components, (Smith, 2005).

Corrugations are minimised by using higher strength steel rails to sustain greater contact stresses and reduce plastic flow. A good wheel/rail profile has a large contact area, which can minimise the contact stress. Corrugations can also be lessened by reducing the track stiffness by using softer and thicker rail pads. Corrugations are removed by grinding the rail until it is flat and smooth, whilst ensuring that it has the correct cross-sectional profile, (Wilson, 2012).

Figure 2.6: Short pitch corrugations (Wilson, 2012)

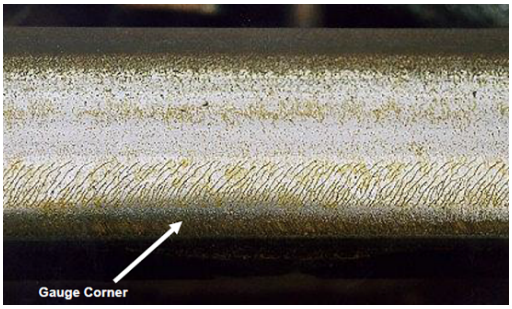
Rolling Contact Fatigue

Rolling Contact Fatique (RCF) is a common defect in most railway systems. Contact fatigue describes defects that arise from the development of excessive shear stresses at the wheel/rail interface. Common types of RCF are checking, shelling and flaking.

Checking occurs primarily on the high rail in sharp corners (gauge corner) and can look like fish scales, as seen in Figure 2.7a. Cracks initiate from the rail surface, or just below, every 2-5 mm along the rail. These cracks can then grow to 2-5 mm deep at which point they tend to break out as small wedges. Flaking is similar, but occurs on both the high and low rails of a corner on the running surface of the rail. An example can be seen in Figure 2.7b.

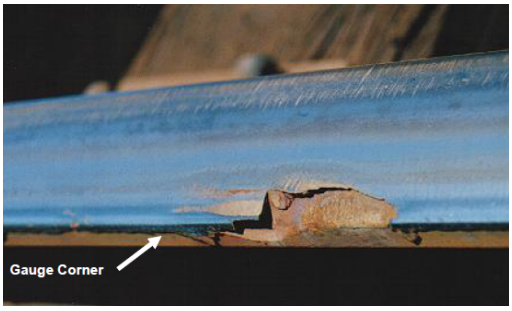
Shelling also generally occurs on the high rail of curved track and is an internal defect about 2-8 mm deep on the gauge side of the rail head, with the early stages of shelling noticeable as dark spots at the gauge corner of the rail. The defect can break off into a shell, as in Figure 2.7c, or can grow downwards and form a defect in the transverse plane, as seen in Figure 2.7d, which, if not discovered, can result in a rail failure, (Smith, 2005).

RCF occurs due to shear stress above the serviceable limits of the rail material in the wheel/rail contact area. This stress is caused by nominal, dynamic and impact wheel loadings which are affected by many factors including track geometry, bogie characteristics, wheel and rail irregularities and track cant. RCF is normally found on the high rail around corners due to the larger forces experienced on this rail, (Wilson, 2012).

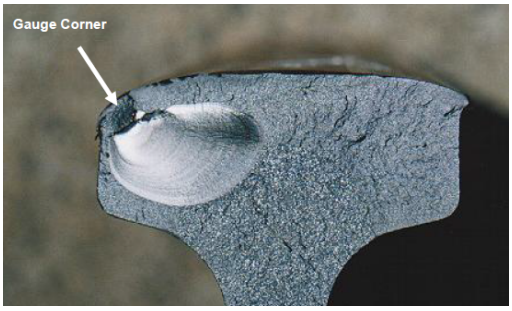


(a) Gauge corner checking

(b) Flaking with minor spalling



(c) Severe stage of shelling



(d) Transverse defect initiated from shelling

Figure 2.7: Rolling Contact Fatiques defects (Wilson, 2012)

Squat Defects

Squats are subsurface laminations which initiate at small cracks. These then extend diagonally downwards, reaching depths of 4-6 mm. They then spread laterally and longitudinally across and along the rail surface, (Wilson, 2012).

There are two categories of commonly occurring squats: running surface squats and gauge corner squats. Running surface squats are caused by thermal traction effects associated with wheel slips, which cause elevated surface hardness. Gauge corner squats initiate from pre-existing cracks, such as RCF, and usually occur on the gauge side of the rail head on the high rail of corners.

Squats are visible as a darkened area on the running surface of the rail head caused by a slight depression and are often double-sided kidney shapes, as seen in Figure 2.8. Squats can occur in irregular patterns on one or both of the rails and gradually develop over many months or years, (Grassie, 2011).

The depression in the running surface caused by squats can cause increased vertical impact loading, which exacerbates the degradation of other assets in the track structure, as well as creating increased vibration and noise. Ultrasonic testing can be used to see the depth and length of the squats sub-surface cracks. Once detected, squats can be removed by grinding of the rail but due to their depth a large quantity of metal often needs to be removed (as much as 6 mm), which greatly reduces the usable life of a rail, (Wilson, 2012).

Figure 2.8: Running surface squat (Wilson, 2012)

Tache Ovale

These are internal defects that propagate from near the centre of the rail head, growing transversely. An example can be seen in Figure 2.9. Due to the internal nature of tache ovale they cannot be detected by visual inspection but are instead discovered by regular ultrasonic rail inspections.

Tache ovales are created by the presence of excessive levels of hydrogen in the rail steel or weld (caused by cooling steel too quickly). The build up of these hydrogen molecules leads to increased internal pressure and can cause a crack to initiate, (Smith, 2005).

Due to being a rail production problem, it is possible that the same piece of rail develops many tache ovales. This can lead to catastrophic failure of the rail, especially under high impact loads.

To avoid the appearance of tache ovale defects the critical hydrogen content in the rail is reduced by appropriate steel making and heat treatment procedures. If rails affected by tache ovale are in the track, the growth can be inhibited by reducing normal, dynamic and wheel impact loadings by methods such as speed reductions. Ultrasonic testing is used to monitor the tache ovale until they become a critical size, at which point the rail is replaced, (Wilson, 2012).

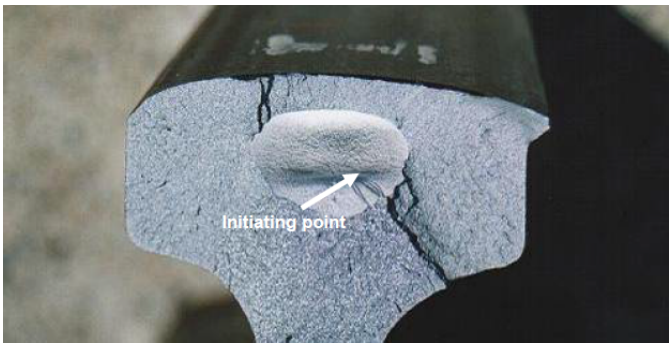


Figure 2.9: Tache ovale defect (Wilson, 2012)

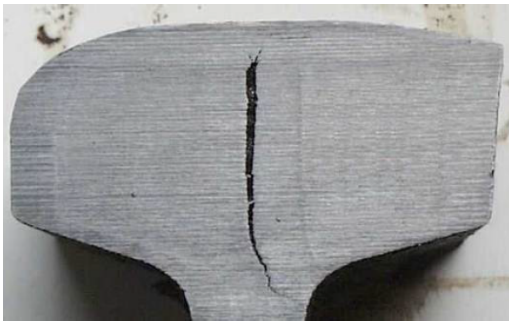
Vertical and Horizontal Split Head

When a vertical crack appears within the rail head, as in Figure 2.10a, it is known as a vertical split head and this can cause the rail head to split in two. Alternatively a horizontal defect can occur. The defects cannot be detected visually until it is already large. An example of an externally visible vertical split head can be seen in Figure 2.10b. The size of a split head can be determined using ultrasonic testing, (Profillidis, 2006).

The vast majority of split heads occur due to impurities in the steel. They are much more common in older rails. The initial crack propagates vertically (or horizontally) as well as experiencing longitudinal growth. Split heads are often caused by heavier axles and impact loads as well as off-centre contact between hollow wheels and a worn rail.

Where the crack extends along the rail it can lead to a large part of the rail head being weakened. If this is not detected and rectified a complete vertical failure of the rail head can occur, which would increase the risk of derailment, (Profillidis, 2006).

Split heads are becoming less common due to better steel rail manufacturing but can still occur in modern rails under extreme rail wear conditions. Once detected on the track the crack growth is monitored using ultrasonic inspections, and can be minimised by reducing the applied nominal, dynamic and most importantly wheel impact loading stresses by measures such as, speed restrictions or changes to maximum axle loads. Split heads can be minimised by grinding the rails to a shape where the wheel loading is concentrated on the centre of the rail head, (Wilson, 2012).



(a) Vertical split head internal view

(b) Vertical split head external view

Figure 2.10: Vertical split head defects (Wilson, 2012)

Wheel Burns

Wheel burns are defects on the running surface of the rail that appear in pairs opposite each other on the two parallel rails and can reach lengths of more than 50 mm.

Wheel burns are caused by the continuous slipping of the wheels, increasing the temperature near the surface of the rail, which then cools quickly causing the steel rail to transform into the hard and brittle martensite phase. Wheel slipping is commonly caused by excessive track gradients, rapid acceleration of trains or by contamination of the running surface.

Wheel burn defects tend to break up, leading to a depression in the running surface, and hence increased impact loads. These impact loads exacerbate the degradation of both the track and the bogie. As with other transverse defects, if they are not detected in time a complete vertical failure of the rail head may occur. Due to the high impact loading caused from the wheel burn these transverse defects tend to propagate at a faster rate than other defects.

To reduce the likelihood of wheel burns, rail lubrication must be applied correctly in the right amounts. The trains must also be driven sensibly with no rapid acceleration. It is also recommended that the rail head is cleaned of any contaminants by high pressure water spray, especially after a long dry spell of weather. Ultrasonic inspection should be undertaken regularly to monitor the transverse defects under the wheel burn before they reach a critical size, (Wilson, 2012).

2.3.3 Sleepers

Sleepers are traverse ties on which the rails are connected. The sleepers have many purposes, including, (Swarnakar, 2012):

- Even transfer of the load from the rails to the ballast;
- Holding rails at the correct gauge and alignment;
- Supporting the rails evenly;
- Absorbing blows and vibrations from moving loads;
- Providing longitudinal and lateral stability;
- Providing an easy way to correct track geometry faults.

Sleepers are typically manufactured from timber, reinforced concrete or steel. In the UK there are 22,200 km of concrete, 4,700 km of timber and 2,800 km of steel sleepers, with 200,000 timber sleepers being replaced by concrete sleepers every year (Williams, 2012). Each material choice brings advantages and disadvantages, which are discussed below.

Timber Sleepers

Timber sleepers are usually made from hardwood such as oak but softwoods can be used on lighter lines. In the USA timber sleepers are primarily used, with 93% of all sleepers in use being made of timber. The main advantages of timber sleepers are, (Li, 2012; Profillidis, 2006):

- Are relatively light and easy to handle and install, with no required plant;
- Are cheap and easy to manufacture (dependent on country of production);
- Can be sustainably sourced;
- Have a good capacity to absorb shocks and vibrations;
- Acts as an electrical insulator allowing track circuits for signalling;

- Are flexible, resulting in better load distribution and hence are useful on poor subgrades;
- Are adaptable and available in multiple sizes with the ability to be cut down on site making them especially useful for non-standard situations such as S&C;
- Can be reused as garden landscaping or biomass fuel if untreated.

The main disadvantages are:

- Their limited lifetime, typically 20-50 years if treated, otherwise 7-12 years;
- They can be damaged by wear, decay and vermin; Environmental factors such as sunlight and moisture have a large effect on a timber sleepers service life;
- Their low weight leads to low transverse resistance, making them unsuitable for high speed lines;
- Smaller sleeper gaps are required when compared to concrete, resulting in more sleepers per km;
- If treated with a product such as coal tar creosote, which is a toxic hazard, there are added disposal costs and the sustainable sleepers lose environmental credibility;
- Maintaining gauge can be difficult as the sleeper degrades;
- Susceptibility to fire hazards.

Degradation Mechanisms

Wooden sleepers generally degrade slowly due to wear beneath the baseplate, as well as wearing of the fixing holes for screws and spikes that hold the rail or baseplate in place. Both of these processes occur due to rail traffic causing movement between the separate parts of the fastening system. The other major degradation mechanism is the natural process of rot and decay of the timber which can be exacerbated by vermin. The timber sleepers on most highly trafficked lines in the UK have been replaced by concrete sleepers, with timber sleepers tending to remain on low trafficked lines. This means that wear from traffic has reduced, leaving rot and decay as the primary reason for sleeper replacement. Timber sleepers are visually inspected and replaced individually if required (Williams, 2012).

Concrete Sleepers

Pre-stressed, reinforced, mono-block, concrete sleepers have become a replacement for timber sleepers in many parts of the world including Europe, Australia and Japan. The main advantages of concrete sleepers are, (Swarnakar, 2012; Li, 2012):

- Their long service life, typically between 45-55 years;
- Their high weight leading to increased track stability, which makes them better suited to CWR than the alternatives;
- Increased stability makes concrete sleepers the best option for high speed lines;
- The elastic fastening systems used with them helps maintain the correct gauge,

cross level and alignment;
- They act as an electrical insulator allowing track circuits for signalling;
- They are inflammable and not affected by vermin with minimal corrosion;
- They are easy to mass produce;
- Larger sleeper spacing possible leading to fewer sleepers per km in comparison to alternatives.

The main disadvantages are:

- Their vulnerability to impact damage;
- Their high weight, typically between 215 kg and 270 kg, makes them difficult to handle meaning that plant equipment is required for installation, resulting in longer installation times;
- They are relatively expensive, especially if outside of standard sizes as required at S&C;
- They are susceptible to heavy damage in derailments, thus requiring replacements (Zakeri and Rezvani, 2012);
- Having greater height than other sleeper types makes concrete sleepers less well suited to routes with minimal height clearance, such as in tunnels;
- They have minimal scrap value;
- They are not sustainable and have a large environmental cost due to their production requiring a large amount of energy;
- They can be damaged in transportation.

Degradation Mechanisms

Concrete sleepers generally degrade slowly. The rail seat area tends to wear with the rate being increased by thin or worn rail pads. Sleepers can also be physically damaged by activities such as tamping. Poor ballast conditions can lead to soffit attrition and cracking due to inconsistent support. It is also possible for spalling to occur due to rusting of the steel reinforcement causing them to delaminate the concrete. Concrete sleepers can still function as required with cracks but when they fail the load on adjacent sleepers and rail is increased and this can lead to localised geometry faults. Concrete sleepers can be individually replaced but rail seat wear due to ineffective rail pads will generally affect a large section of sleepers, meaning that all need to be replaced. This means that the rail pad is of great importance (Williams, 2012).

Steel Sleepers

Steel sleepers are hollow, open bottom rectangular cuboid's with rounded edges that taper out near the ends for transverse track stability. Their main advantages are, (Li, 2012; Profillidis, 2006):

- Their long service life, typically between 30-60 years;

- They are light and easy to install but tamping is required to force ballast inside;
- They are inflammable and not affected by vermin;
- They are recyclable with good scrap value.

The main disadvantages are:

- They are liable to corrosion;
- They are unsuitable for areas with track circuits for signalling due to electrical conductivity;
- They have low transverse resistance resulting in poor high speed line performance;
- They can develop cracks at the rail seat.

Degradation Mechanisms

Due to the relatively recent introduction of steel sleepers, their failure mechanisms in the field are not understood to the same degree as concrete and timber sleepers. Fatigue is the primary source of failure which causes cracking in and around the rail seat (ARTC, 2009). Since failure is caused by fatigue, failure occurs suddenly with no prior indication. To reduce fatigue it is important to prevent excessive fatigue loads such as those caused by poor ballast, which results in a particular sleeper taking more load than adjacent sleepers. Steel sleepers can rust in certain conditions, which can greatly weaken them. Like other sleepers, steel sleepers can be spot replaced if required.

2.3.4 Ballast

The ballast is the platform on which the sleepers are laid. It consists of crushed rock which can vary due to requirement and available local sources. Granite, traprock, quartzite, basalt, carbonate rocks and slag are some of the most commonly used materials, (AREMA, 2012). Good ballast is composed of angular, crushed hard stones and rocks, uniformly graded, free of dust and dirt and not prone to cementing action, (Selig and Waters, 1994). The stones have to be able to withstand large loads without being crushed, and they must be angular to enable interlocking in order to create a strong assembly of particles. Ballast has many functions, including, Profillidis (2006); Calla (2003):

- Distributing stresses transmitted by the sleepers down to the sub-ballast;
- Attenuating train vibrations and sound;
- Resisting the track shifting both transversely and longitudinally;
- Allowing rainwater drainage;
- Facilitating track geometry to be restored through maintenance.

Under the ballast layer there is a sub-ballast layer that consists of gravel and is commonly around 15 cm thick. The sub-ballast functions to, (Profillidis, 2006):

- Protect the subgrade from the intrusion of ballast stones;

- Distribute stresses from the ballast down to the subgrade;
- Allow rainwater run-off;
- Create a transverse slope (of 3-5%) in the upper part of the subgrade to allow water to run-off.

The granulometric composition of ballast is important for it to fulfil its role correctly. For fresh, clean ballast a comparison of two common composition limits can be seen in Figure 2.11, which was created using information from BS EN 13450 (2002) and AREMA (2012). It can be seen from Figure 2.11 that the AREMA boundaries tend to be more well graded, allowing smaller particles but also having more large particles than the BSi option which is more uniformly graded. Despite the slight differences, both agree that the majority of the ballast should have particle sizes of 20-60 mm.

The material and particle size are not the only geometric characteristics for good ballast. It is also important that the ballast has a minimal quantity of fines and fine particles, which are classified as passing sieve sizes of 0.063 mm and 0.500 mm respectively, (BS EN 13450, 2002; Profillidis, 2006). The large stone sizes and minimal inclusion of fines helps the ballast to drain quickly. The particle shape is also an important factor with a flakiness index, shape index and particle length boundaries being checked for new ballast. The mechanical properties of the ballast are also tested to quantify its resistance to fragmentation and wear, which has set boundaries (Aursudkij, 2007; BS EN 13450, 2002).

Track settlement mainly occurs within the ballast and is measured using a specially designed train, such as the New Measurement Train used by Network Rail. This train can travel at up to 125 mph and uses laser scanning equipment to measure the track geometry changes every 0.2 m. This is possible with the inclusions of gyroscopes, accelerometers and transducers to detect the movement of the train. As a train travels down the track it will be effected by undulations in the track profile. When it comes to effecting the passage of the train, long smooth undulations do not affect the safety of the trains passage and are not felt within the train so these are removed to leave only the more sudden changes in vertical profile that may have a detrimental effect on the running trains. To remove these smooth changes, which are represented by long wavelengths within the vertical top profile, a filter is used. Network Rail (NR) use a 35 m wavelength four-pole high pass filters (Lewis, 2011). The vertical Standard Deviation (SD) is the most-used track geometry measurement for indicating the track condition, as vertical geometry tends to degrade fastest and is also the main factor that effects the ride quality and possible maximum speed (Network Rail, 2012). The SD is typically calculated from the average height of both rails (after the 35 m filter has been applied) for 220 yard sections of track and can be found using Equation 2.1, where m is the total number of values, a is the mean value and a_1 is the sample value, (Dingwall, 1998).

$$SD = \sqrt{\frac{1}{m}\Sigma(a_1 - a)^2} \tag{2.1}$$

An inflated example of the vertical top profile of a poskey, after a 35 m wavelength filter has been applied, is shown in Figure 2.12. The left and right rails are recoded separately and an average profile found, which is then used to calculate the SD in mm.

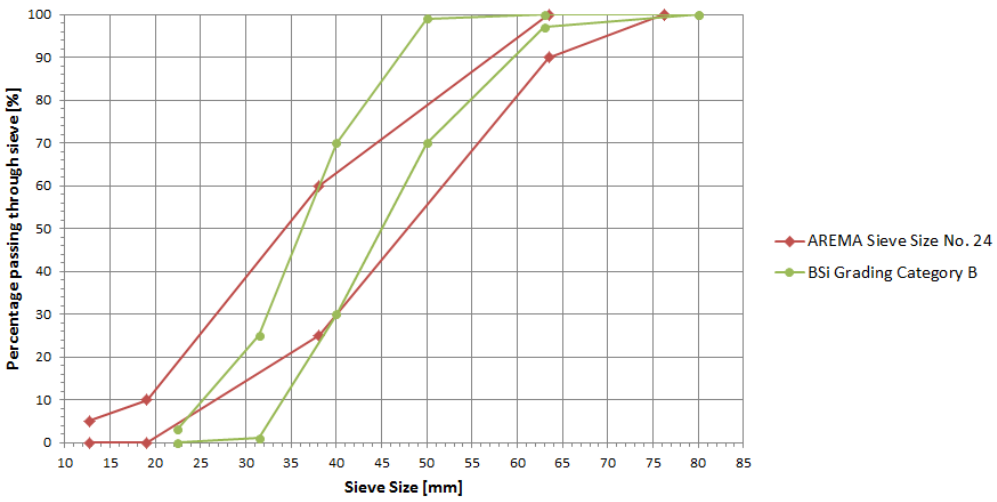


Figure 2.11: Comparison between AREMA Number 24 and BSi Category B ballast grading boundaries

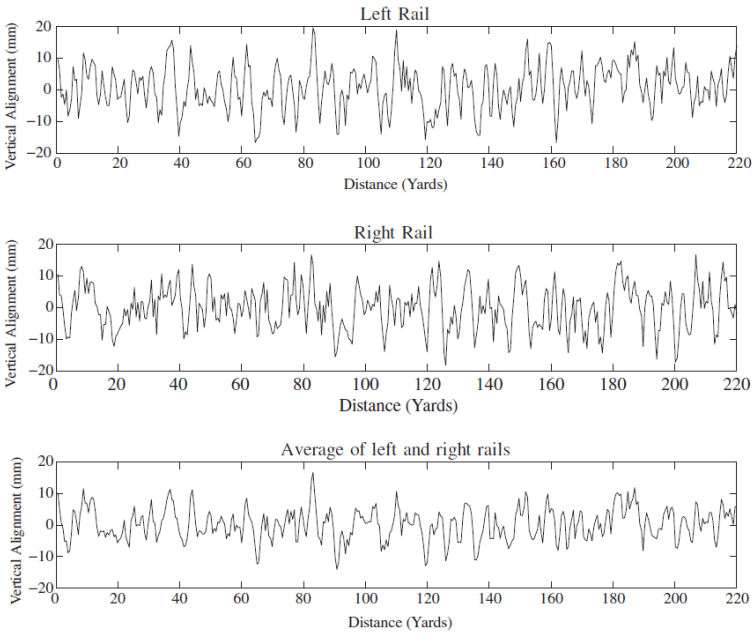


Figure 2.12: Exaggerated vertical top profile after 35 m wavelength filter, based on geometry recordings from a measurement train (Audley and Andrews, 2013)

Degradation Mechanisms

Fouling

Ballast fouling occurs when the ballast contains small particles, which are described by Selig and Waters (1994) as particles smaller than 4.76 mm in diameter, and in BS EN

13450 (2002) as those particles that pass through a 5 mm sieve. The causes of fouling are, (Selig and Waters, 1994; Williams, 2012):

- Ballast breakdown caused by dynamic forces from traffic leading to the ballast particles to rub together as well as maintenance activities such as tamping;
- Infiltration from the surface (such as coal spilling and airborne dirt);
- Sleeper wear;
- Infiltration from underlying granular layers;
- Infiltration from the subgrade.

The main sources of fouling in British and North American railways are given by Selig and Waters (1994). In British railways the main causes of fouling is wagon spillage at 43%, followed by traffic and sleeper wear, which includes ballast breakage at 21%, and tamping at 20%. In North America 76% of fouling is due to ballast breakage, 13% related to underlying granular layers and 7% caused by surface infiltration.

Ballast fouling impedes the drainage path, resulting in the subgrade becoming saturated. This increases the soil's plasticity, resulting in faster permanent deformation (Kirchhof, 2006). Lichtberger (2005) indicates that the critical Ballast Fouling Index (BFI) is 30%, as above this level the amount of run-off is minimal as seen in Figure 2.13, where the BFI is used to quantify the degree of fouling. Wet beds can appear in highly fouled ballast; these can be seen from the surface where a slurry of water and fines mix and can increase the degradation rate of sleepers and rails. As fouling increases the distribution of loads is reduced, as seen in Figure 2.14. This leads to faster settlement within the subgrade due to the increased concentration of the load and also increased ballast contamination from subgrade infiltration (Williams, 2012). Increases in fouling also decrease the ballast's resilience to deformation as the fines coat the contact points between large particles, acting like a lubricant, lowering the strength of the particle network and increasing the ballast stones movement, (Ebrahimi, 2011b). The degree of fouling affects how quickly the vertical geometry degrades, with higher fouling resulting in faster degradation as seen from Figure 2.15, (Williams, 2012). The vertical geometry can be maintained by tamping, but this greatly increases ballast breakup with it being estimated that every tamping operation produces as many fines as 20 MGT of traffic, (Ottomanelli et al., 2005).

Fouling also has an impact on the effectiveness of tamping and hence the ability to restore good track geometry. As fouling occurs and the voids in between the large ballast stones fill, resulting in denser ballast. Tamping then loosens the stones and particles creating voids, which under load, as the small particles move, are quickly filled. This leads to a higher rate of ballast settlement after tamping than experienced with new, clean ballast (Aursudkij, 2007). This leads to a higher frequency of maintenance being required as shown by Figure 2.16, (Williams, 2012). It is noted that Figure 2.16 includes stoneblowing maintenance, which is much more effective on fouled ballast, so for higher levels of fouling the relative frequency would be much greater for tamping alone,

which is estimated to be four times higher for choked ballast, where all the voids are filled with fines and fine particles. It is for this reason that stoneblowing is the maintenance of choice for heavily fouled ballast (Williams, 2012). Tamping and stoneblowing are described in Section 2.3.4.

High degrees of fouling such as those that are present when wet beds appear can be identified during visual inspection. However, lesser degrees of fouling cannot be identified by visual inspection. In such cases equipment such as Ground Penetrating Radar (GPR) is used. From the readings obtained from the GPR, it is possible to calculate the BFI, enabling the correct decision on maintenance or renewal action to be made. The BFI is a percentage that is used to categorise the degree of fouling. It can be calculated from Equation 2.2, where P_4 and P_{200} are the percentage of ballast particles that pass through the 4.75 mm and 0.075 mm sieves.

$$BFI = P_4 + P_{200} \tag{2.2}$$

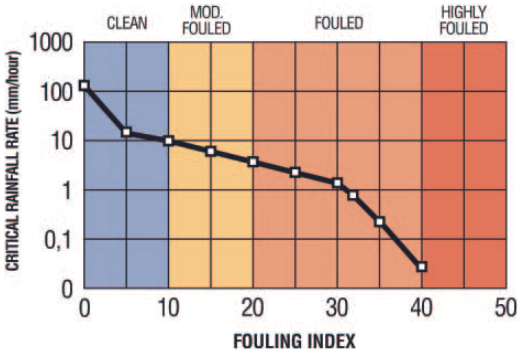


Figure 2.13: Critical quantity of precipitation depending on the degree of fouling (Lichtberger, 2005)

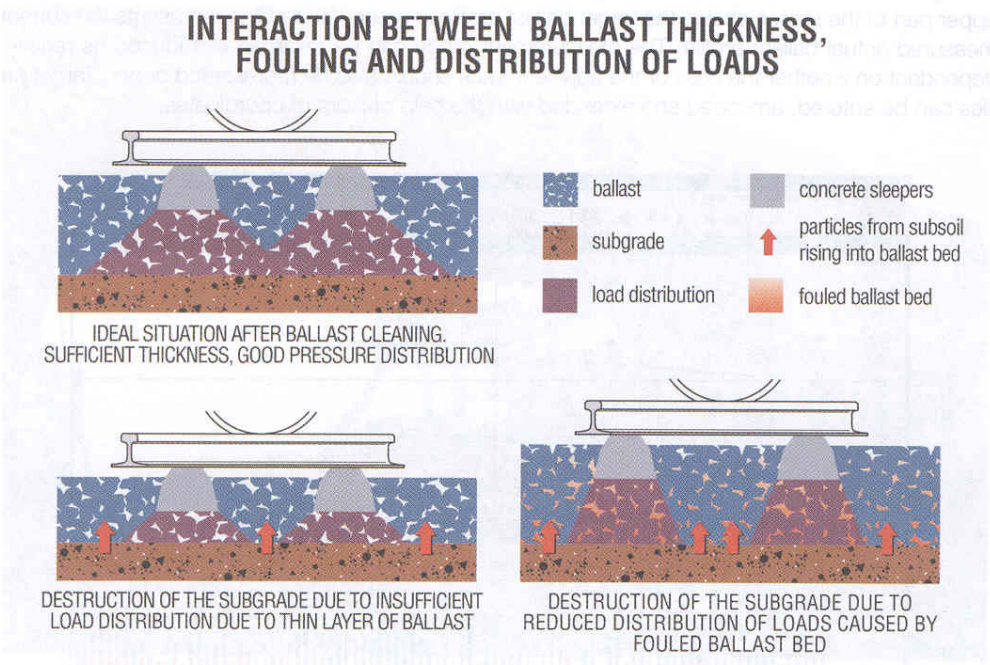


Figure 2.14: Schematic representation of fouled ballasts reduced bearing capacity (Lichtberger, 2005)

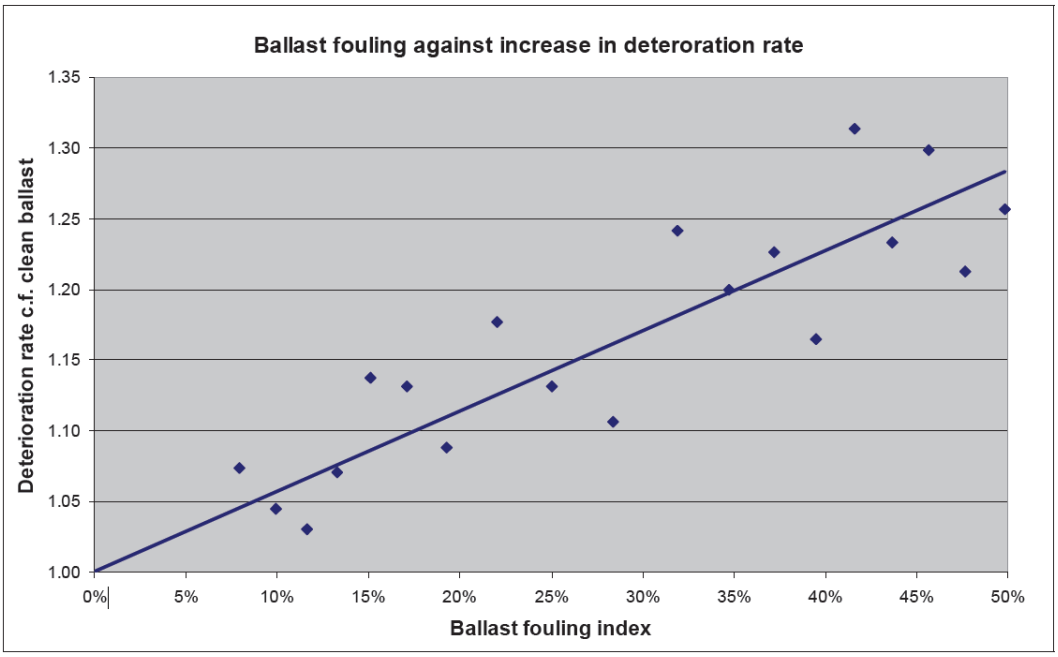


Figure 2.15: Increase in track geometry degradation rate with ballast fouling (Williams, 2012)

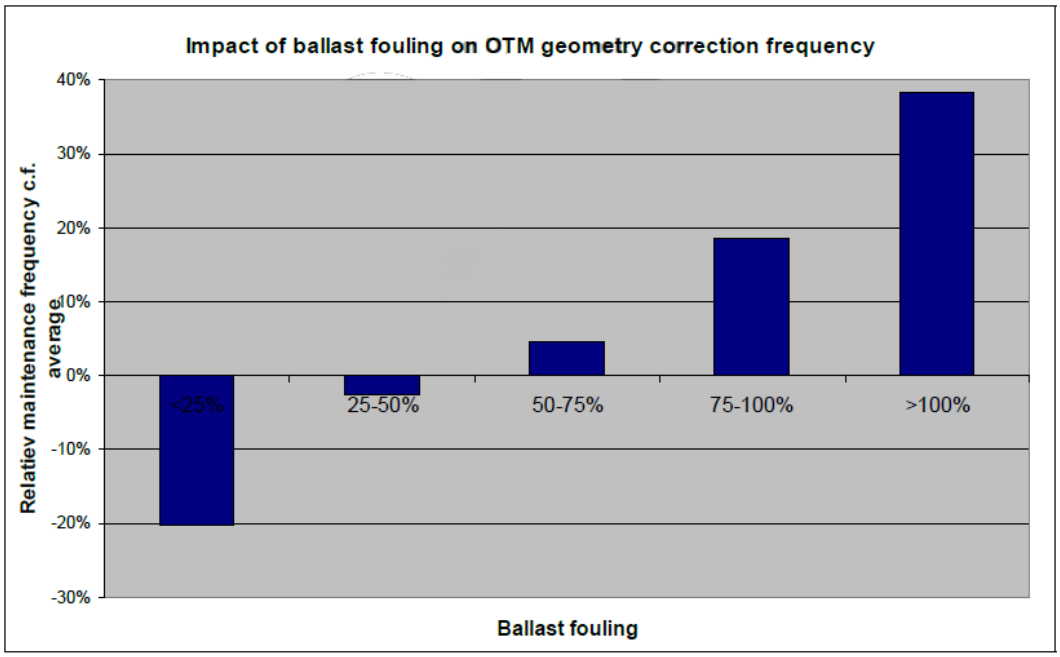


Figure 2.16: Increase of maintenance frequency due to ballast fouling (Williams, 2012)

Maintenance Techniques

Ballast maintenance restores the track geometry and involves either tamping or stoneblowing, with small faults being corrected manually. When the ballast reaches a highly fouled state general maintenance has less of an effect so more radical maintenance, such as a renewal, is performed. This can involve cleaning the ballast and reusing a proportion of it or carrying out a complete renewal in which all the ballast is replaced.

Tamping

Tamping is the most common form of maintenance for the correction of track geometry faults. Tamping is undertaken by a specialised tamping train which uses tines to correct the track level. The process is outlined by Figure 2.17a, where:

A: Track profile needs correcting.
B: Tamping train lifts the rail and sleeper to a target level, leaving an empty space below the sleeper.
C: The tamping tines are inserted on both sides of the sleeper. This causes ballast breakage.
D: The vibrating tines squeeze ballast into the gap under the sleeper, recovering the correct position of the sleepers. This also causes ballast breakage.
E: The tines are removed, leaving the track in the correct position and the tamping machine moves on to the next sleeper.

After a tamping operation is complete, fast degradation of the track geometry occurs due to the ballast skeleton being loosened by the tamper. Once an equilibrium has been

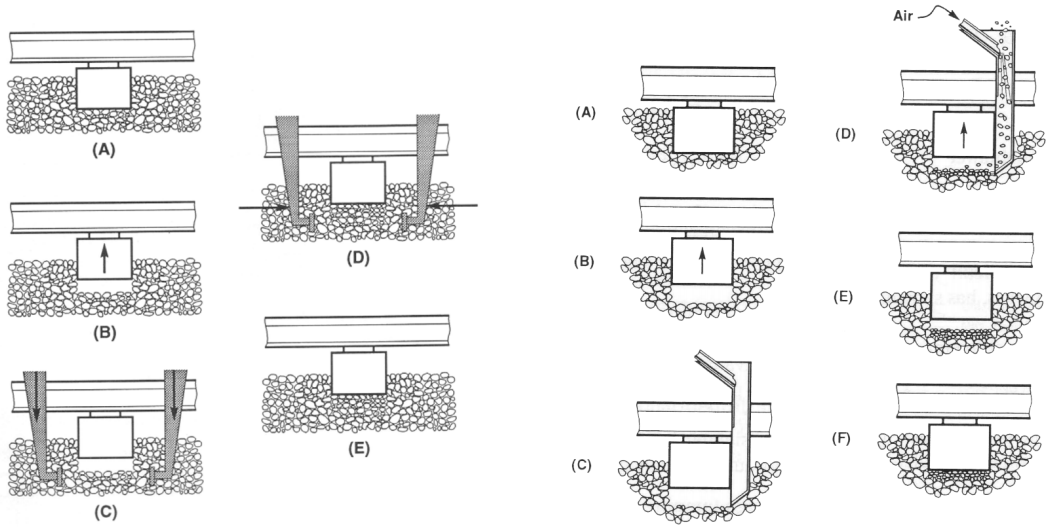
reached in the ballast, the rate of degradation will reduce but the track geometry will reach the same level as observed prior to tamping quicker than it did before the tamping operation took place. This effect is known as ballast memory and is reduced by lifting the sleepers to a higher level than ultimately required in order to account for the fast settlement which will occur directly after tamping (Aursudkij, 2007). Ballast tamping is most effective on single graded ballast stones. Hence, as ballast fouling increases tamping has less of an effect on prolonging degradation. Acting in combination with ballast memory, this effect of fouling reduces the time interval between required tamps after each tamping operation, until it is no longer economical to undertake tamping operations (Selig and Waters, 1994). When this is the case, stoneblowing can be performed. Ballast cleaning or complete renewal of the ballast can also be considered.

Stoneblowing

Stoneblowing is a more modern alternative to tamping but is more expensive and slower. However, it has many advantages over tamping, including causing much less ballast breakup with Wright (1983) (as cited by Aursudkij (2007)) showing that tamping produced eight times more particles below 14 mm per operation than stoneblowing. The process of stoneblowing is demonstrated in Figure 2.17b, where:

A: Track profile needs correcting.
B: Stoneblowing train lifts the rail and sleeper to a target level, leaving an empty space below the sleeper.
C: Stoneblowing tubes are inserted into the ballast next to the sleeper.
D: Compressed air blows a measured quantity of stones into the space below the sleeper. The stones are normally about 20 mm in size.
E: The tubes are removed from the ballast.
F: Sleeper is lowered onto the new stones.

Another advantage stoneblowing has over tamping is that it can be used even when ballast is fouled, returning the track to a better condition than tamping would with a reduced degradation rate after. This leads to a longer time period between successive stoneblowing operations. Stoneblowing is particularly effective for smaller geometry corrections where small lifts, less than one inch, are used. Due to the small size of the injected stones stoneblowing is not as well suited to larger lifts which are greater than one inch. For these larger lifts the ballast rearrangement caused by tamping is more effective (Zarembski, 2005). The general rule of thumb is that for fresh, clean ballast tamping is used and later in life, as the ballast becomes fouled and tamping becomes less effective, stoneblowing is used, until the ballast condition due to fouling becomes unacceptable. At this point ballast cleaning or a ballast renewal is undertaken.



(a) Tamping sequence (b) Stoneblowing sequence

Figure 2.17: Ballast maintenance sequences (Selig and Waters, 1994)

2.3.5 Formation and Subgrade

The formation layer below the ballast layers is typically made from sand and designed with a slope from the track centre to the side to enable good drainage and run-off. Below this is the subgrade, a prepared (compacted) natural layer, which forms the base of the track system. The subgrade can be strengthened by activities such as lime stabilisation and the use of geogrids, reducing the amount of settlement that occurs. These layers can be inspected using GPR but maintenance is not possible without a complete track renewal.

Degradation Mechanisms

The formation and subgrade are affected by settlement in the same way as the ballast, where loading causes the particle skeleton to change, reducing the void ratio, which ultimately causes track geometry faults. Despite having similar degradation mechanisms it is shown by Selig and Waters (1994) that the majority of the overall track settlement and hence geometry faults occur within the ballast and not the subgrade, Figure 2.18.

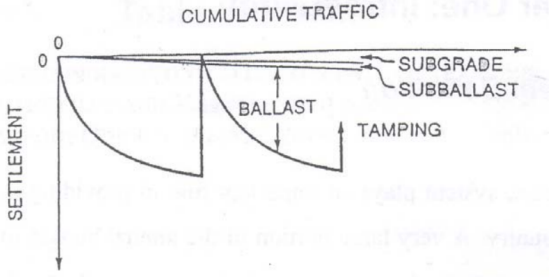


Figure 2.18: Track layer contributions to settlement (Selig and Waters, 1994)

2.3.6 Summary of Railway Track Assets

A railway track consists of rails, fastenings and rail pads, sleepers, ballast, formation and subgrade, with the forces produced from passing trains occurring on the rails and moving downwards.

The rail which provides a smooth running surface can either be CWR or jointed, with CWR being the modern preference. The fishplated joints are the weakest part of the rail, with the increased dynamic forces from the dip damaging the rails and the other assets. CWR also has problems; changes in temperature, which causes the rail to expand or contract can lead to buckling or breaks unless there is the required lateral resistance given by the ballast on the sleepers. The majority of rail defects are caused by fatigue and are more likely to occur when under higher stresses, such as the top rail around a bend, or impact damage, such as by joints. These rail defects can cause increased forces to occur, due to irregularities in the rail running surface, leading to the faster degradation of other assets, such as sleepers and ballast.

There are three main types of sleepers used; timber, reinforced concrete or steel. The sleepers transfer the load from the rails whilst supporting them evenly and providing longitudinal and lateral stability. Timber sleepers are cheap, renewable, recyclable, light, adaptable and easy to install, but they have a limited lifespan between 20-50 years and their low weight decreases their transverse resistance making them unsuitable for high speed lines. Gauge tends to degrade faster with timber sleepers due to the available fastening systems and the timber degrading. Timber sleepers are more dependent on environmental conditions than other sleeper types with moisture or sunlight causing rot and decay as well as vermin causing damage. As timber sleepers are more prone to environmental damage the lifespan of timber sleepers can be more time dependant than usage. Wear of the sleeper below the baseplate is also common and can cause geometry problems. Concrete sleepers have longer serviceable life spans between 45-50 years and are much more suited to higher speed lines due to their high weight. Embedded anchors for the fastenings also reduces the degradation of the gauge. Being very heavy, concrete sleepers are slower to install as they require plant equipment. Unlike timber sleepers, concrete sleepers can not be sustainably sourced, require large amounts of energy to produce and cannot be recycled easily leading to minimal scrap value. Concrete sleepers tend to fail due to either wear of the rail seat area or cracking caused by inconsistent support from the ballast. Steel sleepers are the newest form of sleepers used in the UK and can have a long service life between 30-60 years. They are light and easy to install and recyclable, but they can corrode in certain climates and their low transverse resistance makes them unsuitable for fast lines. Fatigue causing cracking around the rail seat is the most common failure mode. With all sleeper types, if a sleeper is in a highly degraded state it will not support the rail correctly, increasing the load on the nearby sleepers. This increased load will cause higher rates of degradation of the sleepers, fastenings and rail pads and ballast. Due to the uneven loads the degradation will not

be constant leading to possible geometry faults. Sleepers that are not of the required condition to full-fill their role are replaced.

Ballast is uniformly graded and primarily consists of stones between 20-60 mm. The stones tend to be angular to enable stronger interlocking. The ballast transmits stresses down into the formation and allows rainwater drainage. It tends to settle quickly to begin with, as the ballast particles move into a compacted, strong skeleton, after which it tends to settle slower. Uneven settlement is the primary cause of geometry faults and hence speed restrictions. The track geometry is recorded by a measurement train, with the primary used measurement being the SD of the vertical geometry. Ballast degrades as it becomes fouled with fine particles which reduce drainage, leading to higher levels of saturation within the subgrade which causes faster settlement. The fine particles are generally produced from a combination of subgrade infiltration, wagon spillage, ballast breakage and tamping. These fine particles lubricate the ballast stone's edges, causing greater movement of the stones and increased geometry degradation. Fouled ballast is also less efficient at spreading the load of passing trains causing increased stresses within the subgrade and hence faster geometry degradation. Maintenance activities such as tamping and stoneblowing are used to rearrange the ballast particles and return the track to the required geometry. Tamping is generally more effective on newer ballast, with fouled ballast causing the track to tend to return to its previous condition at a faster rate. When tamping is not being effective, with the track geometry degrading quickly after completion, stoneblowing is commonly used instead, as it tends to be more effective on fouled ballast, reducing the rate at which the track geometry degrades after completion compared to tamping.

Below the ballast is the formation and subgrade. The formation layer is typically on a gradient and made from sand to allow rainwater run-off helping to maintain an unsaturated subgrade and ballast. The subgrade is the compacted, natural geological layer beneath the track. Stresses from passing trains cause settlement to occur, which, if not constant along the track, can cause geometry faults but the majority of the overall track settlement and hence geometry faults occurs within the ballast and not the subgrade layer.

2.4 Existing Railway Track Degradation and Maintenance Models

2.4.1 Introduction

There are many existing models for mapping the degradation of many of the railway track assets. These models are mainly used to analyse track geometry and defects, rail defects and sleeper degradation and can be classified as either deterministic or stochastic. Deterministic models are either empirical, developed through experimentation or

field data, or mechanistic, where the model is based around physical knowledge. They relate the factors that contribute to the degradation process being modelled generally through equations. They can be highly complex, requiring large quantities of input data, and during construction the modeller must ensure that all contributing factors are considered. Deterministic models also do not take into account the random nature of degradation, which can be illustrated by considering two identical pieces of track that, although used and maintained in the same way, will degrade differently (Selig and Waters, 1994). Mechanistic models are advantageous for new track sections or track with minimal prior information, as they do not use past degradation patterns to predict future behaviour. Stochastic models are usually constructed using historical records and data and use probability distributions to account for variability. Stochastic models account for all the factors that effect degradation since they are based on actual measurements of track performance, but without good past data their accuracy can be limited. Whilst stochastic models are based on actual track behaviour, they do not give any insight into the underlying physics of the degradation so it's causes can be unclear.

2.4.2 Track Geometry Degradation Models

Deterministic Models

Alva-Hurtado and Selig's Strain Model

The primary cause of track geometry degradation is plastic strain within the ballast. Many mathematical models have been developed to calculate this permanent strain, with one of the most widely used being developed by Alva-Hurtado and Selig (1981) (as cited by Shi 2009). By using results obtained from triaxial tests, Alva-Hurtado and Selig describe how the permanent strain, ε_N, after a number of cycles, N, is related to the permanent strain after one cycle, ε_1, by Equation 2.3. The model accounts for the fact that the initial loading of ballast causes a large amount of plastic deformation due to rearrangement of the stones as they try to reach a point of equilibrium. In subsequent loadings the amount of plastic strain reduces. The permanent strain after N cycles is given by:

$$\epsilon_N = \epsilon_1 \left(1 + C\log N\right) \tag{2.3}$$

Where C is a dimensionless constant, which controls the growth of deformation, with Selig and Waters (1994) recommending values between 0.2-0.4. Research by Office of Research and Experiments (ORE) and British Railway (as cited by Profillidis (2006)) suggested a C value of 0.2 whereas the American Railways suggested a value between 0.25-0.40.

Selig and Waters' Strain and Settlement Models

Selig and Waters (1994) proposed a different relationship that takes into account a non-linear relationship between strain and the log of the number of cycles. Using triaxial tests to relate the permanent strain, ε_N, the permanent strain after one cycle, ε_1, and the amount of load cycles, N, similarly to Alva-Hurtado and Selig (1981). Field measurements were used to replace cycles with traffic, as MGT. To take into account the non-linear relationship found, Selig and Waters (1994) compared semi-log, hyperbolic, parabolic and power mathematical relationships. It was found that the power law function, seen as Equations 2.4 and 2.5, closely fitted the data obtained:

$$\varepsilon_N = \varepsilon_1 N^b \tag{2.4}$$

$$\varepsilon_T = \varepsilon_{T1} T^c \tag{2.5}$$

Where b is a constant that within their experiment resulted in a value of 0.17, ε_{T1} is the strain after 1 MGT of traffic and T is the traffic load in MGT. Using the field measurements Selig and Waters found that $c = 0.21$, $\varepsilon_1 = 0.0035$ and $\varepsilon_{T1} = 0.026$.

As plastic strain is settlement divided by the layer thickness, Selig and Waters (1994) notes that a similar power law function can be used for the settlement, with the same value of b, where the settlement, S_N, after N cycles is given by:

$$S_N = S_1 N^b \tag{2.6}$$

Selig and Waters noted that when the log of the settlement was plotted against the log of the number of cycles a linear relationship was visible. Similarly to Equation 2.3, the equations suggested by Selig and Waters model very fast initial strain/settlement of the track with the rate reducing with increasing amounts of load cycles. Using the recommended values for Equations 2.3 and 2.6 to plot Figure 2.19 it can be seen that the model proposed by Alva-Hurtado and Selig (1981) is linear on a semi-logarithmic scale, whereas the Selig and Waters (1994) tends to increase slowly. The figure also demonstrates, taking the log scale into account, the fast initial degradation with the rate decreasing over additional load cycles as well as how Selig and Waters (1994) is much more pessimistic about the rate of strain with higher values obtained from Equation 2.5.

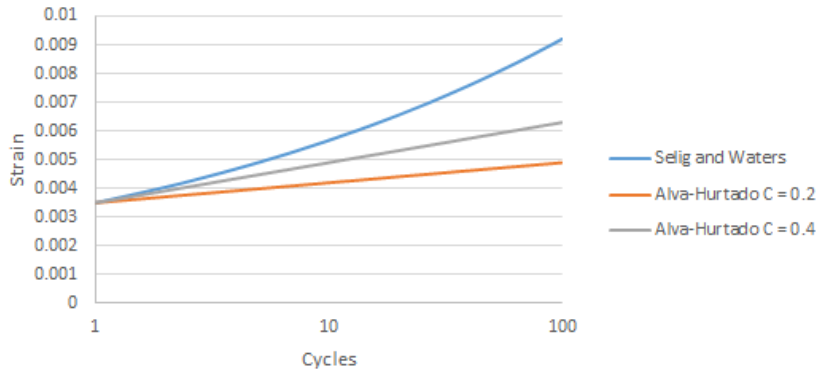


Figure 2.19: Comparison of strain/cycle models

Shenton's Settlement Model

An alternative settlement law was explored by Shenton (1985). Based on laboratory and field experiments, Shenton (1985) (as cited by Dahlberg 2007) developed the settlement law:

$$S_N = K_1 N^{0.2} + K_2 N \tag{2.7}$$

Where S_N is the settlement after N load cycles. K_1 and K_2 are constants, which depend on many factors, including axle load, rail section, sleepers spacing as well as the track and subgrade stiffness, previous maintenance and lift. Axel load is noted as the most important with lighter axles shown to have minimal impact on track geometry, (Shenton, 1985).

The experiments showed that settlement was proportional to the fifth root of the number of load cycles up to 10^6 cycles but that above 10^6 cycles this relationship did not hold. The values of K_1 and K_2 must therefore be specified in such a way that the first term of Equation 2.7 is dominant for $N < 10^6$ and that the second term is only significant for $N > 10^6$. As with the previously discussed settlement/strain models Equation 2.7 with appropriate values for K_1 and K_2 demonstrates early fast settlement which slows with further load cycles.

Sato's Settlement Model

Another model for determining ballast settlement under repeat loading, S_N, was suggested by Sato (1995) based on laboratory studies and was expressed as:

$$S_N = \gamma(1 - e^{-\alpha N}) + \beta N \tag{2.8}$$

Where the loading factor, N can be either the number of load cycles or cumulative traffic tonnage. The constants γ and α describe the long term settlement of the track and β

describes the short-term. One α is the vertical acceleration required to initiate slip, which can be measured using spring-loaded plates of the ballast material on a vibrating table. β is proportional to the sleeper pressure and peek acceleration experienced by the ballast particles and is effected by the condition of the ballast material as well as the presence of water. The value of γ is based on the quality of the initial packing of the ballast material.

TU Munich Settlement Model

At the Technical University of Munich, Iwnicki et al. (2000) developed another settlement model using laboratory experiments representative of vehicles passing over a dipped joint. A dipped joint is a joint which due to degradation dips below the normal rail top (height position) when traversed by a train. Three equations were developed to account for an optimistic, S_{opt}, pessimistic, S_{pes}, and median, S_{med}, levels of settlement as (Iwnicki et al., 2000):

$$S_{opt} = 1.57p\Delta N_a + 3.04p^{1.21} \log N_a \tag{2.9a}$$

$$S_{pes} = 2.33p\Delta N_a + 15.2p^{1.21} \log N_a \tag{2.9b}$$

$$S_{med} = 1.89p\Delta N_a + 5.15p^{1.21} \log N_a \tag{2.9c}$$

Where N_a is the number of axle passes and p is the pressure in the ballast which is calculated using the Zimmermann method, which involves a theoretical longitudinal sleeper placed under the track (Kuttelwascher, 2012). The first part of the equations gives the initial, quicker settlement which occurs immediately after maintenance and the second part relates to the more gradual, longer term settlement that occurs after 10,000 axle passes where an axle pass is one load cycle (Iwnicki et al., 2000). The decision that settlement begins to slow at 10,000 load cycles differs from Shenton (1985) who suggested that settlement starts to slow later, after 1,000,000 load cycles but this difference maybe due to the degree of packing the ballast was initially compacted to.

Office of Research and Experiments Track Geometry Models

Due to maintenance being a major factor of railway assets life, the effect of maintenance on the track quality and further degradation has been thoroughly explored in literature. Profillidis (2006) describes how after maintenance (such as tamping), longitudinal defects progress rapidly up to around 2 Million Standard Axles (MSA) beyond which the rate decreases as the track has become more stable. This proves that after maintenance the track is not fully stabilised with it being estimated that it is only about 50% of when the track is fully stabilised. The ORE of the International Union of Railways (UIC) developed a semi-logarithmic empirical relation for mean track settlement related to cumulative axle load which took into account the initial fast rate settlement that then slows under a greater amount of loading operations. The mean track settlement, m_e, is

related to the cumulative traffic tonnage, T by (Profillidis, 2006):

$$m_e = a_1 + a_0 \log \frac{T}{T_r} \tag{2.10}$$

Where T_r equals 2×10^6 tonnes or 2 MGT, the point at which the rate of settlement slows. a_1 is the mean settlement for traffic load T_r, with values typically between 5-15 mm, and a_0 is the settlement increase rate in mm/decade which is mainly depended on subgrade quality and is typically 2-6 mm/decade. A more useful characteristic of longitudinal defects is the SD as this represents the inconsistent settlement which causes uneven vertical geometry. For this ORE developed the empirical relationship relating the SD of longitudinal defects, SD_{LD}, to the cumulative traffic tonnage, T by (Profillidis, 2006): .

$$SD_{LD} = c_1 + c_0 \log \frac{T}{T_r} \tag{2.11}$$

Where c_1 is the SD of longitudinal defects for a traffic load of T_r, with values varying between 1.0-1.35 mm. It is related to early settlement and quality of maintenace, increasing after each maintenance action due to how the ballast condition degrades with each tamping activity and with time as it becomes fouled. c_0 is the rate of increase of the SD of longitudinal defects, with values varying between 0.1-0.2 mm/decade. It can be noted that Equation 2.11 is in the same form as 2.10 but with differing variable values. Equation 2.11 can be rearranged to find a simple traffic interval between maintenance, T_{LD}^{lim}, using a SD limit, SD_{lim}, which is when maintenance occur. The amount of traffic in tons between maintenance can be calculated by:

$$T_{lim} = 2 \times 10^6 \times 10^{\left(\frac{SD_{LD}^{lim} - c_1}{c_0}\right)} \tag{2.12}$$

Velt's Track Quality Model

It is noted by Velt (2007) that good quality track degrades slower than poor quality track, so the stability is dependent on the condition. This can be mathematically expressed for the track quality, Q after a certain amount of time, t, by:

$$Q = Q_0 e^{bt} \tag{2.13}$$

Where Q_0 is the initial quality when $t = 0$ and b is the degradation rate. The use of an exponential relationship demonstrates slow early degradation which later speeds up which is the opposite of most other research undertaken and discussed in this section but is backed up by work undertaken by Quiroga and Schnieder (2012), which is discussed later.

Office of Research and Experiments' Track Quality Model

Shafahi and Hakhamameshi (2009) state a simple model for track degradation first developed by ORE in 1988 and then converted to a Combined Track Record (CTR) index resulting in:

$$E = 36.57 \times T^{-0.0418} \times P^{0.2955} \tag{2.14}$$

Where E is the track degradation index as a CTR, T is the total accumulated tonnage since the track was new in MGT and P is the design axle load in tonnes. CTR is a type of Track Quality Index (TQI) which defines the quality of the track as a function of four Track Geometry Parameters (TGP) which degrade including unevenness, twist, alignment and gauge. The CTR index value varies between 0-100 where 0-50 classes as a failed track and 80-100 is excellent. By using a TQI, the four aspects of geometry degradation of the track are all taken into consideration when describing the track quality but by doing this detail of which TGP is the main problem is unknown and hence the required type of maintenance is also unknown.

Sato's Track Damage Model

Sato (1995) developed an empirical track damage model using field data obtained on the Japanese railway network using track geometry recording vehicles. The model uses rail vibrations to predict the growth of irregularities and is given by:

$$S = 2.09 \times 10^{-3} \times T^{0.21} \times V^{0.98} \times M^{1.1} \times L^{0.21} \times P^{0.26} \tag{2.15}$$

Where S is the increase in track irregularities in mm/100 per day, T is the traversed tonnage in MGT/year, V is the mean velocity of the trains on the track in km/h, L is the influence of the rails where $L = 1$ for CWR and $L = 10$ for jointed rail. P is the influence of the subgrade where $P = 1$ for good soil and $P = 10$ for poor soil and M is the structural parameter given by:

$$M = P_b y_z S_i \tag{2.16}$$

Where P_b is the quasi static pressure in the ballast, y_z is the acceleration of the rail and S_i is an impact coefficient which is a function of the rail properties.

Wisc-Rail

Wisc-rail is a ballast degradation model that predicts how the track geometry degrades over time as a SD of the vertical geometry. The models focus is the effect of fouled ballast in varying moisture contents on the rate of track degradation. It was developed as part of a thesis completed by Ebrahimi (2011b) and was written in Matlab, with a

graphical interface which can be seen in Figure 2.20.

To establish the effect of differing ballast fouling conditions with varying moisture content, a Large-Scale Cyclic Triaxial (LSCT) test was used. The chosen stresses were equivalent to a 264 kN (30 tonne) axle load, and with the number of load repetitions, converted to MGT.

The effect of mineral, coal and clay fouling were then measured by adding the material prior to compaction and then increasing the moisture content by adding water to the top of the LSCT sample. The plastic strain was then constantly measured up to, $N = 2{\times}10^5$. From Figure 2.21 it can be seen that the semi-logarithmic rate of plastic strain ($r_p = d\varepsilon_p/ln(N)$) is constant through the initial compaction phase ($N < 1 \times 10^4$ or 0.3 MGT) and is linear during the fouling impact phase ($N > 1 \times 10^4$) for constant fouling and water content conditions (Ebrahimi, 2011b). This results in the relationships for the rate of plastic strain to be defined as:

$$r_p = \frac{d\epsilon_p}{d\ln N} = b \qquad N < 10^4 \ \ (0.3 \text{ MGT}) \tag{2.17a}$$

$$r_p = \frac{d\epsilon_p}{d\ln N} = b + a\log\left(N - 10^4\right) \qquad N > 10^4 \tag{2.17b}$$

Where b and a are related to other aspects such as axle loads, ballest fouling (as BFI) and moisture content.

Due to the constantly changing traffic loads, moisture condition and degree of fouling, to give a total plastic strain over a large time period Equation 2.18 is used.

$$\epsilon_p\left(N\right) = \sum_{i=1}^{N}\left(\int_{N_i}^{N_{i+1}}\left(\frac{d\varepsilon_p}{d\left(\ln N\right)}\right)_i d\left(\ln N\right)\right) \tag{2.18}$$

The deformation of the subgrade calculated as a strain, ϵ_p, is then calculated using Equation 2.19 which was developed by Li and Selig (1996), where the parameters d, c and m are related to the subgrade's classification by Table 2.3. σ_s is the soils static strength.

$$\epsilon_p\left(N\right) = c\left(\frac{\sigma_{ds}}{\sigma_s}\right)^m N^d \tag{2.19}$$

The change in the vertical geometry, measured as a SD can then be found using Equation 2.20 which was developed by Chrismer and Selig (1994). d_L is the total vertical deformation from plastic strain in the ballast and subgrade and δ_{v0} is the SD at the start of the time period. They equation presumes that a SD increase is equivalent to 15% of the total track settlement.

$$\delta_v = \delta_{v0} + 0.15 d_L \tag{2.20}$$

Table 2.3: Railway track subgrade deformation model parameters (Li and Selig, 1996)

	Subgrade Classification			
Model Parameters	ML	MH	CL	CH
d	0.1	0.13	0.16	0.18
c	0.64	0.84	1.1	1.2
m	1.7	2.0	2.0	2.4

The ballast deformation model that is at the centre of Wisc-Rail is a highly developed method, based on a series of well planned laboratory experiments that seem to closely match situations in the field. The model has not been checked against field data and may require alterations before it can be used. From Figure 2.20 it can be seen that the model requires a substantial amount of input data, some of which will be hard to obtain for an in-situ track such as the moisture content in each season and the unconfined subgrade strength. The models ability to estimate ballast deformation with the inclusion of water content and a fouling index is highly important as these have a great influence on the rate of plastic strain as seen in Figure 2.21.

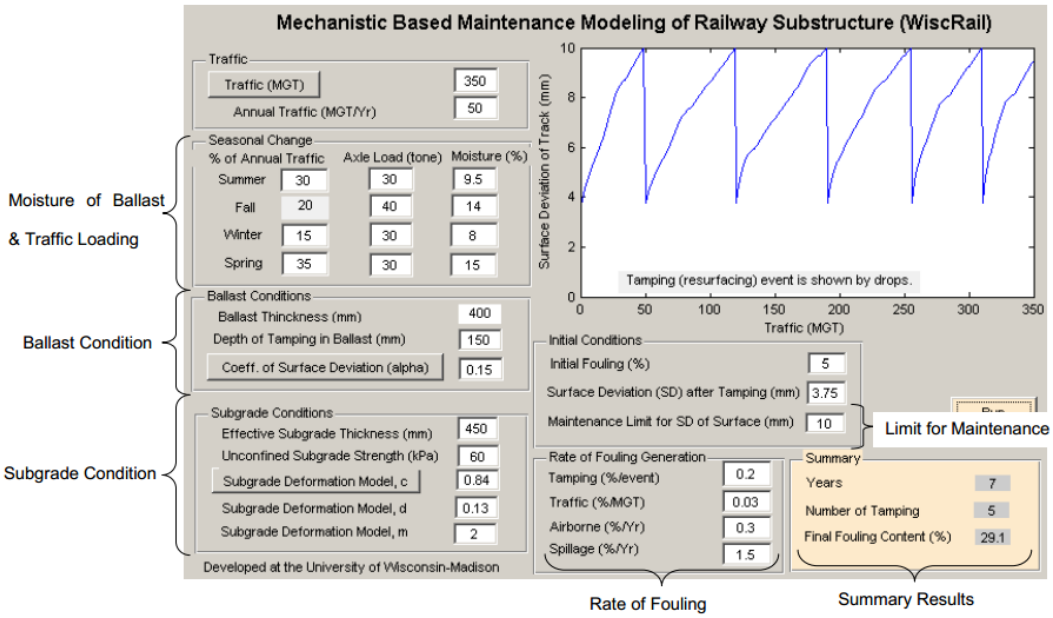


Figure 2.20: Wisc-Rail's graphical interface (Ebrahimi and Keene, 2011)

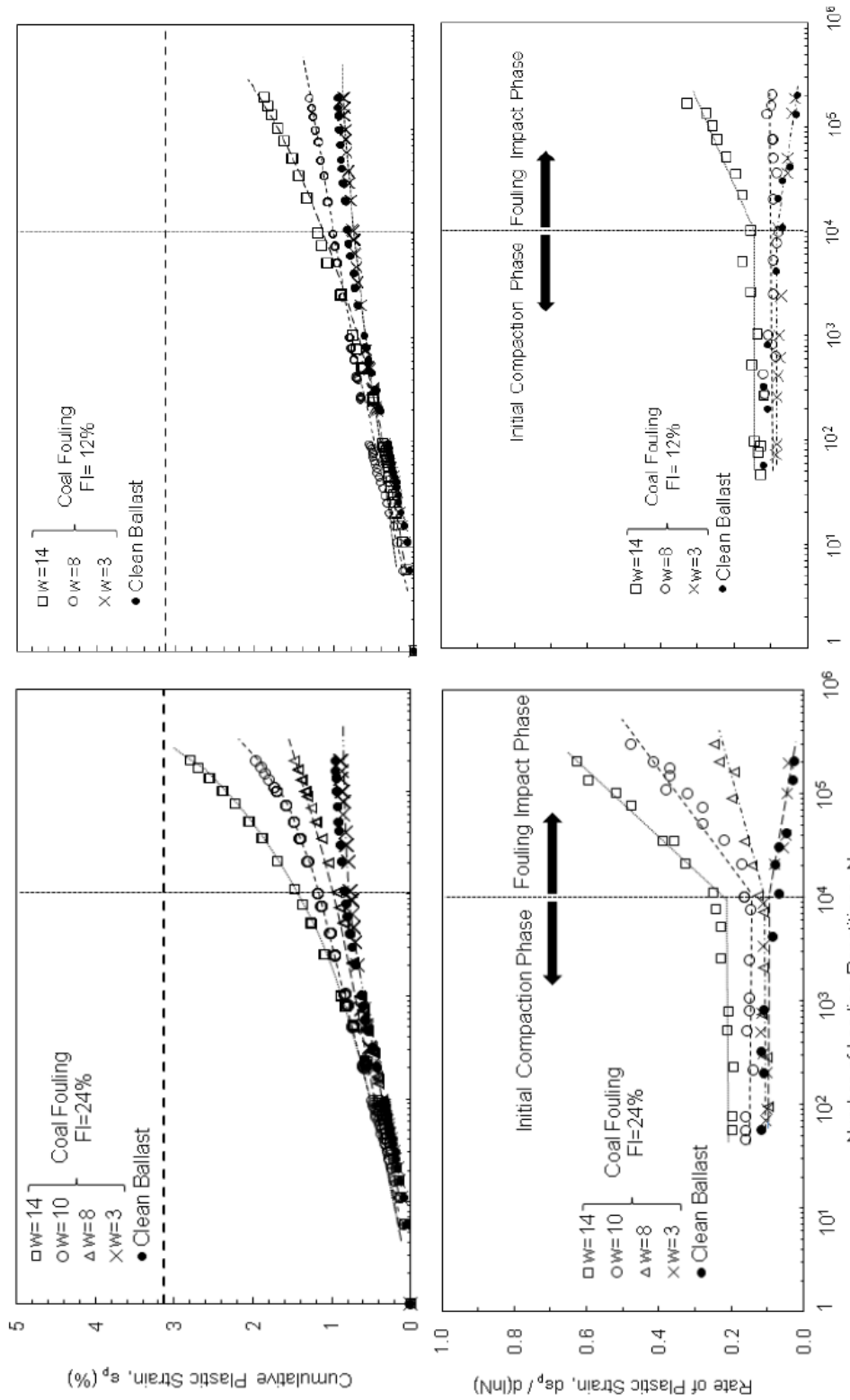


Figure 2.21: Effect of fouling and water content of plastic strain and rate of plastic strain (Ebrahimi, 2011b)

Multivariate Statistical Analysis Approach

Gular et al. (2011) introduced a method of predicting track geometry degradation, including twist, gauge, alignment, cant and level. This involved relating the degradation rates to: traffic loads, x_1; velocity [km/h], x_2; curvature $(1/R)$ [1/m], x_3; gradient [%], x_4; cant, x_5; sleeper-type, x_6; rail type [kg/m], x_7; rail length [m], x_8; falling rocks, x_9; land-slide, x_{10}; snow, x_{11} and flood, x_{12}.

Gular et al. used data from 180 km of track in Turkey which was then split into homogeneous sections, as demonstrated by Figure 2.22. This resulted in 820 segments, with an average length of 220 m. The data available included 7 years of track geometry recordings, with two undertaken each year.

To account for the unknown track quality when maintenance actions were performed, it was presumed that the track quality was at the maintenance threshold whenever a maintenance record existed, as seen in Figure 2.23. The resultant dataset had 820 rows, one for each section and 84 columns, one for each month of the 7 years of data. This was created for each of twist, gauge, alignment, cant and level.

From this a degradation rate for each track parameter is calculated by using Equation 2.21 to calculate the gradient of a linear line between the initial and final degradation points, as demonstrated in Figure 2.23. An average degradation rate was then found for each section of track.

$$\tau_{ij} = \frac{D_j - D_i}{Y_j - Y_i} \tag{2.21}$$

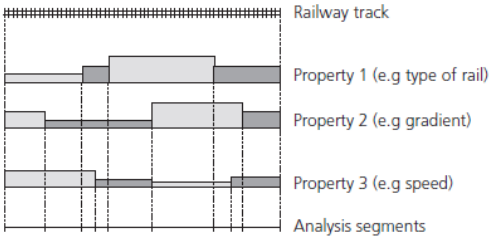


Figure 2.22: Track segmentation process (Gular et al., 2011)

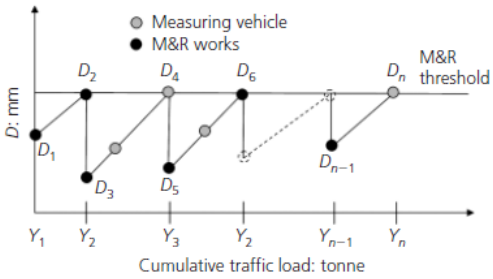


Figure 2.23: Determination of degradation rates (Gular et al., 2011)

To take into account uncertainty and the degree that the relating factors, x_1-x_{12} affect the degradation rate, in mm/tonne, Equation 2.22 was formalised.

$$\begin{aligned}\bar{\tau_i}=&\alpha_{i1}x_1 + \alpha_{i2}x_2 + \alpha_{i3}x_3 + \alpha_{i4}x_4 + \alpha_{i5}x_5 + \alpha_{i6}x_6 + \alpha_{i7}x_7 + \alpha_{i8}x_8 + \\ &\alpha_{i9}x_9 + \alpha_{i10}x_{10} + \alpha_{i11}x_{11} + \alpha_{i12}x_{12} + \alpha_{i0}\end{aligned} \tag{2.22}$$

To enable qualitative variables such as sleeper type to be included dummy variables are used, for example a timber sleeper has a value of one and a concrete sleeper has a value of zero.

A statistical process consisting of; correlations, hypothesis testing, stepwise analysis, case statistics and multicollinearity tests, gave values of how strong of a relationship there is between the variables or if it should not be included. This forms the values of α_{in}, which can be seen in Gular et al. (2011) for twist, gauge, alignment, cant and level.

The model is a good example of all the possible factors that effect the degradation. The resulting R^2 values for twist, gauge, alignment, cant and level were 0.624, 0.718, 0.694, 0.775 and 0.684. These are reasonable values but show that some errors will occur when using the model. There are also issues with the initial data as taking Maintenance and Renewal (M&R) data and presuming that it occurred at the point of the threshold value to increase the amount of data points is going to lead to much of the data being incorrect. This is because the M&R activities may have occurred late due to issues such as unavailable plant or early due to another track geometry measurement being at the threshold (tamping maintains many geometry issues at once). Another aspect that may lead to M&Rs occurring before the threshold would be opportunistic maintenance. Additionally, by taking a mean degradation value for each section, the increased rate of degradation seen after each maintenance activity, due to reductions in ballast quality, is not taken into account.

Despite some shortcomings within the model, the statistical mathematics used to develop the model are well utilised for the situation enabling the many factors that effect track degradation to be quantified. The method of splitting the track into homogeneous sections is also an effective way of allowing all the factors that have an effect on the degradation rate of the track geometry to be analysed while keeping to data clean.

ECOTRACK Geometry Degradation Rate Model

ECOTRACK is a decision support system that calculates the costs of different maintenance policies allowing the most efficient M&R activities to be undertaken. Models are included for all aspects of the track degradation with these models relying on a database that contains all the information about a railway network, including all previous M&R activities, as well as previous faults, track geometry recordings and even layout and operational data (Jovanovic and Pearce, 2000). The model for calculating track degradation is described below and the rail model is discussed in Section 2.4.4.

The ECOTRACK model calculates degradation rates from past track geometry recordings, including rates of vertical level, alignment, twist and cross level. The principle behind the ECOTRACK model can be seen in Figure 2.24, which demonstrates how geometry degradation occurs. If left unmaintained a track section would follow the initial dotted curve, which shows initially very fast degradation in the period marked a. This includes the bedding in period where large settlement occurs as the ballast stones

move as they become compacted. The initial fast period of degradation then slows and becomes linear as shown by area *b* in Figure 2.24. After this period the degradation rate tends to increase rapidly in the period marked *c*. Following the dashed line, the effect of maintenance can be seen. When maintenance is initially completed the track quality is improved but not returned to the initial new state. Faster degradation than when new then occurs which, with the reduced initial quality, results in the track reaching the maintenance limit quicker. After each maintenance activity the quality returned is lower and the rate of degradation is higher, hence the usage until maintenance is required is reduced. When a replacement takes place at 60 MGT, in Figure 2.24, the track is returned to a new state and the degradation process starts again, following the same as before due to being a direct replacement. When an upgrade is completed at 120 MGT, the improved track degrades slower, increase the usage until maintenance is required. The degradation is measured as a roughness increase, which is calculated using the equation on Figure 2.24. Due to the main part of the degradation which occurs up to the maintenance limit having a constant degradation rate, area *b* in Figure 2.24, the ECOTRACK model uses linear regression to calculate this part, ignoring the fast degradation which occurs in area *a*. The fast degradation that occurs in area *c* is also ignored with it presumed that maintenance will occur before the degradation rate will increase (Jovanovic and Pearce, 2000). This leads to the ECOTRACK model actually calculating the red saw-like shape in Figure 2.24. The linear regression is calculated through the past geometry recordings (ERRI, 1994). The ECOTRACK model is designed to estimate the quality improvement, drop in the red line, with occurs with maintenance as well as the change to the degradation rate after each maintenance activity.

The ECOTRACK model was a well-developed commercial product, which was successfully implemented in many European countries. Despite its simplification of the non-linear track geometry degradation to a linear one the model was proved in the field to be useful at predicting the occurrences of M&Rs.

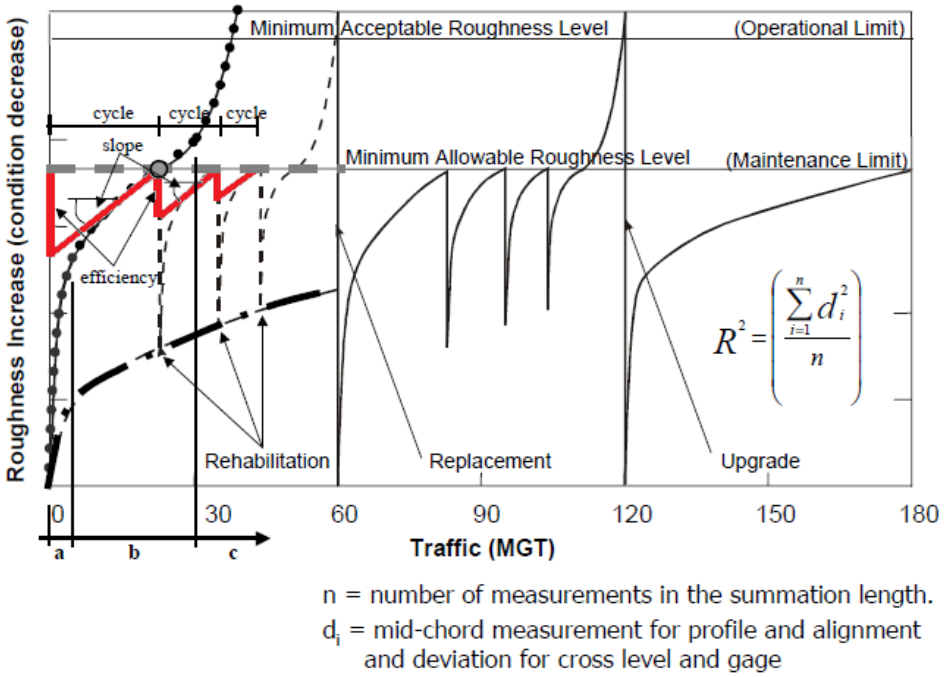


Figure 2.24: ECOTRACK geometry degradation analysis reasoning (Ebersohn and Selig, 1994)

Network Rail's Geometry Degradation Model

As part of NR's extensive selection of LCC models there are multiple tier 2 LCC models which include many designed for track. The tier 2 models look at a single asset and are used to examine the range of possible renewals, maintenance and utilisation options to establish an optimal policy. The model consists of elements for modelling the degradation of many parts of the track including ballast, sleepers and rail defects. The ballast degradation is discussed here, the geometry fault model is considered in Section 2.4.2 and the rail defect model is described in Section 2.4.4 (Halcrow, 2012).

The geometry and ballast model uses the 35 m vertical top geometry SD as the primary condition measurement. The geometry at a particular time, G_t, is related to the time by (Network Rail, 2012):

$$G_t = K\ e^{at^b} \tag{2.23}$$

Where K, a and b are the track section specific parameters, where $K \leq 0$. These values are based on empirical knowledge and are obtained by fitting the equation to SD values obtained from NR's tier 1 Track Strategic Planning Application (T-SPA) model which is part of their Vehicle Track Interaction Strategic Model (VTISM). The T-SPA model requires inputs such as; current infrastructure and conditions, traffic data, static and dynamic forces, RCF and unit costs of work, then allows the user to set criteria such as maintenance, renewal and inspection rules. With this data the model is able to output data such as the volumes of track maintenance and renewals, as well as the associated

costs over a certain time span (Williams, 2012).

The T-SPA model is a complex model of nearly 40 equations and many parameters to estimate the track SD at a particular time. Due to this the model is instead used to find just 3 points at 0.5, 1.0 and 2.0 years.

The T-SPA model calculates the damage accumulation by using traffic forces, which depend primarily on the number of axles, axle loads and train speed as well as track properties such as sleeper spacing and support stiffness. The impact of geometry fault (related to the total unsprung mass), dip joints (track irregularities at joints and welds) and total ride force are considered.

The three contributing factors are then combined by:

$$SD_T = \sqrt{SD_{DJ}^2 + SD_{UM}^2 + SD_{RF}^2} \tag{2.24}$$

Where SD_T is the total SD in mm at the calculated time, SD_{DJ} is the SD caused from the dip joints, SD_{UM} is the geometry fault SD and SD_{RF} is the SD caused by the total ride force.

Values for the SD are calculated using the T-SPA model for 0.5, 1.0 and 2.0 years and the Equation 2.23 fitted using a computer program. This then gives the values of K, a and b for that particular track section.

Equation 2.23 gives the expected SD at any given time, for clean ballast. To account for ballast fouling and its effect on geometry degradation a Ballast Condition Factor (BCF) is used. This is related to the BFI by:

If BFI=0% then:

$$\text{BCF} = 1 \tag{2.25a}$$

If 0% < BFI < 80% then:

$$\text{BCF} = 1 + (0.625 \times BFI) \tag{2.25b}$$

If 80% ≤ BFI < 100% then:

$$\text{BCF} = 1.447 + (4.668 \times 10^{-6} e^{11.638 \times BFI}) \tag{2.25c}$$

If BFI = 100% then:

$$\text{BCF} = 2 \tag{2.25d}$$

With the BFI being estimated by:

$$\text{BFI} = \frac{EF + (TF \times \text{MGT}_{All}) + (DF \times \text{MGT}_{Dirty}) + (MF \times Tamps)}{VA} \tag{2.26}$$

Where the parameters have been defined in Table 2.4 (Williams, 2013).

As each section of track degrades differently, even if they are seemingly identical, and because the model does not take into account all the factors that effect track geometry degradation, a Local Track Section Factor (LTSF), is applied. This is calculated by using least squares regression to map the line found from Equation 2.23 multiplied by the BCF on to past SD recordings. The SD of the track section at any time can then be found using Equation 2.27.

$$SD_{VerticalTop} = K(\text{BFC})(\text{LTSF})\ e^{at^b} \tag{2.27}$$

Table 2.4: Parameters used to calculate Ballast Fouling Index (Williams, 2013)

Parameter	Description	Default Value
VA	Available ballast void after renewal	22.95%
EF	Environmental fines	0.209% a year
TF	Traffic fines	0.0214% per MGT
MF	Maintenance (tamping) fines	0.578% per tamp
DF	Dirty wagon fines	0.18% per MGT

Adaptive Networks and Fuzzy Logic

Fuzzy logic is used to overcome the boolean results given by classical systems based around binary logic which follows the principle where a state can be either true or false. By applying fuzzy sets as a method of representing and manipulating data that is not precise, this allows the system to take uncertainty into account. It also allows recordable data, such as geometric quality of the track, to be combined with subjective information, such as visual inspections. Fuzzy logic uses human decision-making propositions like; $IF\ x\ is\ s$; $THEN\ y\ is\ t$, where x and y are linguistic variables and s and t are possible values. To achieve fuzzy sets, Membership Functions (MFs) are used to replace the standard indicator function used for classical systems. The MF can be any shape and represents the degree of truth, which can vary between 1 and 0, whereas classical systems which use binary logic has to be either 0 or 1 (Rojas, 1996; Dernoncourt, 2013).

Adaptive networks are networks of nodes and links which combine dynamics on the network as well as dynamics of the network. This produces a network where the links between the states adapt with respect to its states, enabling the network topology to dynamically change (Gross and Sayama, 2009). There are multiple types of adaptive networks including Artificial Neural Network (ANN) and Adaptive Neural-based Fuzzy Inference System (ANFIS) models, where an ANFIS is a mixture of an ANN and a Fuzzy Interface System (FIS). Kalogirou (2006) summarises the advantages of adaptive

networks as requiring less data than stochastic modelling with improved accuracy but states that they have limited applicability and are non-interpretive.

Type 2 Fuzzy Reasoning Model

Dell'orco et al. (2002) proposes a FIS that uses type 2 fuzzy reasoning and is able to logically make a decision on how long it is possible to delay maintenance based on a number of inputs. A simple type 2 fuzzy reasoning inference system with two inputs, x and y, one output and two fuzzy rules can be seen in Figure 2.25.

Standard data obtainable for a railway track was suggested consisting of; alignment, longitudinal level, cross-level, gauge, rail wear and rail corrugations. Instead of normal maintenance thresholds, a MF is used. Inspection data such as geometry SD is converted to fuzzy logic using MFs, linked to the maintenance MF and then though defuzzification, an estimate of time until maintenance is required is obtained.

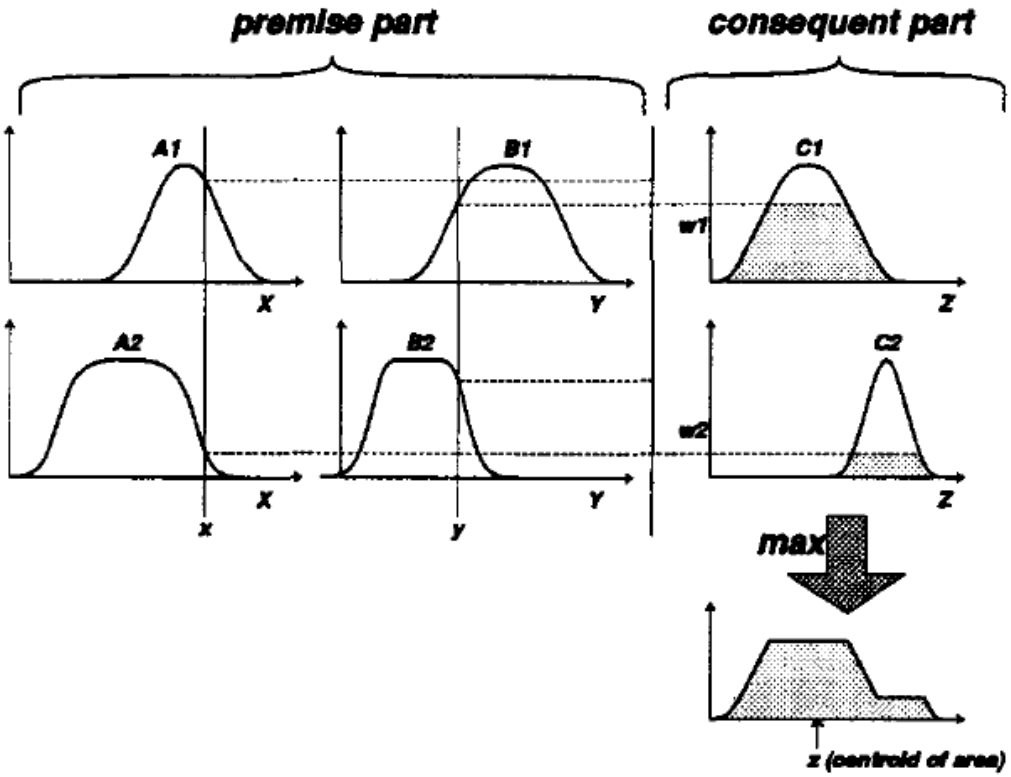


Figure 2.25: Type 2 fuzzy reasoning (Jang, 1993)

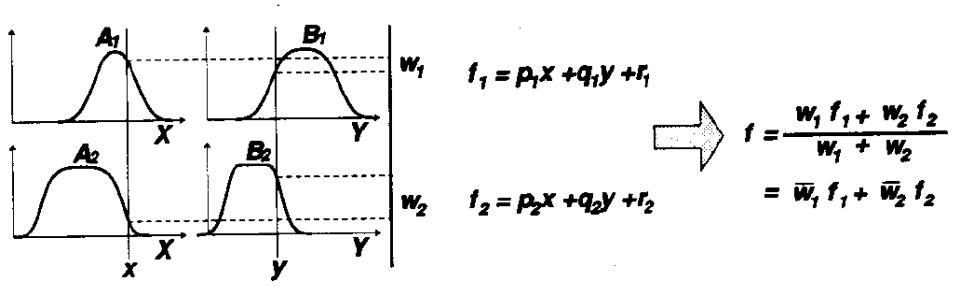
Ottomanelli et al.'s Maintenance Occurrence Model

Ottomanelli et al. (2005) expands on the work carried out by Dell'orco et al. (2002) by using a simplified version of an ANFIS with three inputs and one output for 200m long pieces of track. An ANFIS uses type 3 fuzzy reasoning, which is demonstrated in Figure 2.26a with two inputs x and y and one output f. The ANFIS architecture is a neural network that integrates fuzzy logic principles into a single framework (Jang

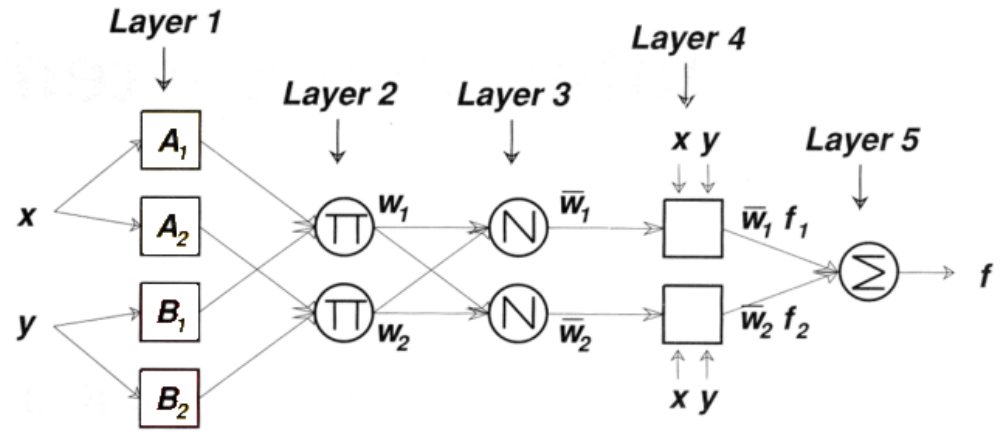
et al., 1997). It uses IF-THEN fuzzy rules that have learning capabilities to approximate non-linear functions from given data (Jang, 1993). The equivalent ANFIS to the FIS in Figure 2.26a can be seen as Figure 2.26b.

The three inputs are: the difference between the track alignment SD and vertical SD to a given threshold, with negative values meaning that the threshold has been passed. The third input is the time period since tamping was last completed. The output given is an estimate of the number of days until tamping works will have to be undertaken.

The robustness of the system is proved by Ottomanelli et al. by adding in an error for the input values ranging from the most recent (last input) up to the 5^{th} to last from this it was possible to see that the output error caused from a wrong input was minimal varying between 0.5-7.29% depending on when the error occurred and the degree of the error.



(a) Type-3 Fuzzy Interface System



(b) Equivalent type-3 Adaptive Neural-based Fuzzy Inference System

Figure 2.26: Simple example of an Adaptive Neural-based Fuzzy Inference System (Jang, 1993)

Dell'orco et al.'s Maintenance Occurrence Model

Another method of predicting tamping works was proposed by Dell'orco et al. (2008) by using a complete ANFIS network. The SD of the longitudinal level, alignment, cross level, number of tamping works previously carried out and number of days since last tamping operation at the date of measurements were used as the inputs. The output

given is the number of days since the last tamping operation that the next one will be required.

The ANFIS is then used to find the predicted date of the next tamping operation for each track segment. It is then possible to group close segments that require tamping at similar predicted dates to be completed at the same time by using the Fuzzy C-Means clustering method. This is able to group the interventions in location and time. It can also take into account possible working dates set by the railway organisation and time limits for maintenance.

The ANFIS was tested by using a 8.4 km piece of track, split into fourteen 600m long segments, with data being supplied by the Italian railways. The data consisted of 150 measurements for each segment, with Dell'orco et al. noting that after each successful tamping operation the returned SD values for the track were lower than the tamping operation carried out before, which is expected. The error obtained during the training phase of the ANFIS reached a constant value of 37 days.

Shafahi et al.'s Artificial Neural Network and Adaptive Neural-based Fuzzy Inference System Models

Another expert system has been proposed by Shafahi et al. (2008) using an ANN. This consists of a number of simple processes, which are known as the neurons which communicate with each other via weighted links. An ANN network consists of a number of inputs and outputs with hidden layers in-between consisting of any number of neurons, which are connected by the weighted links. An ANN involves back propagation using known inputs and outputs, from which the network learns, optimising the weighted links (Gershenson, 2003).

A cycle length of one year and a CTR index were used and track was split into blocks of a chosen length (Shafahi et al., 2008). The system of classification is the same as done by Shafahi and Hakhamameshi (2009), which is discussed in Section 2.4.2. Six inputs where used consisting of:

1. CTR index for the year and previous four years classified into 5 states.
2. Traffic volume or load classified into light and heavy.
3. Maximum allowable speed of the line sorted into five classes.
4. Geographical location classified into plain, hilly and mountainous, a class value is then given combining this and traffic volume.
5. Maximum gradient of the block classified into five classes.
6. Minimum radius of curves in the block

The output desired was the CTR state, which from the data, it was known that each year the track either stayed in the same state or moved to the state below. The same data was used as by Shafahi and Hakhamameshi (2009), who's Markov chain method is described in Section 2.4.2. After the training period the ANN correctly predicted the

next state for 67% of track sections.

Converting the ANN model into an ANFIS model improved the results by 6% leading to 73% of the estimations being of the correct state. A table comparing the linear regression results obtained from the expert systems developed by Shafahi et al. (2008), the Markov Chain method by Shafahi and Hakhamameshi (2009) and the ORE equation, as seen from Equation 2.14 can be seen in Table 2.5. Where the linear regression is obtained by plotting the estimated answers against the observed.

Table 2.5: Comparison of four different degradation models using Combined Track Record indexes

	Observation-Estimation Relationship		
Model	a	b	R^2
ORE	0.1188	72.823	0.1190
Markov Chain	0.7805	19.871	0.8317
Neural Network	1.0352	0.0171	0.7243
Neuro-fuzzy Network	0.8749	0.6443	0.8096

Stochastic Models

Quiroga and Schnieder's Track Quality Model

Quiroga and Schnieder (2012) introduce a model to predict track quality which takes into account tamping operations and how degradation increases after each tamping operation. The model links a track quality measurement, Q, with time, t, by:

$$Q = a\ e^{b(t-t_0)} + \varepsilon(t) \tag{2.28}$$

Where a and b are log-normally distributed stochastic variables and ε is normally distributed with a mean value of zero. t_0 is the time of the last tamping operation. Due to the parameters values being distributed once these have been defined, a Monte Carlo simulation can be run and the results used to optimise the occurrences of tamping. The use of distributions for the parameters, instead of the strict empirical relationships at the heart of the models in Section 2.4.2, results in a spread of answers for the rate of degradation. This spread allows for confidence in the results to be quantified. The ε parameter, adds variability to the results and the first part of Equation 2.28 follows a similar exponential format to Velt's track quality model in Section 2.4.2.

Andrade and Teixeira's Track Geometry Degradation Model

A similar style stochastic model to the one proposed by Quiroga and Schnieder (2012) is presented by Andrade and Teixeira (2011). Vertical top geometry recordings expressed as

a SD were obtained from the Portuguese railway. The recordings demonstrating a linear relationship between the SD and usage expressed as MGT. Andrade and Teixeira (2011) found disparity in the degradation rates of the sections that were virtually identical, with the same assets, speed and usage. This gives a good indication that a stochastic modelling technique is more suitable for modelling the railway track as it allows for this disparity to be accounted for. The linear relationship used consisted of:

$$\sigma = c_1 + c_0 T \tag{2.29}$$

Where, T is the accumulated tonnage, σ is the resulting track SD, c_1 is the initial track SD and c_0 is the rate of degradation. To account for the disparity in degradation rates, distributions were fitted to the parameters c_1 and c_0, where log-normal distributions was shown to best fit the data. To increase the accuracy of the distributions the track sections were grouped by what they contained, in this case; bridges, PL, stations and switches. The model can then be solved using Monte-Carlo sampling, with the results giving the possible spread of the track condition.

Caetano and Teixeira's Track Geometry Degradation Model

As with the degradation model described by Andrade and Teixeira (2011), Caetano and Teixeira (2016) employs a linear degradation relationship between the track geometry and usage in MGT with a stochastic variable solved by Monte Carlo. Unlike Andrade and Teixeira (2011) who placed variability on the initial quality and degradation rate with log-normal distributions, to achieve the variability in the degradation rate Caetano and Teixeira (2016) multiplies the current track linear degradation rate, r_i, by a stochastic variable, δ_i:

$$SD = SD^0 + Z_i l \tag{2.30}$$

Where SD^0 is the current track geometry condition and l is the future usage in MGT. r_i is found for each track section by fitting a linear line to the past geometry records, where i is the amount of recordings available for the fit. The stochastic variable δ_i is related to i so that for lower values of i, where there is less confidence in the linear fit and the value of r_i, the spread is increased. How the degree of errors change as more geometry recordings were used to predict the degradation is demonstrated in Figure 2.27, where it can be seen that as i increases the errors reduce, initially quickly until $i = 6$ where it starts to slow. It was found that the logistic distribution was the best fit for the errors. A track sections future geometry as a SD can then be solved by Monte Carlo.

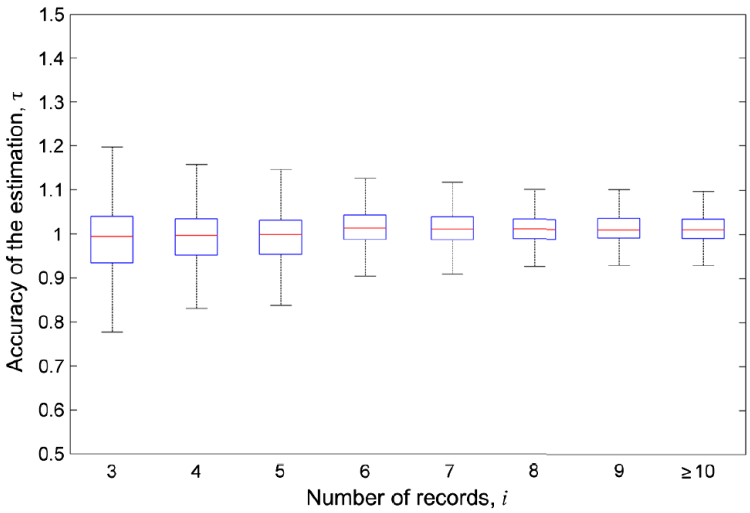


Figure 2.27: Accuracy of the predicted value of the Standard Deviation of the vertical top geometry (Caetano and Teixeira, 2016)

Markov Models

Introduction to Markov Models

A Markov model is a random process which is memoryless, meaning that only the current state of the process influences where it will go next. The reason for this is that Markov is based on the exponential distribution, which has a constant hazard rate. Another assumption of Markov is that the probabilities are constant. Markov models have a finite amount of states, i.e. degradation states. Markov chains are discrete time models, where after a set amount of time, i.e. an hour, week or year, there is a probability of which state will be moved to the next Markov processes are continuous time, where the time to move to another state is governed by an exponential distribution with the density function of:

$$f_T(t) = \lambda e^{-\lambda t} \qquad \text{for } t \geq 0 \tag{2.31}$$

Which state a Markov process will move to next is found by calculating the waiting time, T, for each possible next state, choosing the smallest, it will then move after that amount of time to that state (Norris, 1998).

Shafahi and Hakhamameshi's Markov Chain Degradation Model

Shafahi and Hakhamameshi (2009) outlines a method using Markov chains for determining the degradation of track condition. To split the track condition into different states a Track Quality Index (TQI) is used. The TQI used is a Combined Track Record (CTR) which can vary between 0-100 which was split into 5 states; failed, medium, good, very good and excellent. Due to the data used only showing a maximum decrease of 10 CTR in one year it is not possible for the track state to make more than one transition in a year

i.e. from 5 to 3. As there are 5 states and it is only possible to move one transition in a year or stay in the same state and no works are completed to improve the degradation state, the Markov Chain seen in Figure 2.28 is produced. The accompanying transition matrix for this Markov Chain is:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q_2 & p_2 & 0 & 0 & 0 \\ 0 & q_3 & p_3 & 0 & 0 \\ 0 & 0 & q_4 & p_4 & 0 \\ 0 & 0 & 0 & q_5 & p_5 \end{bmatrix} \tag{2.32}$$

In 2.32, p_n equates to the probability of staying in the same state, with the top left equalling one as once the track state 1 has been reached it cannot degrade further. The q_n values are the probability of degrading to a lower track state in the given time period of a year. The probability of either staying in the same state or degrading to a lower state must equal 1, therefore $q_n = 1 - p_n$. The state of the track is expressed as a row vector $p(n)$ which for a track that is in an initial state with excellent condition would be; $p_0 = [0 \quad 0 \quad 0 \quad 0 \quad 1]$. As part of a Markov chain the condition after the n^{th} year can be found by:

$$p_{m+n} = p_m P^n \tag{2.33}$$

Where m is the age of the track at the start of the chain and n is number of years that are calculated.

Data was then taken from the Iranian railway. This consisted of topographical, annual traffic and axle load, date of construction or reconstruction and track condition data for each block of track taken for 3 years. The traffic amounts were used to split the blocks into light and heavy traffic and the topography conditions into plain, hilly and mountainous. An average CTR index was then found for each track age for each group. This average CTR was then used to place the track in a track state, (Shafahi and Hakhamameshi, 2009).

The sum of the square differences between the expected state that was predicted from the Markov chain, Equation 2.33, and the track condition at each age is used as the objective function to estimate the transition matrix. The objective function was then minimized by the quasi-Newton method. This produced a transition matrix for each track class, with the p_n values found seen in Table 2.6. From the table it is possible to see that with a track quality of excellent the probability of the block degrading is low, but as the track quality worsens the probability increases, as seen by the lower p_2 values. Shafahi and Hakhamameshi used linear regression between the observed, x, and expected, y results to see the strength of the model. An R^2 value of 0.832 was achieved. The results show that there is some strength to the model to predict track deformation but also demonstrates that there is a degree of error.

Despite the ability of the model to predict degradation, the use of a TQI reduces the result's usefulness as it is unknown what geometry aspect needs to be maintained and hence what maintenance action to undertake is also unknown. Additionally, by using a Markov chain important aspects of a complete model would be hard to implement such as inspections to discover the current system state. This is due to the discrete time-step used within the process, as an inspection maybe every 6 months but the time step would be much smaller than this.

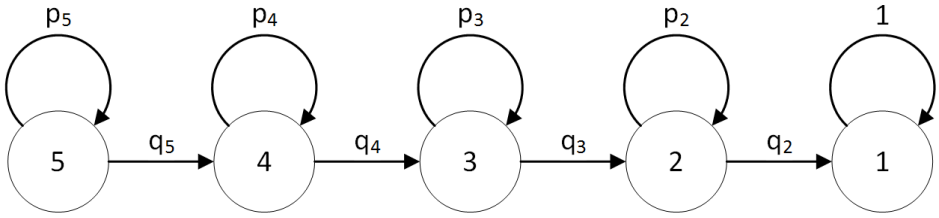


Figure 2.28: Markov chain degradation model

Table 2.6: Elements for the transition matrix for each track class (Shafahi and Hakhamameshi, 2009)

Track Class (K)	Terrain	Traffic	p_2	p_3	p_4	p_5
1	Plain	Light	0.3957	0.6104	0.7565	0.8641
2	Hilly	Light	0.1264	0.6497	0.7365	0.8151
3	Mountainous	Light	0.1931	0.5247	0.7310	0.8732
4	Plain	Heavy	0.3404	0.5314	0.6753	0.8390
5	Hilly	Heavy	0.3296	0.5546	0.7085	0.8067
6	Mountainous	Heavy	0.1633	0.4897	0.6593	0.7417

Prescott and Andrews's Markov Process Model

A Markov process model is proposed by Prescott and Andrews (2015) and includes degradation, inspections, maintenance and renewal of a $1/8^{th}$ of a mile track section. The model uses vertical top SD as the measure of quality with the degradation related to time. The track quality is split into four states; good, critical, speed restriction and line closure, with exponential rates, λ, to calculate the speed of degradation. States are related to particular vertical top SD values and the exponential rates are dependent on the amount of previous tamping operations. The Markov chain model is shown in Figure 2.29. Within the model the track section degrades through states 0, 1, 2 and 3, with these being the actual state of degradation, A, which is unknown until an inspection occurs, hence in states 0-3 the known condition, K, is still good. Inspections occur at set time intervals and are represented by the dotted lines on Figure 2.29. When an inspection occurs the state moves so the known state is the same as the actual state, e.g.

if the track section is in state 5, where the track condition is in a speed restriction but the last known condition was that it was critical. When the inspection occurs the track condition becomes known, so the state moves to state 7, where the actual and known conditions are both speed restriction. Now that an inspection has identified a section requires a speed restriction one can be put in place.

Maintenance is scheduled when the known state is critical, speed restriction or line closure. There are two types of maintenance scheduling embedded within the model. Normal tamping maintenance is scheduled when the condition of the track is known to be critical, and urgent tamping is scheduled when the track is known to have a speed restriction or line closure. The rate of degradation increasing after each tamping activity, resulting in smaller times to reach the chosen quality levels. This requires smaller exponential rates but due to the exponential rates within a Markov model being constant when tamping occurs the system cannot be returned to state 0 as from here the exponential rates are for track that has been renewed and had no tamping activities taken place. Due to this the model demonstrated in Figure 2.29 is copied for after each additional tamping activity with the related exponential rates. When maintenance occurs either by normal or urgent scheduling the system is moved to a good state in the associated part of the model for the current amount of previous tamping activities. This is seen in Figure 2.30, where states 0-9 are for a track which has just been renewed and 10-19 is for track that has been tamped once. This chain of small individual models for each amount of tamping activities goes up to seven previous tamps and hence there are 80 states within the whole model. After seven tamps, the rate of degradation is presumed not the change. Any additional tamps return the track to a good condition, but stays in the model part for seven tamps. Renewals are scheduled after a certain amount of time and return the state back to state 0.

The Markov model for the track section creates 80 different equations which are then solved numerically using a forth-order Runge-Kutta algorithm. This results in the probability of being in each state which can be summed to give the total probability of being in a good, critical, speed restriction or line closure state. These probabilities can then be multiplied by 365 to give the average amount of days a year that the track section is in each state. Due to the model using different states for known and unknown conditions it is possible to calculate the amount of time spent where the condition has changed state but has not been inspected yet, so it is unknown it changed. The results can also demonstrate the probabilities of being in each state after any amount of days since renewal so the results can show aspects such as how the chance of a speed restriction increases the longer it has been since a renewal. The amount of maintenance actions can also be found, with the probabilities of the amount that have been undertaken for any amount of days since renewal.

The model has multiple inputs which are related to asset policies, so by changing these inputs the effect of changes to policy can be seen in the outputs discussed above. These inputs consist of the track quality as a SD that maintenance is requested, the mean time

to scheduled and undertake normal maintenance, inspection interval and renewal period. Prescott and Andrews (2015) notes that the model is currently relatively simple but can return useful information that important decisions can be based on. It is also noted that increasing the aspects of the model to include aspects such as; multiple track sections, stoneblowing and other degradation aspects such as alignment and rail faults may lead to a state-space explosion. This is a common issue with Markov models, where the more complex the system gets the larger the Markov model becomes, which makes analysing difficult to perform.

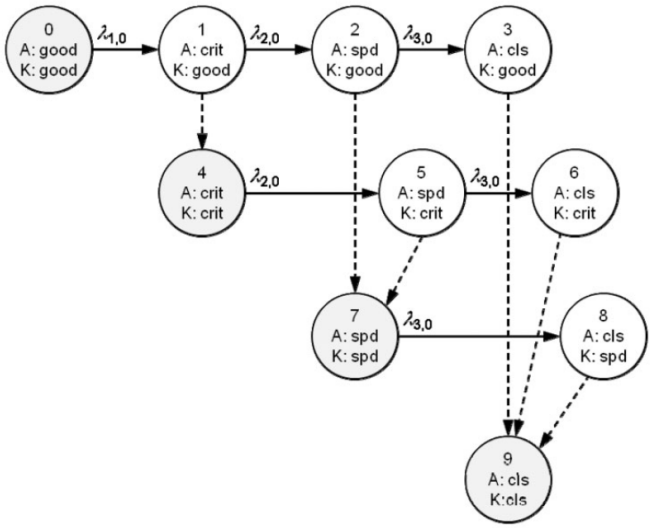


Figure 2.29: Continuous-time Markov chain model following renewal (Prescott and Andrews, 2015)

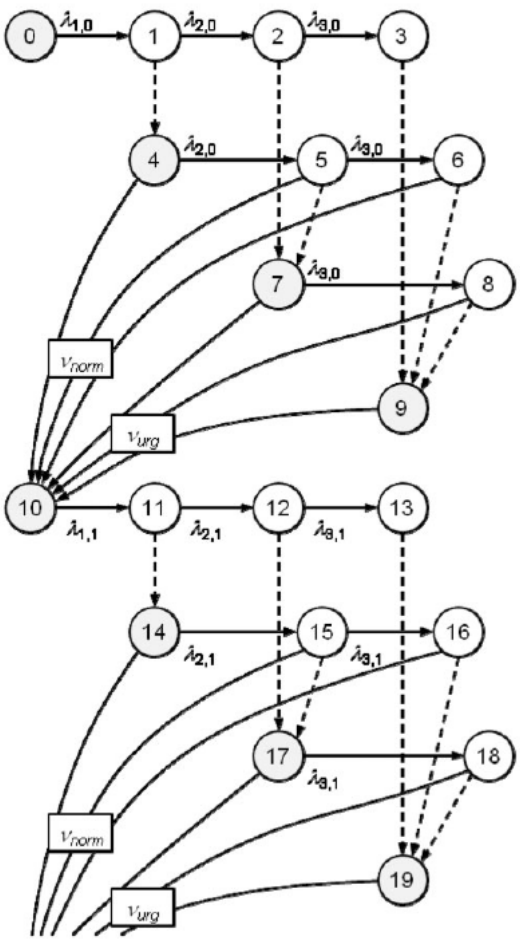


Figure 2.30: Continuous-time Markov chain model with maintenance (Prescott and Andrews, 2015)

Petri Net Modelling Approach

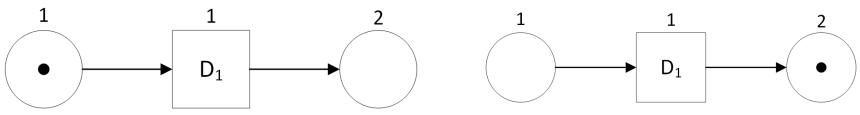
Introduction to Petri Nets

Petri Nets (PNs) are directed bipartite graphs that can be used as a pictorial representation of dynamic processes. PNs consist of two different types of nodes; places which are depicted by small circles and transitions depicted by small rectangles. The places can be either a state or condition of the system and can contain tokens, which are visualised by a dot. These tokens are moved around the PN, from place to place, by the transitions, which after a certain amount of time, fire the tokens from the input places into the output places, with the time for a transition to fire being a non-negative real number which is marked on the transition. Figure 2.31 demonstrates how the transition works, in the initial state the token is in place 1, p1, which activates transition 1, t1. After the time, D_1, the transition fires taking away a token from p1 and placing one in p2. If the transition time is zero, it is known as a immediate transition and occurs without delay. The places are connected to the transitions by edges which have arrows to give the direction of movement and can also contain a number depicting the amount of tokens required in the previous place until the transition is activated. For a transition to be activated

all the input places must contain the required amount of tokens depicted by the edges. Once the transition fires, all the tokens required to activate it are removed from the input places and tokens added to the output places, equalling the amounts depicted on the edges. An additional edge type that is used is the inhibitor edge which is depicted by a filled in circle instead of an arrow. The inhibitor edge stops a transition from activating while the leading place contains the required amount of tokens, (Schneesweiss, 2004). An example of numbered and inhibitor edges can be seen in Figure 2.32. First the PN starts in the initial state with the inhibitor restricting the activation of t1 and with t2 being activated, so after time, D_2, the transition t2 fires, so a token is removed from p3 and a token is added to p4. Now that the inhibitor place is empty and as there are the required two tokens in p1, t1 is activated. After time, D_1, two tokens are removed from p1 and one is added to p2. Inhibitors have many uses, allowing more logical operations as a transition can only fire if another place is empty, such as to ensure a maximum of one token is in a place (for example modelling a queue).

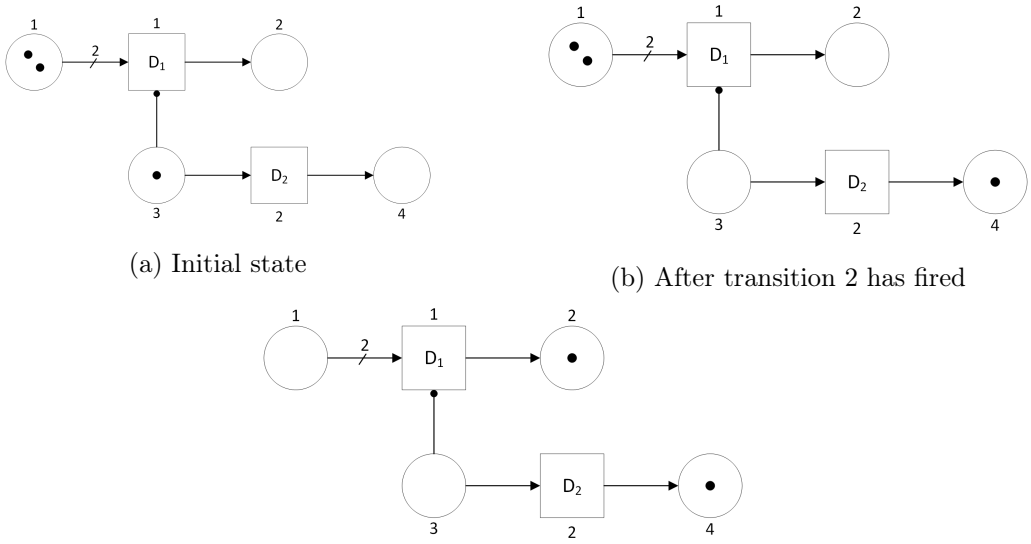
A more complete example of a PN is demonstrated in Figure 2.33. The PN starts in its initial state, in which only t3 is activated. After time, D_3, a token is removed from p3 and a token added to p2 and p4. Now only t1 is activated, leading to the removal of a token from p1 and p4 and the addition of a token in p1 and p2, after time, D_1. Now that there are two tokens in p2, t2 is activated. After time, D_2, two tokens are removed from p2 and a token is added to t3. This returns the PN back to its initial state and the loop keeps repeating.

PNs have the additional ability for the transition times to be any kind of distribution, which is especially useful for stochastic modelling. This freedom of distributions allows for any data to be represented. When distributions are used Monte Carlo sampling is employed to sample transition times from the distributions as they are activated. After running multiple simulations, with each time consisting of differing transition times due to the use of distributions, the running average of the required outputs from the PN, such as time spent in p1 or amount of time t1 fires, will start to convergence.



(a) Initial state (b) After transition 1 has fired

Figure 2.31: Typical example of a transition firing in a Petri Net

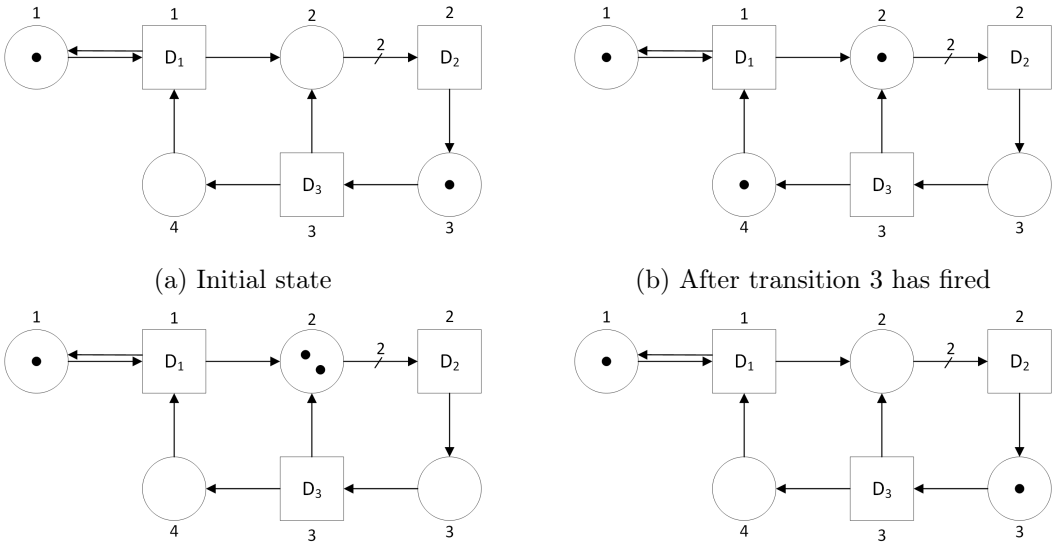


(a) Initial state

(b) After transition 2 has fired

(c) After transition 1 has fired

Figure 2.32: Simple example of an inhibitor transition within a Petri Net



(a) Initial state

(b) After transition 3 has fired

(c) After transition 1 has fired

(d) After transition 2 has fired

Figure 2.33: Simple example of a looping Petri Net

Andrews's Petri Net Asset Management Model

Andrews (2012) discusses the use of PNs to model track degradation, including inspection, interventions and renewal processes. A PN was used as they have many advantages over other modelling techniques such as Markov models, as they allow the use of variable degradation rates and the use of any distribution type for the transition times, which allows for a closer representation of how track geometry degrades. The distributions required for the degradation transitions were found by using multiple track geometry measurements as a SD for each track section to find linear degradation rates and hence the times to certain SD values. Distributions were then fitted to these times, where it

was decided that the Weibull distributions was the best fit, which is supported by Audley and Andrews (2013), where a comparison of the fits of multiple distributions was undertaken using the coefficient of determination, R^2, as the goodness of fit measure. Audley and Andrews found that the two and three parameter Weibull distributions returned the largest R^2 values of 0.986 and 0.987 respectively and due to the closeness of the fits the two parameter Weibull distribution was chosen over the three parameter for simplicity.

Conditional transitions were used to allow transition times to be related to previous activities, such that track degradation immediately after tamping is related to the previous amount of tamping operations undertaken. This is accomplished by changing the Weibull parameters, used for the degradation transitions, depending on the amount of previous tamping operations. By doing this the quicker degradation that occurs after each tamping operation, which is caused by the reduction in ballast quality, can be modelled within the PN. Convolution transitions were used for degradation of the track geometry, of which the vertical top 35 m SD was used, with four states consisting of; new, maintenance required, speed restriction required and line closure required. The convolution transitions are used as the Weibull distributions of times to the different states all originate from the same initial time, e.g. renewal to a needing maintenance state or renewal to a speed restriction required state. Figure 2.34 demonstrates how the blue parts, including times to $SD = \sigma 1$ and $\sigma 2$, are known, but the red section, $F_{\sigma 2|\sigma 1}(u)$, is unknown but required. To calculate the times between the degraded states, $F_{\sigma 2|\sigma 1}(u)$, Equation 2.34 is solved using the trapeze rule (Andrews, 2012).

$$f_{\sigma 2}(t)\,dt = \int_{u=0}^{t} f_{\sigma 2}(t)\,du \times f_{\sigma 2|\sigma 1}(t-u)dt \tag{2.34}$$

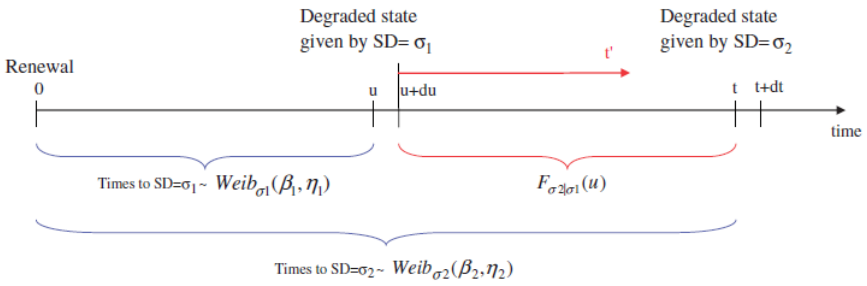


Figure 2.34: Degradation process - convolution transition (Andrews, 2012)

The PN is made up of four sections. The first is the degradation process, then there is the measurement train condition monitoring, maintenance and renewal processes. This leads to a complete track section PN containing 17 transitions and 13 places. The complete PN and a description of all the transitions and places can be seen in Andrews (2012).

To solve the PN Monte Carlo sampling is used and involves repeated random sampling of

a value, X, between zero and one, the same boundaries as the probabilities in a distribution's Cumulative Distribution Function (CDF). The CDF for the Weibull distribution is:

$$F(t) = 1 - exp\left(-\left(\frac{t}{\eta}\right)^{\beta}\right) = X \tag{2.35}$$

To obtain the sample time, t, the time until the transition fires, for each sampled value of X, Equation 2.35 can be rearranged to get:

$$t = \eta[-(\ln(X))]^{1/\beta} \tag{2.36}$$

The Monte Carlo simulation was run 65,000 times, with each time different random values of X being used until the running average of the number of interventions converged. Once this happens it is possible to obtain from the model information such as the minimum, maximum and average number of line closures and speed restrictions. It is also possible to find the total amount of days that speed restrictions were imposed and how many days maintenance was required. The model also gives the amount of tamping interventions and renewals that were undertaken for the given lifetime ran in the simulation.

The PN approach to modelling a railway track has many strong points. The ability to have changing distributions depending on previous works makes PNs much more adept for the track degradation process than other techniques such as Markov processes. PNs, unlike Markov processes, are not memoryless and so allow the degradation to be related to the track section's M&R history. The PN model proposed by Andrews (2012) contains the important aspects of track degradation, M&Rs and inspections allowing lots of useful information to be pulled from the model and policies, such as when M&Rs occur and inspections are undertaken, to be altered. The model can then be rerun with the new policies and the effect on performance, such as time spent with a speed restriction, compared with differing policies.

From research into other existing models it is possible to see that this PN approach does not take into consideration other inputs such as line speed and sleeper type, which other researchers, like Sato (1995), Gular et al. (2011) and Halcrow (2012), classified as important enablers of the track geometry degradation. This means that this PN approach is very general and hence may give results which are different to what is found in the field. The accuracy would be improved by calculating a greater amount of distributions for different inputs which categorise a particular track section in greater detail.

Network Rail's Geometry Faults Model

As part of NR's Track Tier 2 model, which was described in Section 2.4.2, there is also a method of estimating the amount of geometrical faults that will occur over a given

stretch of track in a set time period. This model is based on empirical relationships that have been derived from field data and relates the probability of faults, P_t, to the vertical track geometry as a SD in mm, G_t, by:

$$P_t = \ln(A + \alpha \log(G_t) + B + (\log(G_t))^2) \tag{2.37}$$

Where A, B and α are variables. A Poisson process simulation is then used to derive the specific fault occurrences and calculate the amount over a chosen length of time. The probability is taken to be the Poisson rate for a single time step, which is equal to one month. A CDF is then evaluated and compared with a random sample to determine its fault occurred in each time step (Halcrow, 2012).

2.4.3 Maintenance Effectiveness Models

Knowing the track geometry degradation is useful to predict the next time the thresholds are passed and maintenance is required. After the maintenance activity the track is not returned to an as good as new state, with each subsequent maintenance activity returning the track to a lower quality state, as shown by Figure 2.24 and discussed in Section 2.3.4. Due to this it is important to be able to predict accurately the quality of the track after a maintenance activity has been carried out.

Deterministic Models

Network Rail's Model

As part of NR's tier 2 track model the effectiveness of tamping and stoneblowing are calculated. The model uses the vertical top SD of the track before maintenance is undertaken, SD_{Pre}, to find the SD after maintenance, SD_{Post}, by using the linear function of:

$$SD_{Post} = A_G + (SD_{Pre} \times B_G) \tag{2.38}$$

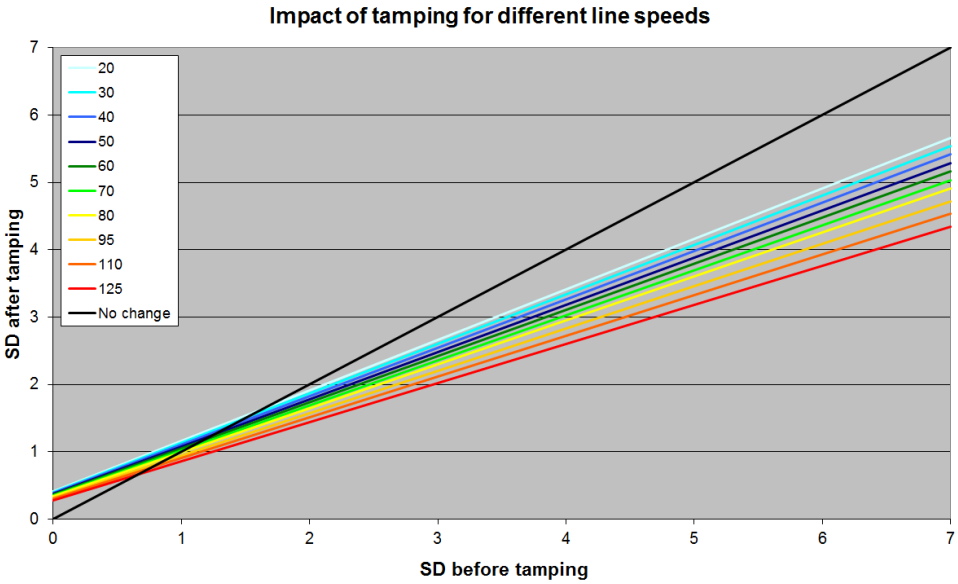
The values of A_G and B_G are related to the line speed and can be found in Table 2.7. The use of a linear relationship between the before and after tamping condition of the track is supported by M Miwa and Oyama (2000) and is used in their tamping restoration model. The values in Table 2.7 were found by using past data of SD measurements of within 30 days before and after a maintenance activity on different line speeds, (Halcrow, 2012). Looking at the values in the table it can be seen that stoneblowing is less effected by the initial condition of the track with lower B_G values than tamping. A comparison of the effect speed has can be seen in Figure 2.35.

From Figure 2.35 it is possible to see that both tamping and stoneblowing can actually

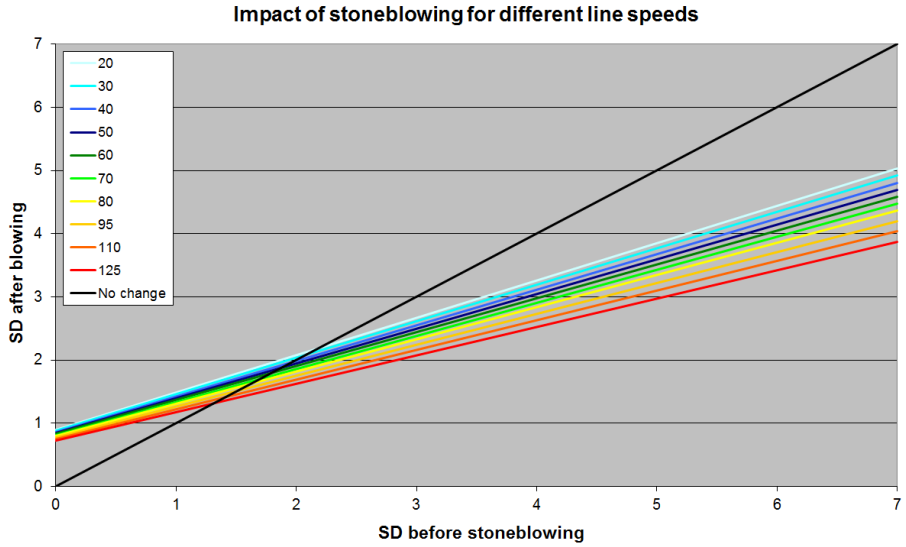
increase the track geometry SD if undertaken at an already low SD value (coloured lines are above the black line (no improvement)). It can also be seen that in general tamping is more effective on tracks with SD values between 1-4 mm whereas stoneblowing is more effective on tracks where the SD has already past 4 mm. Figure 2.35 also shows that the track speed alters the effectiveness of maintenance, with higher speed lines SD tending to be improved by maintenance more than slower lines (seen by the lower gradients). The reason why the line speed alters the effectiveness of maintenance is not known. It could be due to the degree of fouling in the ballast as slower lines tend to have poorer quality ballast due to maintenance, such as ballast renewal, not being undertaken as the permissible SD limits are much higher. An issue with the model may arise from the decision of using data from 30 days before and after the maintenance activities as this maybe ignoring some bedding in period. The data may also be skewed by the use of over lifting for some sections of the track to reduce ballast memory and may lead to a little or negative quality improvement in the short term. Once the track is fully bedded in the full improvement will be seen, but this may not occur in the 30 days after maintenance. It is also noted that the model does not take into consideration the amount of previous maintenance activities that have occurred or the ballast condition.

Table 2.7: Linear parameters for effectiveness of maintenance (Network Rail, 2012)

	Tamping		Stoneblowing	
Line Speed (mph)	A_G	B_G	A_G	B_G
20	0.365	0.754	0.880	0.577
30	0.360	0.742	0.864	0.571
40	0.356	0.730	0.847	0.566
50	0.351	0.718	0.831	0.560
60	0.347	0.706	0.814	0.554
70	0.342	0.693	0.798	0.548
80	0.338	0.681	0.782	0.543
95	0.331	0.663	0.757	0.534
110	0.325	0.645	0.732	0.525
125	0.318	0.626	0.708	0.517
140	0.311	0.608	0.683	0.508



(a) Effectiveness of tamping



(b) Effectiveness of stoneblowing

Figure 2.35: Effectiveness of maintenance on vertical top geometry Standard Deviation (Williams, 2012)

Caetano and Teixeira's tamping effectiveness relationship

As with the NR's maintenance model and other work by Stephen M Famurewa and Kumar (2015) which is discussed below, Caetano and Teixeira (2016) recommends a linear relationship between the initial quality before maintenance and the resultant quality. Using track recordings obtained from the LisbonPorto railway line in Portugal, parameters for the linear relationship have been found for track which was last renewed around 10 and 20 years ago. This results in two equations for the vertical top geometry, which are:

$$\Delta SD^{10} = -0.17 + 0.50SD_f \tag{2.39a}$$

$$\Delta SD^{20} = -0.13 + 0.36SD_f \tag{2.39b}$$

Where SD_f is the current track section geometry and ΔSD^{10} and ΔSD^{20} are the tamping improvement expressed as a SD for track 10 and 20 years old. Looking at the values in the equations it can be seen that tamping performs better on newer track, as seen by the linear gradient of 0.50 compared to 0.36.

Stephen M Famurewa and Kumar's tamping effectiveness relationship

Stephen M Famurewa and Kumar (2015) suggest another linear relationship between the initial vertical geometry SD and the resultant SD from tamping. To quantify the unknown linear parameters sections of track with good quality ballast were used with a linear relationship fit, as in Figure 2.36. The fit resulted in the linear equation:

$$R = 0.5445SD_0 - 0.8893 \tag{2.40}$$

Where R is the recovery of SD and SD_0 is the SD of the track section before tamping is undertaken. It can be seen from the values that Stephen M Famurewa and Kumar (2015) and Caetano and Teixeira (2016) agree about the degree the previous SD effects the quality improvement for newer good quality track, with Stephen M Famurewa and Kumar (2015) noting a value of 0.5445 and Caetano and Teixeira (2016) a value of 0.5. The two models differ though for the general improvement caused by tamping with values of -0.17 and -0.8893 with the value of -0.17 recommended by Caetano and Teixeira (2016) being more optimistic.

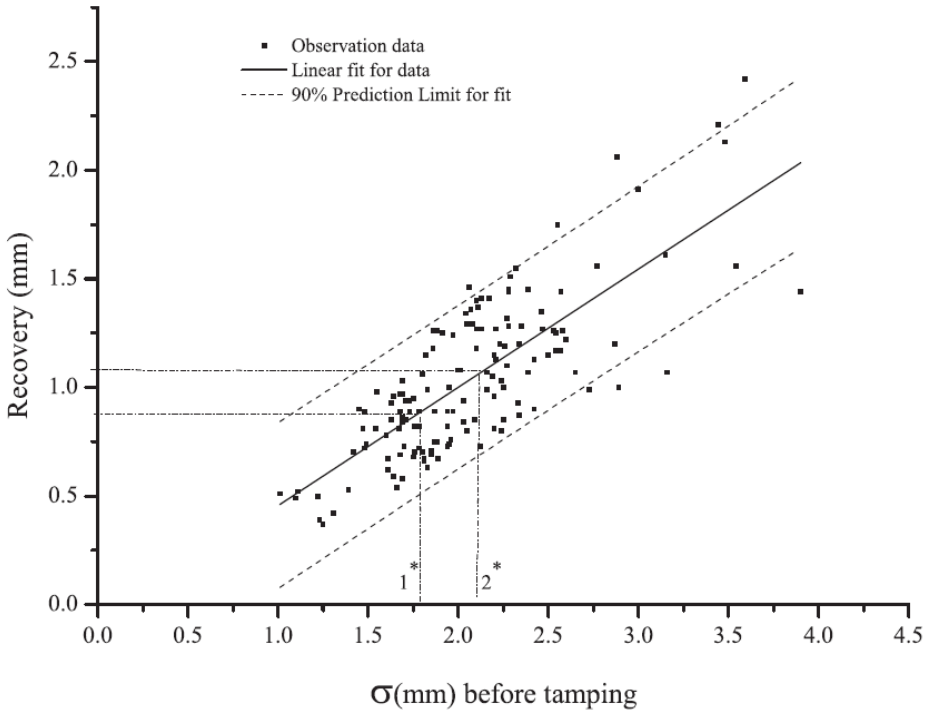


Figure 2.36: Effect track quality has on tamping effectiveness (Stephen M Famurewa and Kumar, 2015)

Velt's Maintenance effectiveness model

Velt (2007) notes that there is a diminishing return of quality of maintenance such as tamping as track ages. To model this relationship an exponential relationship is used to account for a slow early reduction in effectiveness which increase to a point where the maintenance action cannot return the quality back to a serviceable level, as seen in Figure 2.37. The exponential relationship follows:

$$Q = Q_0 e^{gt} \tag{2.41}$$

Where Q is the quality after works are completed, Q_0 is the initial quality, t is time and g is an unknown variable.

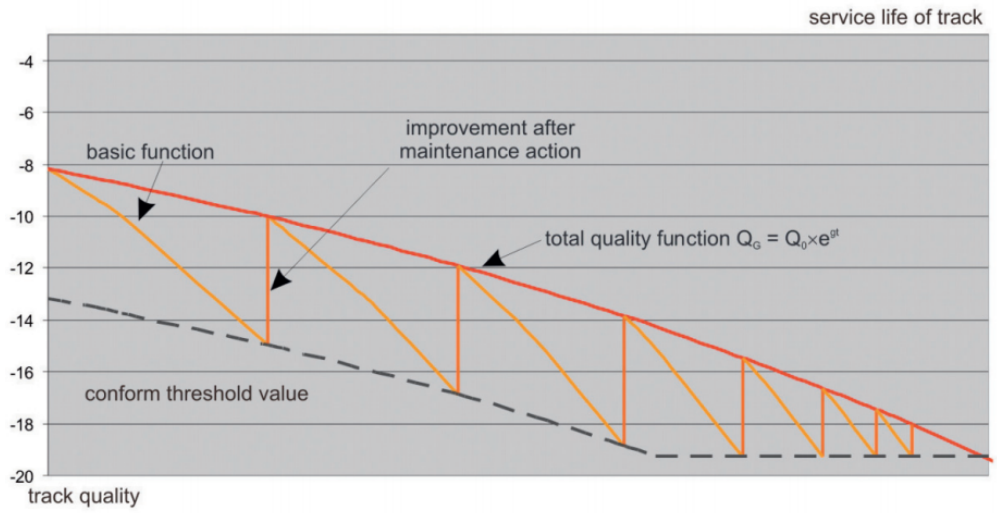


Figure 2.37: Exponential Maintenance Model (Velt, 2007)

Stochastic Models

Log-normal Distribution Model

Audley and Andrews (2013) discusses a method of predicting the likelihood of reaching a particular SD after tamping.

Past track geometry recordings from the UK were used to build up the probabilities of reaching different SD values after a full renewal or a certain amount of tamping operation up to two. Recordings of the minimum SD after the occurrence of tamping were used. Different distributions were then fitted, with the two parameteer log-normal shown to best represent the data, with an R^2 value of 0.995.The two parameter log-normal Probability Density Function (PDF) is defined as:

$$f(\sigma) = \frac{1}{\sigma\sqrt{2\pi\alpha^2}}\, exp(-\frac{(ln\ \sigma - \mu)^2}{2\alpha^2}) \tag{2.42}$$

Where α and μ are the parameters, with α giving the scale and μ giving the location. Values were found for after a renewal, one tamp and two tamps for four speed bands of 115-125, 80-110, 60-75 and 10-55 mph. These can then be used to obtain the probability of reaching any level of SD by using the log-normal CDF.

Like the NR maintenance effectiveness model, the model proposed by Audley and Andrews uses track speed as an effecting factor but unlike the NR model, Audley and Andrews does not consider the track condition prior to maintenance as a driving factor. As seen in Figure 2.35 the previous quality does have an effect on the change in SD achieved from tamping, due to this the log-normal distribution model maybe losing accuracy as it presumes that the SD level reached from tamping is only dependent of the amount of previous tamps and line speed. The use of the amount of previous tamping

operations is in line with other research which states that the achievable track quality after tamping is dependent on the BFI which increases with each tamping operation as the tamping tines cause ballast breakage.

2.4.4 Rail Degradation and Fault Models

The rail on the track is prone to a large amount of possible faults as described in Section 2.3.2. Rail faults can be extremely dangerous due to the weakness they cause in the running service of the track, with the possibility of causing derailments. These faults need to be predicted to allow future work and costs to be forecast.

Deterministic Models

Network Rail Model

Another method that is used by NR as part of their tier 1 model, known as VTISM, is outlined by Denby (2013) and Gordon (2006). The model is based off of relationships between the amount of actionable rail defects, consisting of either squats, tache ovals, bolt holes, welds or others and other aspects such as gross tonnage, track curvature, vertical geometry and rail type. The relationships were found using past data of rail defects obtained from the UK railway network. The model also takes into account the three different track types of CWR, jointed rail and S&C as well as stations and tunnels and can be seen as Equation 2.43 with the parameters being outlined in Table 2.8. Each type of rail defect of squats, tache ovals, bolt holes, welds or others for all three track types has its own values for the parameters.

$$\begin{aligned} D = d \times E^P \times max\left\{minG, a + b \times G\right\} \times (m + n \times min\left\{maxC, |C|\right\}) \\ \times r_{rail} \times s_{station} \times t_{tunnel} \end{aligned} \tag{2.43}$$

It is then possible to find the amount of serious rail defects and rail breaks per mile per year as a proportion of actionable defects.

The NR model that makes up part of VTISM is a very complete model and includes all the factors which are known to effect the amount of rail defects. Equation 2.43 has been developed over many years, using a large amount of field data obtained from the UK railways. Despite the amount of variables in Equation 2.43, only E, G and C are required for each track section with the rest of the variables having values related to the defect type and rail type, jointed, CWR or S&C, with these values being given in Denby (2013). By looking at these it can be seen that more defects occur in bullhead rails, followed by pre-1976 flatbottom rail, post-1976 flatbottom rail and then CEN-60. It can also be seen that for squats and tache ovales a tunnel section will experience more defects but a station would experience less. Weld defects are more common in both tunnels and stations but bolt holes are not effected by location.

Table 2.8: Network Rail's VTISM rail defect model parameters

Parameter	Description
D	Actionable defects per mile per year for the chosen defect group and track type
d	Defect constant
E	Cumulative Equivalent Million Gross Tonnage (CEMGT)
p	Exponent of E
G	Vertical short wave geometry SD [mm]
$minG$	Minimum vertical short wave geometry SD [mm]
a, b	Linear relationship for geometry and relative defect rate
C	Average rail curvature of the track section, 1/radius [m^{-1}]
$maxC$	Maximum curvature [m^{-1}]
m, n	Linear relationship for curvature and relative defect rate
r_{rail}	Multiplier for rail type
$s_{station}$	Multiplier if in a station
t_{tunnel}	Multiplier if in a tunnel

ECOTRACK Rail Replacement Model

The ECOTRACK rail model includes three methods of calculating when to undertake rail replacements, with a rail replacement occurring when one of the models passes its threshold (Jovanovic and Pearce, 2000). There is an age based and usage based model where rail replacements occur after a set time frame or amount of cumulative MGT. The thresholds are based on the rail type and steel used, sleeper type and operational conditions. The third model involves an exponential degradation, related to cumulative tonnage and number of failures. Cumulative tonnage is increased by a factor related to the track curvature, to increase degradation on curved sections of track.

Despite not predicting rail faults this model gives a clear indication on reasons why a rail would be replaced. By using a combination of usage, time and amount of faults thresholds the rail will be replaced before its condition becomes unfit for purpose, and hence dangerous. The conversion of the usage depending on the track radius agrees with other models and research into the rails, discussed in Section 2.3.2, that the curvature of the track is a leading cause of rail faults. This is due to the increased forces exerted on the rails, especially the top rail, around bends.

Stochastic Models

NR Probability Simulation Model

As part of the Network Rail tier 2 model an empirical rail defects model is used to predict

the amount of rail defects that will occur over a section of track in a set time period. The rate of defects is a function of the geometry of the track, G, Cumulative Equivalent Million Gross Tonnage (CEMGT), experienced by the rails, the curvature of the track section, C, and the section rail type and weight, S, by:

$$\text{DefectRate}_t = (A + \text{CEMGT}_t^{\alpha} \times (B + \beta\, G_t) \times (K + \delta\, C)) \times S \tag{2.44}$$

Where A, B, K, α, β and δ are section specific parameters and are altered depending on the section of track being analysed. The different rail section types are bullhead, 113lb flatbottom pre and post 1976 as well as the more modern UIC CEN 60. The curvature is calculated by dividing one by the radius of the curve, where a straight piece of track will have a curve value, C, of zero (Gordon, 2006).

The model then uses the calculated defect rate in a Poisson process simulation to derive the specific defect occurrences. Sections of $1/8^{th}$ mile and a time step of one month are used to find the Poisson rate for each time step. The CDF associated with the Poisson rate is then found. The number of defects in a time step is found by evaluating the CDF against a random number using a constant seed for a given time step. A constant seed for each time step allows different users to replicate results with the same parameters. The amount of serious defects is then calculated as a proportion of the amount of defects.

The model was well researched in its development with a large amount of data being used. It takes into account all the primary causes of rail defects, including, usage, geometry, curvature and rail type. The use of a Poisson process simulation also enables the degree of randomness of the defects to be taken into account. Some factors are not included in the model that other models do. These include; stations, due to the fatigue caused by the slowing down and accelerating of trains or reduced fatigue due to slower moving trains; poor drainage and diesel deposits and tunnels, due to harder inspection and maintenance as well as damp and corrosive conditions.

Weibull Distribution Model

A method of modelling rail breaks is proposed by Chattopadhyay et al. (2005) (as cited by Kumar 2006), where a two parameter Weibull distribution is used. This is backed up by Simoes (2008) who also suggests the use of a Weibull distribution to find the rails Mean Time To Failure (MTTF) and reliability. The Weibull distribution is used due to its ability to provide reasonably accurate failure analysis from a small sample size and because it has been proven for other mechanical components ageing, wear and degradation. Chattopadhyay et al. uses the amount of traffic as MGT which is denoted by m, as the affecting factor. The two parameter Weibull distribution used as a cumulative rail failure distribution , $F_n(m)$, and as a density function, $f(m)$, are given by:

$$F_n(m) = 1 - exp(-(\lambda m)^{\beta}) \tag{2.45a}$$

$$f(m) = \lambda\beta(\lambda m)^{\beta-1} exp(-(\lambda m)^{\beta}) \tag{2.45b}$$

Where λ is the inverse of the characteristic/scale function, η, and β is the shape parameter of the distribution. These values are found by fitting the Weibull distribution to data of times until a rail break occurs in the field using a computer program. This can be done for one section of track if enough data is available or by grouping similar sections together by conditions such as curve radius and type of rail.

It is then possible to calculate a failure intensity function, $\Lambda(m)$ which is an increasing function of m indicating how the number of failures increases with m. This function is found using a combination of 2.45a and 2.45b resulting in:

$$\Lambda(t) = \frac{f(m)}{1 - F(m)} = \lambda\beta(\lambda m)^{\beta-1} \tag{2.46}$$

The total number of failures over a period from i to $(i + 1)$ can be found using:

$$E[N(M_{i+1}, M_i)] = \lambda^{\beta}((M_{i+1})^{\beta} - (M_i)^{\beta}) \tag{2.47}$$

Where λ and β are the Weibull function parameters and M_i is the total accumulated MGT up to the i^{th} inspection.

It is then possible to calculate the MTTF using Equation 2.48a which uses the gamma function outlined in Equation 2.48b.

$$MTTF = \eta\,\Gamma\left(\frac{1}{\beta} + 1\right) \tag{2.48a}$$

$$\Gamma(n) = (n-1)! \tag{2.48b}$$

The use of a Weibull distribution allows the random nature of rail breaks to be modelled. By splitting the track up homogeneously into sections of similar aspects, such as speed, curvature, rail weight, rail type, tunnels and stations it is possible to make this model include all the important factors that cause rail breaks. Chattopadhyay et al. (2005) did not include different faults or seriousness of faults in their analysis. Base on the approach it would be possible to include these by splitting the original defect data into different types of defects before fitting Weibull parameters to the groups. This would give greater detail to the models output, improving its usefulness.

Markov Chain Method - Including Inspections

Hokstad et al. (2005) and earlier work undertaken by Dolven et al. (2004) discus a method

of combining a continuous and discrete-time Markov model to allow the continuous degradation and failures to be modelled with the discrete time transitions of inspections. This enables the probability of not detecting a fault during an inspection to be included. The model is designed to discover what the most cost effective preventative maintenance strategy is by altering the interval times of inspections.

First, a continuous-time Markov model is produced as seen as part of Figure 2.38. A fault free piece of rail occurs in OK, this then moves towards a degradation failure, F_1, through states D_{1u} and D_{2u} with transitions of μ, ω and ν. The u denotes an undetected state whereas when a fault is detected in a given state it is shown with a d. When a fault has been detected in D_{1u} it moves to D_{1d}, from which point it moves to D_{2d} with a transition of σ and then can move to failure, F_1 with transition ν if the fault at D_{2d} is not discovered. Once a fault has been detected in D_1, additional inspections are made. The ρ transition describes the rate of detecting state D_2 in these additional inspections undertaken due to detecting state D_1.

The transitions q_1, q_2 and q_3 are the probabilities of finding a fault, with q_1 being the probability of finding a fault whilst in D_1. q_2 is the probability of finding a fault, D_2, without prior knowledge of a fault and q_3 is the probability of finding a fault, D_2, with prior knowledge of a fault, D_1. When a fault is found at D_1 no maintenance work is undertaken but when fault D_2 is found work is instantly undertaken, returning the rail back to an OK state. Both F_1 and OK^* are absorbing states, after which at the beginning of an interval the state is returned to OK.

Past data is then used to obtain the transition rates with the inspections probabilities, q_n and ρ being chosen from expert judgement. It is then possible to alter both the length of inspection interval and the frequency of additional inspections when fault D_1 is discovered. From this it is possible to find the number of failures over a given track length in a given amount of time. The different inspection intervals are then compared and the most cost effective ones in terms of cost, reliability and safety is chosen.

The results from the model, using the calculated rates from the data show that over a 365 km piece of track over 5027 days the expected number of failures is 213 with a one year interval time and 312 for a two year interval time. As a MTTF per km of track, this results in 23.6 years and 16.1 years respectively. It is noted that different lines will have different transition probabilities which will need to be found from past data.

Continuous Markov models are very suitable to random failures, which occur with rail faults, due to being exponentially based. The use a time discrete Markov model is also well suited to inspections due to their constant time step and probabilities to find faults. By having different states for known and unknown conditions allows for the inspection time to be lowered when a fault has already been found, allowing the fault to be monitored and fixed at the correct time. The model does not completely model real life as once an inspection occurs and a major fault is found the corrective work is assumed to be instantaneous. There is also no other maintenance considered as part of

the model other than renewal which returns the model back to a new state. The models parameters have to be calculated for each different section of track to be used, but it would be possible to categorise similar pieces of track together to obtain a list of failure rates for different track types. This could be split up by factors known to affect rail faults such as speed, rail weight, rail type, curvature and tunnels.Additional failures can also be added such as 'shock' failures that can occur in any state. It is possible to use the same approach for other assets that degrade but the inspections are not guaranteed to find the faults (such as sleepers with visual inspections).

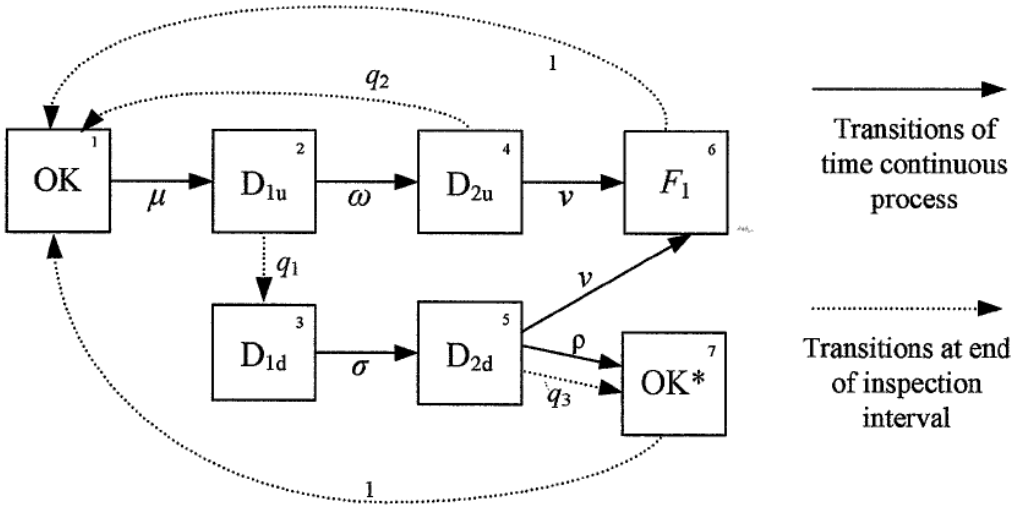


Figure 2.38: Markov rail failure and maintenance model (Hokstad et al., 2005)

2.4.5 Existing Railway Track Degradation and Maintenance Models Summary

Track Geometry Degradation Models Summary

Many models have been proposed to calculate railway geometry degradation, including deterministic, FIS, adaptive networks and stochastic methods. A large selection of models have been introduced within this section and are compared in Tables 8.1, 8.2 and 8.3, in the Appendix.

As part of the deterministic models, there have been many empirical relationships suggested. These tend to link either the amount of cycles or traffic as tonnage with the track settlement or strain. As settlement and strain within a track causes vertical geometry degradation, with it estimated by Chrismer and Selig (1994) that the SD increases by 15% of the settlement, it can be said that the proposed relationships for strain and settlement are similar for the vertical geometry SD. It can be seen in Table 8.1 that in general the degradation rate immediately after a renewal or maintenance degrades quickly but decreases with time/usage. This happens due to the ballast stones after work occurring not being in equilibrium and hence early forces tend to move the stones

a lot, once the ballast is compacted again so the ballast skeleton is interlocked the rate decreases. The point at which the rate reduces varies in literature with Shenton (1985) estimating it happening at 1,000,000 cycles, Iwnicki et al. (2000); Ebrahimi (2011b) estimating its occurrence at 10,000 cycles or 0.3 MGT and Profillidis (2006) estimating 2,000,000 standard axles. Sato (1995) discuses how the point which vertical geometry degradation slows to dependent on the quality of the initial ballast packing, which maybe the reason for the disparity within the literature. Additionally, the load of each cycle may vary between the different experiments, which would change the speed the ballast reaches equilibrium.

The relationships found in the literature to quantify the fast early degradation which decreases vary. Semi-logarithmic relationships are proposed by Alva-Hurtado and Selig (1981); Profillidis (2006) with Ebrahimi (2011b) also recommending its use in the Wisc-Rail model up to 0.3 MGT followed by a linear relationship. A power relationship is recommended by Selig and Waters (1994) and Shenton (1985) suggests its use for up to 1,000,000 cycles followed by linear. Even an exponential relationship with time as the main input is advised by Velt (2007) but this line differs from the other suggested relationships, resulting in slow early degradation which increases with time. Sato (1995) also suggests an exponential relationship but with tonnage as the main input and an additional linear part to the relationship. A linear relationship between the track vertical geometry and tonnage is presented by Gular et al. (2011) for the multivariate statistical analysis approach, Jovanovic and Pearce (2000) for the ECOTRACK system and Andrade and Teixeira (2011); Caetano and Teixeira (2016). Andrews (2012) also uses a linear link between the vertical geometry SD and time. Other non-standard relationships are also proposed for the result of tonnage by Iwnicki et al. (2000) for the TU Munich settlement model, Shafahi and Hakhamameshi (2009) for the ORE track quality model and NR uses an exponential-power equation to link the track SD with time. By assessing the models it can be seen that the consensus is that the track initially degrades fast but slows and becomes linear. The point this slows is important as if it is low, like the 0.3 MGT suggested by Ebrahimi (2011b), then it is acceptable to assume that the initial degradation can be ignored and a simple linear relationship used, as was decided as part of the ECOTRACK system (Jovanovic and Pearce, 2000).

Across the models discussed there are 3 main inputs used; cycles, tonnage and time. Cycles and tonnage are very similar as each cycle is equivalent to one axle pass, whereas the tonnage is a combination of all the axle passes multiplied by the axle load. In general cycles and tonnage are used for the deterministic models, whereas the stochastic models tend to use time as the degradation causing factor. The stochastic models, being based on actual track geometry recordings, tend to be in time as the inspections are recorded as a date and not as a current usage. Using time does not make much sense as a track will not degrade if there are no trains traversing it as there are no forces to cause settlement. Additionally, by using time it is then difficult to model the affect of changes in the amount of traffic. Due to these reasons it is recommended that degradation is calculated

in respect of usage, not time.

There are many factors which affect the rate of geometry degradation, with the additional factors considered compared in Tables 8.1, 8.2 and 8.3. Across the models there are certain factors with are agreed upon. The track speed is generally considered to be important due to the increased dynamic loads caused by faster trains. Many models have taken this into account either directly (Sato, 1995; Gular et al., 2011; Halcrow, 2012; Shafahi et al., 2008) or indirectly as another factor such as pressure within the ballast (Iwnicki et al., 2000; Sato, 1995; Shenton, 1985). Axle loads is another well document factor that affects geometry degradation with increased axle loads causing higher forces to occur within the track and hence settlement. Again it is included either directly (Shafahi and Hakhamameshi, 2009; Ebrahimi, 2011b; Halcrow, 2012), or indirectly (Sato, 1995; Iwnicki et al., 2000; Shenton, 1985). To account for different degrees of traffic many models which used time as the primary factor also include some traffic level element (Halcrow, 2012; Shafahi et al., 2008; Shafahi and Hakhamameshi, 2009). Other factors which have been considered within the models discussed include; subgrade type (Shenton, 1985; Profillidis, 2006; Sato, 1995; Ebrahimi, 2011b), sleeper type (Shenton, 1985; Ebrahimi, 2011b; Gular et al., 2011), sleeper spacing (Shenton, 1985; Sato, 1995; Halcrow, 2012), ballast condition (Shenton, 1985; Sato, 1995; Profillidis, 2006; Ebrahimi, 2011b; Halcrow, 2012), rail type (Shenton, 1985; Sato, 1995; Iwnicki et al., 2000; Gular et al., 2011; Jovanovic and Pearce, 2000; Halcrow, 2012), ballast thickness (Sato, 1995; Ebrahimi, 2011b), topography (Shafahi et al., 2008; Gular et al., 2011; Shafahi and Hakhamameshi, 2009), track aspects i.e. stations and switches (Halcrow, 2012; Jovanovic and Pearce, 2000; Andrade and Teixeira, 2011) and track design i.e. curvature and cant (Gular et al., 2011; Jovanovic and Pearce, 2000; Halcrow, 2012; Shafahi et al., 2008). Prescott and Andrews (2015); Andrews (2012) include the amount of previous maintenance actions as a way of including the ballast condition, as ballast fouling is the reason degradation increases after each maintenance action.

The models discussed fall into three main types; deterministic, adaptive networks and fuzzy logic and stochastic, each with their own advantages and disadvantages. Deterministic models are based on the fundamental laws of physical science and are generally quantified through laboratory experiments. They are applicable to almost any situation, are highly intuitive and laboratory results are easy to achieve, but they have limited robustness and struggle to model uncertainties. Despite requiring no/minimal data to set up, deterministic models tend to require a lot of information to run, as seen by the large quantity of additional factors required for the deterministic models compared in Table 8.1. Stochastic models tend to be fit to historical field data using autoregression. The resulting models are robust and accurate with the inclusion of variability as encountered in the real world and uncertainties, but they require good data, have limited applicability and are non-interpretive. Adaptive networks and fuzzy logic require less data than stochastic models with possible improved accuracy and with the use of fuzzy sets variability and uncertainty are incorporated but like stochastic models they are

non-interpretive and have limited applicability (Kalogirou, 2006). Even though adaptive networks have been shown to make good models for predicting aspects such as next maintenance action and track quality next year for one track section, they are not so suitable for modelling the whole system, including degradation, inspections, maintenance and renewals, for multiple track sections over a long time period, hence they do not tend to be used for LCC models. After assessing the models within this chapter and looking at the advantages and disadvantages of the modelling types, stochastic techniques seem to be best suited for creating a LCC model for the railway track primarily due to the variability of track degradation that occurs in the field but also due to the difficulty of knowing the information required for a detailed deterministic model.

Of the stochastic techniques discussed Markov and PN models have been shown to make effective track models. As both are graphical, the model design tends to be intuitive to the system being modelled. Comparing the applicability of the discrete-time and continuous-time Markov models proposed by Shafahi and Hakhamameshi (2009) and Prescott and Andrews (2015) it could be seen that a continuous-time Markov chain is much more flexible, allowing aspects such as inspections to be modelled and the amount of maintenance actions to be outputted. Both continuous-time Markov models and PNs are flexible and allow for informative graphical results. In general Markov models are quicker to solve than PNs and are well-established as a modelling method but once they become more complex such as including multiple track sections and maintenance actions a state-space explosion can occur, which is where the model becomes too large to solve. Other issues with Markov come from the memoryless property where the model decisions are based only on the current system state and not on the sequence of events that proceeded it. Additionally, Markov is based around the exponential distribution and hence if the movement between states is not exponentially distributed a Markov model is not applicable and will give results that are wrong. This makes Markov models not applicable to railway track vertical geometry degradation, which is shown by Audley and Andrews (2013) to follow a Weibull distribution. PNs on the other hand are much more flexible and can use any type of distribution for the transitions. Additionally, aspects such as changing transition rates can be incorporated with minimal extensions to the net. The main disadvantage of using a stochastic PN approach is the computation time, as they have to be solved using random sampling such as Monte Carlo, which for a complex net may take a long time to converge. By assessing the advantages and disadvantages of the modelling techniques introduced, it was decided that a PN approach would be most suitable for a track LCC model. This is because of its flexibility for modelling all aspects of a dynamic system, the inclusion of any distribution type, usefulness of the outputs, pictorial representation makes it easy to understand and map to the system and variability and uncertainty.

Maintenance Effectiveness Models Summary

The effectiveness of maintenance activities, with accurately predicted returned track quality are an important aspect of a track asset management model. This is because, which maintenance activity to undertake and when, are important decisions that have to be made by the responsible party. The degree of effectiveness of tamping and stoneblowing have previously been related to the track quality before maintenance by linear relationship (Halcrow, 2012; Caetano and Teixeira, 2016; Stephen M Famurewa and Kumar, 2015). An exponential relationship is also demonstrated which shows that as the track ages the level of quality the track can be returned to reduces, with the rate increasing with time (Velt, 2007). This differs from Prescott and Andrews (2015) who suggests that after seven tamping operations maintenance reaches an equilibrium. Another method used is to use distributions of the returned track SD values, of which it was found that the log-normal distribution seemed to be the best fit. Different log-normal parameters were found for different speed bands and the previous amount of tamping operations. From this the probabilities of returning the track to different SD levels could be found (Audley and Andrews, 2013).

Amongst the different models discussed different factors that alter the effectiveness of maintenance activations can be seen. These consist of; track speed (Halcrow, 2012; Audley and Andrews, 2013), track age (Caetano and Teixeira, 2016; Velt, 2007) and the amount of previous maintenance actions (Audley and Andrews, 2013). It can be said that both the track age and amount of previous maintenance actions are both easily quantified values that give an indication of the level of ballast fouling, with the effect of ballast fouling on maintenance effectiveness discussed on Page 26. It would be more accurate to use a BFI instead as different track becomes fouled at different rates, but an in-situ track's BFI is hard to quantified and often modelled instead, but these models are only an estimate and have their own errors.

After assessing the models it can be seen that the initial track quality is probably the most important aspect. It has also been shown that the ballast condition is important. To account for this the amount of previous maintenance actions is a good representation as it is known that each tamping action causes a large quantity of fines. The track speed is also shown to affect the maintenance effectiveness but this is not backed up by any physical reason, hence further research into this is required.

Rail Degradation and Fault Model Summary

There have been many models produced to predict rail faults incorporating both deterministic and stochastic methods. Assessing the models it is possible to see many factors that are related to rail faults. Amongst these are; tonnage (Denby, 2013; Jovanovic and Pearce, 2000; Gordon, 2006), curvature (Denby, 2013; Jovanovic and Pearce, 2000; Gordon, 2006) and rail type (Denby, 2013; Jovanovic and Pearce, 2000; Gordon, 2006).

Other factors are suggested by Denby (2013) including track vertical geometry, stations and tunnels.

Due to the randomness of rail faults and breaks, a stochastic model seems to be more applicable as the resultant outputs can show the mean amount of failures as well as the probabilities of different amounts of failures, as with the Weibull distribution model proposed by Simoes (2008). This is helpful within a LCC model as it quantifies confidence in the estimated costs, with the output costs having a spread of values. The possible inclusion of the probabilities of an inspection, such as incorporated in the Markov model proposed by in Hokstad et al. (2005), is also a useful addition to a rail fault model. Many faults are not easily identified until the defect reaches a large size, and hence faults can be missed in inspections. By adjusting the probability of finding faults, the level of worker competence and equipment errors can be included within the model.

2.5 Literature Review Summary

The effective use of asset management within the railway industry could increase the return of investment, improving safety and performance or reducing costs. To ensure that the best asset policies are chosen, tools such as LCC models are used to predict the future costs and asset conditions of different options, which are then compared and used as evidence to support changes to the policies. Due to this it is imperative that the LCC model correctly represents the system giving accurate future costs. It is important that policies can be changed within the model so comparisons can be made. An LCC model requires detailed information on the systems assets, conditions, degradation mechanisms, effectiveness of maintenance activities, costs and the availability of work access and resources. With these it is possible to design an LCC model that can predict future costs and performance by recording asset conditions and the amount of inspections, maintenance and renewals actions taken over a time period.

A railway track is made up of multiple assets, including; rails, sleepers, ballast and subgrade. The rail is the running surface for the train but due to experience high forces fatigue tends to occur causing faults such as RCF and possible fail breaks. Rails are either inspected visually or by ultrasonic testing and are maintained either by grinding, to remove surface defects, or by rail replacements. There are two methods of joining rails, jointed and CWR, with CWR being shown to experience fewer faults. Other aspects that can affect the rate of rails faults include track tonnage, curvature, rail type, vertical geometry, stations and tunnels. The sleepers hold the rails in the correct place and transfer the load down into the ballast whilst providing longitudinal and lateral stability. There are three main types of sleepers used: timber, concrete and steel. Timber sleepers are good at absorbing shocks and vibrations, but they have a short life and their low weight reduces their transverse resistance. Concrete sleepers have a long life and are very heavy resulting in a high transverse resistance. Steel sleepers are cheap and with

a long service life but, like timber sleepers, are light and hence have low resistance to movement. Sleepers have failed when they no longer support the loading from the train and are replaced when required. The ballast transmits the forces down to the subgrade, absorb vibrations, resists the track shifting and allow drainage. When trains traverse the track, the ballast skeleton shifts, changing the track geometry. It also tends to become full of fines, known as fouling, which reduce the ballast drainage and load distribution, increasing the rate the ballast stones move and the force on the subgrade. Fouling is caused by the infiltration of fines from the environment and subgrade, as well as ballast breakage caused by passing trains and tamping. The ballast is maintained by either cleaning the ballast to remove the fine material or just replaced. The subgrade layer which the track was built on tends to settle with repeated loading, causing changes in the track geometry, but the majority of movement occurs within the ballast.

Railway tracks quality is primary categorised by the vertical geometry, as this affects the quality of the ride for the passengers as well as the maximum safe speed. Poor uneven vertical geometry can lead to derailments so speed restrictions or line closures are required to maintain safety. These are expensive so an effective LCC model would accurately predict the degradation and minimise the amount of speed restrictions and line closures. The vertical geometry is recorded as a SD of an average between the two rails vertical geometry after a 35 m filter is used to remove long undulations in the geometry which do not affect the ride or safety. The track geometry is recorded by an inspection train, and multiple SD recordings for a track section can be used to quantify the degradation. Previous research suggests that after maintenance the geometry initially degrades quickly but slows down to become linear with the time required until it becomes linear is dependent on the initial packing of the ballast. To fit this shape multiple relationships have been suggested including semi-logarithmic, power, exponential and linear (ignoring the initial period of fast degradation). The speed at which the geometry degrades is related to many aspects of the track including track speed, axle loads, usage (tonnage), subgrade type, sleeper type and spacing, ballast condition, rail type, ballast thickness, topography, track aspects such as stations and switches, track design i.e. curvature and cant. There is also a general variability in the rate of degradation where two seemingly identical track sections will still degrade slightly differently from each other. The track geometry is improved by tamping and stoneblowing maintenance actions. Tamping involves lifting the track to the desired level, then vibrating tines push the ballast stones to fill the gap below the sleeper. Stoneblowing also lifts the sleeper but injects small stones underneath it to fill the gap. The returned track quality after maintenance tends to vary, with previous research showing that the initial quality, track speed, age and amount of previous maintenance actions affect the returned track geometry SD. Age and amount of previous maintenance actions are used to categorise the ballast fouling, which has been shown to affect the returned quality. The actual ballast fouling is not used as it is hard to measure in the field.

Many modelling techniques have been used previously to model vertical geometry, main-

tenance effectiveness and rail faults. These fall into three main types: deterministic, adaptive networks and stochastic models. The deterministic models are based on the fundamental laws of physical science, can be applied in almost any situation, are highly intuitive and normally based on laboratory results which are generally easy to obtain. However, they lack robustness, do not model uncertainty and generally require a lot of information which may be hard to obtain in the field. Adaptive networks including ANFIS models require less data than stochastic models, can be very accurate and the use of fuzzy sets adds uncertainty and variability, but they have limited applicability and are not flexible enough to model a complete system with degradation, inspections, maintenance, renewals and resource allocation. Stochastic models are fit to previous measurements taken from the field using auto-regression, resulting in a robust model that can model uncertainty and variability. The outputs achievable from stochastic models are much more informative than deterministic, giving the spread of possible results which can then be used to give the probabilities of certain results. For example, this makes it possible to say that there is a 80% chance the cost will be between two numbers. Stochastic models tend to require a large amount of processing to achieve results, with many methods requiring random number generation and multiple simulations. Due to the ability to model the complete railway system with the variability seen in the field and obtain detailed results, stochastic models are a very good fit for creating a railway track LCC model.

There are many ways of modelling stochastically. Markov models are well-established in the area and can be either discrete or continuous and give a graphical representation of the system. Discrete Markov models are not flexible enough to create a whole LCC model due to the use of probabilities and a constant time step. Continuous Markov models are much more flexible and can be solved quickly using algorithms. The issue with Markov is its memoryless property, which is due to being based around the exponential distribution. Due to this a Markov model bases the next movement only on the current system state ignoring any history. Additionally, the ability to only use exponential distributions means that aspects which do not have random movement cannot be included, and as it is not expected for the track degradation to follow an exponential fit, a Markov model is not applicable to the problem. Markov models also tend to grow in size very quickly as aspects are added, such as multiple maintenance actions, to a point where it may take a very time to solve. Another very flexible, graphical stochastic modelling technique are PNs. They have most of the advantages that Markov models have except can take much longer to solve due to having to be solved via Monte-Carlo sampling, with many simulations. Unlike Markov models PNs are not based around the exponential distribution so are not memoryless and allow transitions to follow other distributions. As PNs are flexible enough to build a whole LCC model with, allow for the general variability seen in the field to be included, produce useful outputs for asset management and can be designed to allow inputs to be changed which enables comparisons between asset policies to be undertaken. Due to this it was decided that a PN approach is the best suited to creating a railway track LCC model in this thesis.

Railway track degradation is caused by trains traversing the track, due to this within the model the degradation should be related to the track usage (tonnage) and not time which is common in previous stochastic track model. This should result in more accurate degradation information and also allows for the model to account for changes in the track usage. Other aspects that previous stochastic models have not included but which will be explored in this thesis include: track degradation related to the many aspects which are known to affect it; maintenance effectiveness related to the current track condition; rail faults and breaks included with track geometry degradation with the link between the two also modelled; data uncertainty brought through the model into the outputs and memoryless degradation. To accomplish these an analysis of the factors which affect the rate of degradation will have to be undertaken so only the most important aspects are included in the model (Chapter 3); analysis of the link between the quality before and after maintenance quantified stochastically (Chapter 4); analysis of the cause of rail faults in particular the link between track geometry and rail faults and find a way to model the link (Chapter 5). Production of a LCC model using PNs is introduced in Chapter 6.

Chapter 3

Track Geometry Degradation

3.1 Introduction

This chapter explores the main measure used to categorise the degradation of the railway track, the data obtained, how it has been organised and used, how the rate of degradation was found and then how this information plus additional asset and usage information was used to identify the leading factors that affect the rate of degradation. This is important as when producing an asset management model an understanding of all the contributing factors to degradation is required so they can all be included and hence lead to a model that better represents the real world. It is also imperative to not include any factors which have no effect on the rate of degradation and hence the output of the model, as by including unnecessary inputs, the model is made more complicated to use with no additional benefits. Additionally, as the model to be developed will be stochastic in nature, it is important to maintain large amounts of data for the model to be based on, with every additional factor reducing the size of the datasets used within the model.

As part of the analysis, the major factors that effect the rate of degradation of the railway track have been identified, such as sleeper type or track speed, and the choices within these factors statistically compared and grouped if possible answering questions such as, does a particular sleeper type lead to faster or slower degradation compared to another. This information will help to identify reasons for different track section's rate of degradation and help to make informed decisions on renewal/upgrade choices and the running of the track.

It was decided that the degradation would be related to the track usage and not to time, as from the extensive literature review and data, it could be seen that the primary cause of degradation on a railway track is the running of trains on the line. The creation of a usage based asset management model would also allow changes in the track usage to be modelled stochastically, which is more difficult in a time based model. This is important for the railway industry as the usage of the train lines changes by varying

degrees each year, which some yearly increases of over 200% occurring on the UK rail network. These changes will increase with additional factors such as the use of new technology, i.e. moving block signalling, and speed changes.

3.2 Available Data and Preparation

3.2.1 Introduction

This section reviews the historical data of the UK rail network and discusses how it has been assembled to allow degradation rates to be found and an analysis of the contributing factors to be undertaken. A large quantity of data was obtained from the owner and infrastructure manager of the UK's railway network, Network Rail (NR), including eight years of track geometry recordings and details of mechanical maintenance interventions, such as tamping and stoneblowing. Renewal information has also been obtained, as well as knowledge of the track assets installed and their locations, track usage history and geology information. This data has been used to calculate the rates of degradation of the track sections over different phases of its life cycle, as well as the effectiveness of maintenance activities. Maintenance activities were chosen as the place to split the data into phases, with the degradation rate being categorised by the track section's maintenance history. This was decided as from extensive research it is known that as the track assets degrade the rate of degradation of the vertical top geometry increases, primarily due to the quality of the ballast reducing as it becomes fouled with fine particles. This occurs due to the build up of fines from the atmosphere and passing trains, as well as from ballast breakup caused by maintenance activities, passing trains and infiltration from the subgrade, (Selig and Waters, 1994). As the ballast fouling increases, the fine particles block the voids reducing drainage and the ballast's resilience to deformation, leading to faster vertical geometry degradation, (Ebrahimi, 2011a). As maintenance, in particular tamping, causes a large amount of ballast fouling, it is a good idea to use the amount of previous maintenance activities to group together track sections, as these will have similar ballast fouling levels.

The acquired data was measured over different length track sections with the UK railway network being split up into many parts. The highest level are the Strategic Route Sections (SRS's), which are tens of miles in length. These are then split up into Engineer Line References (ELRs), and then Segment IDs, which divide the network into sections with the same traffic. These are then further broken up into 220 yard sections, known as position keys or poskeys for short. The poskeys are then fragmented further into track IDs which vary in length from 1 to 220 yards, and are dependent on changes in the track, e.g. a change in sleeper type or an Switches and Crossing (S&C) would result in a different track ID within the poskey. Table 3.1 lists the available data, as well as the track section type they are recorded for. To use all the data together it was first required to relate all the data to the same section type. To maximise the amount of data and

hence confidence in a statistical analysis it is best to split up the railway network into as small of sections as possible. This also increases the amount of sections which will have consistent information, such as sleeper type, across their whole length, which was required for the detailed analysis.

Table 3.1: Available data and their recorded section types

Data type	Section type
Geometry recordings	Poskeys
Maintenance records	Poskeys
Renewal records	Track IDs
Asset information	Track IDs
Track usage history	Track IDs
Track usage breakdown 2014	Segment IDs
BFI history	Track IDs
Subgrade geology information	Varying lengths not related to any other section type

3.2.2 Geometry Recordings

Measurement trains have been used to obtain track geometry measurements on the UK rail network since 2006. The train obtains many geometry recordings, primarily including the vertical top profile and the alignment. The vertical top profile is considered to be the most representative measure of track quality as it affects many aspects such as ride comfort for the passengers and the permissible, safe speed limit for the line. The 35 m wavelength Standard Deviation (SD), which is discussed on Page 21, gives a good indication of the track quality, and with the use of many recordings taken over the years, a good way to calculate and categorise track degradation, as well as the effectiveness of maintenance operations.

NR has supplied a database consisting of 5,622,354 vertical top SD measurements from the year 2006 to 2014, with the date and poskey of each measurement also recorded. The calculated SDs are recorded in general to one decimal place but some more recent recordings are calculated to two decimal places. With the computed SDs being rounded to one decimal place, a degree of accuracy had been lost, making small changes in the quality of the track geometry hard to identify. The obtained data was accurate to 0.1 mm though, which it was decided, still had the required degree of accuracy to see the type of changes in track quality which were expected over the large time periods between railway track maintenance activities.

Due to the vertical top geometry SDs being recorded for the 220 yard sized poskey sections, and with no way of calculating it for different sized sections, it was decided that all the other obtained data would have to be known for individual poskeys too. Due

to inconsistent data such as asset type or track speed, some poskeys were removed from the analysis. They were only removed where required, if sleepers were being analysed, as long as the sleepers on the whole poskey were consistent even if the track speed was different it would still be included. Ideally the geometry recordings would have been calculated for the smaller track IDs as their smaller size would result in more track sections and more data to undertake a statistical analysis of the causes of degradation. The track ID's also have consistent asset data within them, so data would not have to be removed, as it is with the use of poskeys, when differing assets appear in the same section, see Section 3.2.5.

3.2.3 Maintenance Records

Mechanical maintenance records, including tamping and stoneblowing activities, have also been obtained. These are recorded for the poskeys and comprised of 376,006 records in a database. The type of each activity, either tamping or stoneblowing, as well as the date the work was completed, were included in this database. Due to already being recorded for the poskeys, no further work had to be completed to line the maintenance records up with the geometry recordings.

As it was decided that the degradation fits would be grouped by the amount of previous maintenance activities, for the poskeys which underwent their latest full renewal before the start of the maintenance records, an estimate of the amount of maintenance activities prior to the records starting was required. This was needed to understand each track sections full maintenance history and its impact on the railway track degradation. Without the estimated previous maintenance before records, only newer track (less that 10 years old) could be included within the analysis on maintenance histories. The maintenance records started in 2005 and as stoneblowing was rarely used before this, it was decided to presume that all major pre 2005 maintenance would have been tamping. This presumption will result in all track sections which underwent a stoneblowing operation pre 2005 to be categorised in the wrong maintenance history, but as stoneblowing was rare before the year 2005, the amount of poskeys categorised with the wrong maintenance history due to the lack of stoneblowing would be very low. To estimate the amount of tamping operations, Table 3.2, was obtained from NR, where the table was created from past data and asset management policies. Table 3.2 first uses the track category to obtain a base tamping frequency in tamping operations per year, where the track category is set using the poskeys speed and usage using Figure 3.1. The Ballast Fouling Index (BFI), which is introduced and organised in Section 3.2.8, is then used to alter the base frequency, as new track tends to require very little maintenance work whereas poorer condition track with a higher BFI will degrade faster and require more maintenance actions. The overall tamping frequency can be obtained by multiplying the base frequency by one plus the BFI alteration value. This was calculated for each year between the full renewal of the poskey and the start of the maintenance records, then the total sum was rounded to

obtain an estimate of the total amount of tamping operations pre 2005 for each poskey. This can be expressed as Equation 3.1, where y is the year the poskey last underwent a full renewal and n is the year being calculated.

$$\text{Total Tamps Pre 2005} = \lfloor \sum_{n=y}^{2004} \text{Tamping Base Frequency}_n * (1 + \text{BFI Alteration}_n) \rceil \tag{3.1}$$

The maintenance records had to be cleaned to remove additional maintenance activities where the data seemed to have recording errors. There were multiple duplicated maintenance activities that were recorded having to have occurred at similar times to each other in the data. These should not exist as two mechanical maintenance activities would not occur close to each other. To find and remove these duplicated records it was decided that subsequent records that occur within 30 days of the previous maintenance activity should be deleted as these are most likely a record mistake. The choice of 30 days came from discussions with engineers at NR, and by looking at NR's asset management policies to see at what points maintenance is scheduled. From this it could be seen that the closest that two mechanical maintenance activities should occur is about 6 months apart, but can be lower if required, but most of the time if this is the case a renewal will occur instead, as it is not feasible to be maintaining the track more than twice a year with mechanical tampers and stoneblowers.

Table 3.2: Tamping frequencies [obtained from Network Rail]

Track Category	Tamping Base Frequency	BFI <25	25< BFI <50	50< BFI <75	BFI >75
1A	0.5	-0.26	0.11	0.09	0.25
1	0.33	-0.24	0.12	0.02	0.21
2	0.33	-0.24	0.04	0.12	0.09
3	0.33	-0.22	0.07	0.06	0.08
4	0.25	-0.15	0.04	0.11	0.01
5	0.16	-0.27	-0.03	0.16	0.21
6	0.10	0.09	-0.12	0.16	0.55

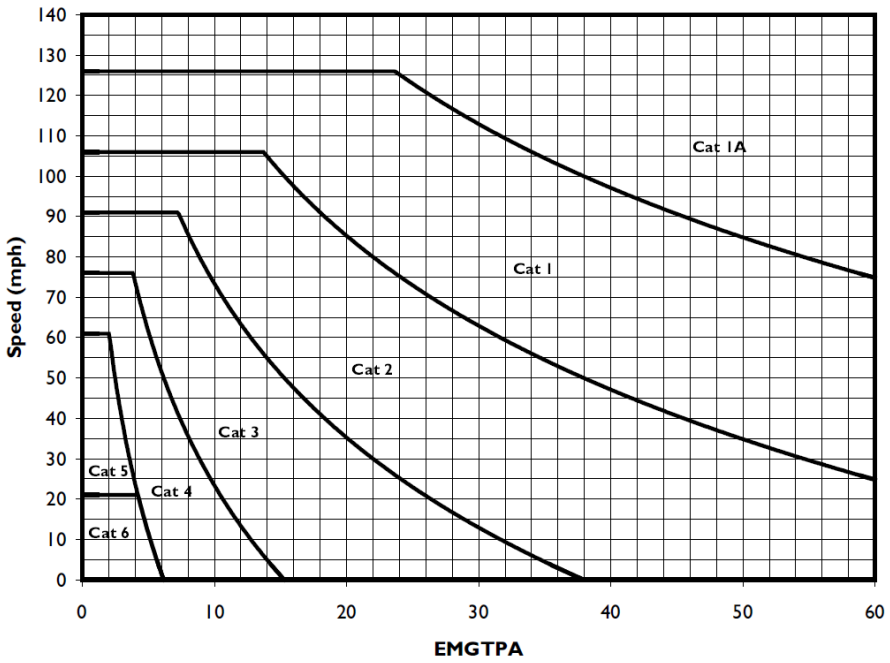


Figure 3.1: Calculating Track Category, (Network Rail, 2011)

3.2.4 Renewal Records

Records of the year of all rail, sleeper and ballast renewals for each Track ID were given as an 11,085,265 row database. The database included a record of each track ID for the years 1995 to 2006, and the year of the latest rail, sleeper and ballast renewal there was in that year. Using this database it was possible to find the latest full renewal for each track ID. Due to being recorded for each track ID, these had to be related to the larger poskey sections to line up with the geometry recordings. As part of this the entire poskey had to have a consistent rail, sleeper and ballast renewal dates and these all had to be the same year for the poskey or the poskey data was deleted. This was undertaken as it was desirable to have all poskeys starting from a full renewal where the rails, sleepers and ballast for entire poskey had been renewed at the same time, as it meant that all poskeys that were included in the further analysis and in the stochastic model data started in the same condition. This resulted in 81,965 poskeys which had undertaken a known full renewal.

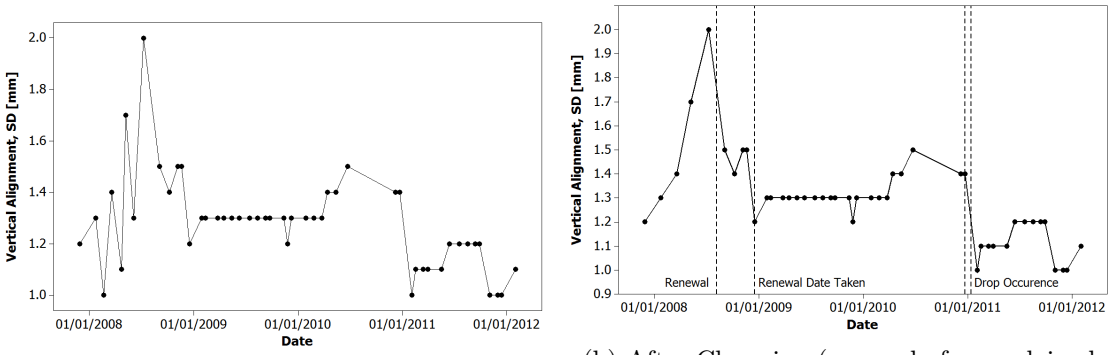
Due to the renewals being recorded as only the year of occurrence and not an exact date, the track geometry recordings were used to obtain an informed estimate of the actual renewal date. The largest drop (improvement) in SD was used to identify the full renewal in the year, with the lower SD recording taken as the renewal date. Before this was undertaken the vertical top geometry recordings had to be cleansed of outliers, which could possible show a large drop in the SD and hence be picked up as the renewal date when they should not be. These outliers were classified as unexplained large drops and peaks and temporarily removed from the data and later returned before the calculation

of the degradation fits in Section 3.3.2. It was decided that an unexplained drop or peak with a SD change of 0.4 mm or more would be removed. This was decided as from research it could be seen that despite general geometry fluctuations occurring such as seasonal track quality changes these would be unlikely to be as high as 0.4 mm, hence it was presumed that recordings creating drops or peeks over 0.4 mm were probably errors, unrecorded maintenance actions or manual maintenance. An example of how the data was cleansed can be seen in the year 2008 of Figure 3.2, where the recordings of the large unexplained dips have been removed.

After a renewal a large bedding in period can occur, where the ballast skeleton, which is made up of ballast stones, moves causing track geometry improvements, until the skeleton reaches an optimum equilibrium. To take this into account, track recordings over the first 240 days after full renewal were checked for further geometry improvement. A new renewal date was taken as either a lower SD recording within 120 days, a SD recording between 120-180 days which has an improvement of more than 0.1 mm or a SD recording between 180-240 days which has an improvement of more than 0.3 mm. These were chosen as a bedding period of 120 days was expected but it may take longer. The reason that greater improvements were chosen to be required the more days past the renewal date was to stop a minimum being found in the first 120 days and then a similar or very slight improvement occurring at a further point, which could be due to general fluctuations and were not enough of an improvement to move the new renewal date for, as that would affect the degradation curve fitting later. The lowest of these recordings is used as the new full renewal date, with all previous recordings removed. Figure 3.2 demonstrates how the bedding in period is taken into account with the largest improvement in SD occurring around August 2008, but the actual renewal date taken as in December when a lower SD recording occurs. This would then be taken as the start of the data that would be used to calculate the track degradation.

As the track recording records only go back to 2006, for full renewals that occurred in a year with no track recordings it was assumed that the renewal occurred in the middle of the year. This results in the minimum possible error, with the estimate being at most half a year out. The middle of the year has also been used as an estimate of the full renewal date for all poskeys which did not have at least three track recordings, spread over the whole year. It was decided that three records was not enough to have confidence the largest drop was when the renewal happened. Additionally, if the largest drop, due to the lack of data, occurred between two geometry recordings 200 days apart, the lower value would have been presumed as the renewal date, when the renewal could have actually occurred very close to the first recording and hence the estimate would be nearly 200 days out. Due to the added uncertainty in estimating the full renewal date without track recordings these poskeys are recorded separately. For reference in this thesis the poskeys for which the full renewal dates were estimated using track recordings will be called Data A and the poskeys with the estimated dates as the middle of the year will be known as Data B. Data A contains data with the full history known information

from records (track renewed prior to the start of geometry recordings and maintenance actions). As Data B contains poskeys which were last renewed prior to records began, the maintenance histories had to be estimated prior to 2006, as discussed in Section 3.2.3. This adds more uncertainty as the poskeys in Data B may be categorised into the wrong maintenance history.



(a) Before Cleansing

(b) After Cleansing (removal of unexplained dips, first half of the year 2008)

Figure 3.2: Example of Track Recordings of Vertical Alignment

3.2.5 Track Asset Information

Detailed information of the UK's rail network's track assets and lines was also obtained in the form of a database. This included useful information for 632,741 track ID's such as: rail and sleeper type, maximum permissible track speed and axle load, operating route and its criticality, existence of tunnels, stations, S&C, electrification types, embankments and cuttings, track curvature and cant. As this data was recorded for each track ID, again the data had to be related to the larger poskeys so the information could be used in conjunction with the geometry recording SDs for analysis of the track geometry degradation. When a poskey's track IDs had inconsistencies for a data type such as sleeper type, the data was not recorded for the poskey. For example, if a poskey was made up of four track IDs, each with the same sleeper type, this is recorded and used to categorise the poskey, whereas if one of the track IDs consisted of a different sleeper type then the sleeper type is not recorded for the poskey. This will ensure that when grouping similar poskeys together, as required to analyse the degradation rates, that all of those being grouped consist wholly of the chosen factor type.

3.2.6 Track Usage History

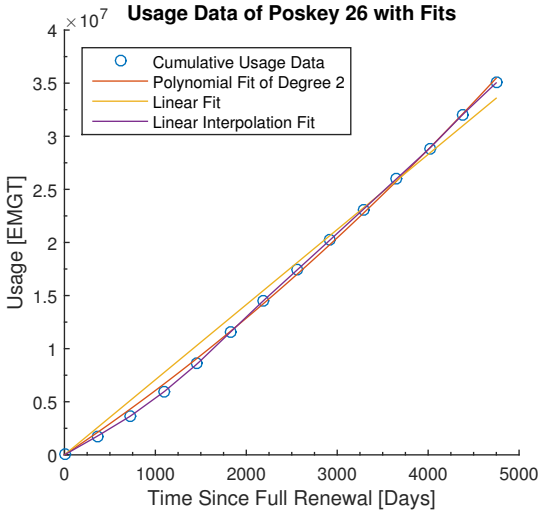
Track usage is recorded as the Equivalent Million Gross Tonnage (EMGT) that went over the line each year for the years from 1995 to 2013. A database was obtained containing 11,085,265 records including the track ID, year of recording, EMGT for the year and cumulative EMGT for the rail, sleeper and ballast since renewal, with the renewal taken as the middle of the recorded renewal year. The EMGT differs from just Million Gross

Tonnage (MGT), as it also takes into account the variations in track damage caused by different traffic types, due to reasons such as axle loads, speed and different bogies with different suspensions. Traffic is split into three types: hauled passenger, hauled freight as well as tractive units, where tractive units are the unit of a train which delivers power through its wheels. For each train type in each category the MGT is multiplied by a factor to obtain the EMGT, which are then summed together to obtain the total EMGT in each year for the track ID. The hauled passenger trains factors are based on axle loads and speeds; the hauled freight train's factors are based on the axle loads and percentage of freight trains; the tractive units factors are based on the axle loads and the power of the unit, where the higher these values are, the higher the factors associated with them are. With all the train types a reduction is given for bogies designed to reduce track damage, (Network Rail, 2011).

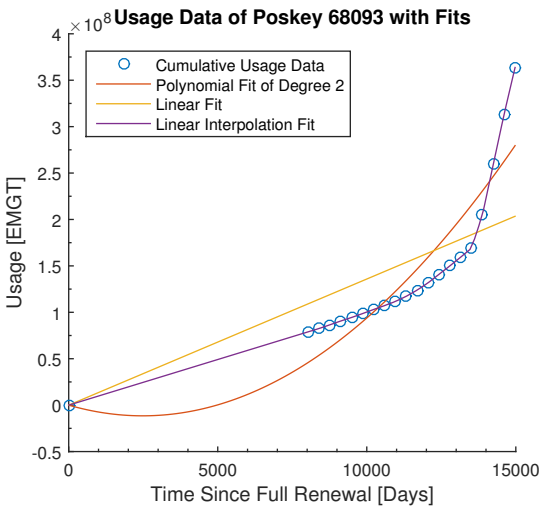
As the usage history was recorded for separate track ID's, these had to be related to the larger poskeys as with the track asset information and renewal records. Similarly to before, when multiple track ID's make up a poskey, they were checked for consistency and if all the track ID's had the same usage history then this was recorded for the poskey, if not, the usage was not recorded. As the majority of the track ID's last full renewal was before 1995, which was when the usage history records began, NR supplied estimates for the usage since renewal were used, with these being included in the cumulative EMGT records. This estimate adds uncertainty into the cumulative EMGT, but this was not an issue for calculating track geometry degradation in respect to usage as this only required the yearly usage values for 2006 onwards, as this is when the geometry recording records began. It does add extra uncertainty to the calculation of ballast fouling history as discussed in Section 3.2.8. The presumption of middle of the year renewals for the cumulative usage history adds additional uncertainty to the values but for track sections that were renewed post 2006, when the track recording measurements started, an estimated renewal date was found, see Section 3.2.4. With the estimated renewal date, linear interpolation could be used to calculate the usage since renewal in the renewal year, using the renewal years usage, e.g. if the renewal was estimated to have occurred on the 3$^{\text{rd}}$ May, which is the 123$^{\text{rd}}$ day of the year, and the yearly usage was 20 EMGT, then the usage for the renewal year after the renewal would be; $20 * ((365 - 123)/365) = 13.260$ EMGT. The calculation alters slightly for leap years when the equation would be; $20 * ((366 - 124)/366) = 13.224$ EMGT.

As the proposed model has a driving input consisting of usage instead of time it was required to have all the data, which is recorded as dates, expressed as a usage since renewal. Once the yearly cumulative usage for each poskey was plotted against time, using the renewal date as zero EMGT, a line had to be fitted to the data to allow estimates of the usage at any date to be found. Initially, linear and second degree polynomial fits were used, with the fit being forced through the origin (no usage when fully renewed). These fits had mean r^2 values of 0.9713 and 0.9966 and median r^2 values of 0.9902 and 0.9992. The polynomial fit is much better fit than the linear fit as it allows

changes in yearly usage to be modelled. Despite the high r^2 values it was decided to use interpolation instead, as even though the polynomial fit in general was very close as seen in Figure 3.3a, some occasional poskeys had poor fits, as seen in Figure 3.3b. These fits would result in significant errors in the cumulative usage estimates of works or inspections, resulting in further errors when calculating the degradation rates. The poor fits generally occurred when the poskey underwent large changes in yearly usage and/or had data missing between the most recent full renewal and the start of the usage data. Due to the difficulty of fitting an equation to the cumulative usage date, because of the possibility of different shapes and large yearly changes, a linear interpolation model was chosen to find the usage since renewal for all track recordings and maintenance activities in each poskey. By using a linear interpolation model instead of regular fits such as polynomial, even large yearly changes in usage can be taken into account, and the best estimates obtained. By using the linear interpolation, a presumption is made that the track usage is linear across a whole year, with all changes in traffic occurring between the end of one year and the start of the next. This could lead to errors in the estimates of usage at the time of an inspection or maintenance activity as the traffic increase or decrease could have occurred in the middle of the year, but as there was no more detailed data of the usage but the yearly values, it was decided that linear interpolation would give the best estimates for the usage at all stages of a track sections life history.



(a) Minimal yearly traffic changes



(b) Unknown usage and a large traffic increase

Figure 3.3: Usage history data with fits

3.2.7 Track Usage Breakdown

A track usage breakdown from the year 2014 was used to give a detailed categorisation of the traffic types that use each poskey. The UK rail network is a mixed traffic network with most lines being used for both passenger and freight trains, due to this it is important to not only know the usage as EMGT but to understand the actual traffic that traversed each track section. For this multiple databases were used including; a detailed

traffic database with 4,988,913 records, including traffic data for each segment ID which consisted of the train types, if they were loaded or unloaded, the sum of their axle load, the amount of times in the year they traversed the segment ID and the combined tonnage of these journeys. To understand the actual speed of the traffic the maximum speed of each of the train types was used. This was given in a database with 1,472 records, where for each train type, and whether it was loaded or unloaded, the maximum possible speed of that train was recorded as well as if it is classified as a freight or passenger train. This was required, as the trains traversing the track may not do the maximum permissible speed. When exploring the link between traffic speeds and the rate of geometry degradation, the track section should be classified by the actual speed (and hence dynamic forces) of trains that traverse the track.

Due to the data being recorded in segment IDs, it was required to relate the data to the smaller poskey level. This was straightforward as all poskeys that make up a segment ID have the same traffic as the segment ID. Once all the data was related to poskeys it was possible to find the percentage of traffic which was freight, the percentage with combined axle loads over 50 tonnes, the mean and median averages of the maximum speeds of the trains that passed over the poskey in a year. The amount of traffic over 50 tonnes was calculated in addition to freight as in general, passenger trains are below 50 tonnes axle load and freight are above. Using axle load moves empty freight into the lower group with passenger trains (were normally empty freight would still be classified as freight). This splits the traffic by more directly relating to the forces the track will experience instead of a simple train classification (passenger/freight) split. The percentages are calculated as both the amount of trains and also tonnage. Higher percentages tend to occur when calculated as a percentage of the usage as even if the same amount of passenger trains traversed the track as freight, which gives 50% freight by the amount of trains, due to freight trains being heavier, the percentage of fright trains by usage would be higher than 50%. Despite the calculation of the mean and median track section speeds using the maximum train speeds, it could be seen that these were virtually identical to the track sections maximum permissible speed, so for the further analysis this has been used instead.

3.2.8 Ballast Fouling History

NR also supplied their estimates for the ballast condition, as a BFI, which were calculated by one of their existing models. The provided BFI history database contained 11,085,265 records, including the last ballast renewal year, cumulative BFI and BFI increase for each year from 1995 to 2013 for each track ID. The model used estimated the BFI's using Equation 3.2, where the parameters have been defined in Table 3.3, (Network Rail, 2012). Despite the ballast fouling model developed by NR incorporating all the major factors that contribute to ballast fouling, as discussed in the literature review and seen in Table 3.3, the resultant BFI values are still estimates, and may encompass multiple

errors. These could arise from issues with the numbers in Table 3.3, or from errors in the maintenance and usage histories. This is especially the case for track sections where the last ballast renewal occurred before the year 2006, which was when the maintenance records began, as the quantity of maintenance activities prior to this had to be estimated. Additionally, the estimates of track usage prior to when usage records began in 1995, as well as unknown percentages of dirty wagons, could also lead to further errors in the BFI estimate obtained from NR's ballast fouling model. Due to these possible errors in the BFI estimates, it was decided to not use the values for analysing a link between ballast condition and vertical geometry degradation but to explore the factors that lead to fouled ballast, such as the factors included in Equation 3.2, instead.

$$BFI = \frac{EF + (TF \times MGT_{All}) + (DF \times MGT_{Dirty}) + (MF \times Tamps)}{VA} \tag{3.2}$$

To obtain the BFI values required to estimate the amount of previous work before the records started, which was undertaken in Section 3.2.3, the BFI's prior to 1995 had to be estimated. This was accomplished, similarly as with the usage data in Section 3.2.6, by plotting the cumulative BFI data against time in days, with the last full renewal recorded being as zero days. Unlike with usage where it is known that at zero days the usage was zero, we do not know the BFI value at renewal, as even new ballast contains a degree of fines. Due to this the chosen fit would have to allow extrapolation back from the start of the BFI data back to the renewal. This meant that the use of linear interpolation, which was used for the usage history, could not be used. Due to yearly changes in the rate of ballast fouling, a linear line fit also was not an acceptable fit for the data. It was decided that a polynomial line best fitted the data while allowing for the slight changes in yearly accumulated BFI that occur, also unlike the linear interpolating fit used for the usage data, it allows for extrapolation back towards the renewal. The fit was given boundaries to maintain the expected shape; a boundary of between 0.2-0.3 being specified for when time equal zero to guarantee that the new ballast has a fouling between 20-30% and the fits were checked for any negative gradient as this would not be possible without maintenance. The values of 0.2-0.3 was chosen due to the recommended initial ballast fouling recommended by NR of 0.2295, as seen in Table 3.3. This value is not zero as even new ballast will have a degree of fines. Unlike with the usage data, discussed in Section 3.2.6, where the polynomial line was not used as it struggled to fit the data when extreme changes in the rate of usage occurred, the changes in the rate of ballast fouling were much smaller allowing the polynomial equation to fit the data very well. The fit was then used to estimate the BFI for each year since a full renewal for each poskey for use in the estimation of the amount of maintenance actions that occurred before the maintenance data started. This calculation can be seen in Section 3.2.3.

Table 3.3: Parameters used to calculate BFI, (Williams, 2013)

Parameter	Description	Default Value
VA	Avaliable ballast void after renewal	22.95%
EF	Environmental fines	0.209% a year
TF	Traffic fines	0.0214% per MGT
MF	Maintenance (tamping) fines	0.578% per tamp
DF	Dirty wagon fines	0.18% per MGT

3.2.9 Geology Information

A spreadsheet containing information on the geology under the track was supplied by NR. This spreadsheet included information on the bedrock, which is the in-situ geology under the track as well as any additional superficial and/or artificial layers that were placed on top of the bedrock before the construction of the track. The spreadsheet contains information of the layer type, ELR, start and end mile and yardage of the section and its geology type. The geology sections vary greatly in size from tens of yards to over 20 miles, with a mean length of 1,055 yards. There was 3,227 records of sections with an artificial layer with equalled 829 miles of track, 20,793 records of a bedrock layer with equalled 11,006 miles of track and 20,848 records of sections with a superficial layer with equalled 7,212 miles of track. There was two unique artificial layer types, 338 unique bedrock geology types and 36 unique superficial layer geology types.

It was required to reduce the amount of geology types, as when analysing the affect the geology type has on track degradation, it was important to keep the size of the datasets as large as possible so a statistical analysis was possible. So by combining similar geology's, and hence increasing the size of the datasets, a more detailed analysis could be undertaken. To accomplish the reduction it was decided to classify the geology by their primary material for example 'sand and gravel' would be classified as just 'sand', whereas for 'gravel and sand' where gravel is the primary material, it was classified as 'gravel'. Combining similar geology's using this method does remove some detail as it increases the spread of geological properties within the classifications such as by classifying 'sand' with 'sand and gravel' where these would have slightly differing mechanical properties such as stiffness and drainage speed. Despite this it was decided that the size of the datasets of poskeys for the ungrouped geologies would not be large enough to perform a detailed analysis of the link between geology and rate of vertical geometry degradation, with only a couple of the ungrouped geologies occurring in more than twenty poskeys. A more accurate method of grouping similar geologies would be to use the geologies mechanical properties instead of the primary material, but this information was not recorded with the geology types. After the geology grouping was accomplished the amount of unique bedrock geology types was reduced to 141 and superficial layer geology reduced to 15.

Due to the geology sections varying in length, no direct link between the poskeys, which is the section size that the data was required in, and the geology sections existed. Due to this the locations of the sections had to be used. To identify the locations, the particular line was required for both the geology and poskeys, for this the ELR was used. The distances down the line for both the geology sections and poskey locations were then used to find which geology made up each poskey. The artificial, bedrock and superficial layers were checked separately for each poskey to ensure that the entire poskey has the same geology across each layer and when the poskey had differing geology within a layer the geology of that layer in that poskey was not classified. This results in 141,211 poskeys with a known consistent bedrock layer, 123,937 with a known superficial layer (including the layers where it is known that a superficial layer does not exist) and 149,021 poskeys with a known artificial layer. From these there are 107,332 poskeys with a known bedrock, superficial and artificial layer.

3.3 Calculation of the Vertical Top Geometry Degradation

3.3.1 Introduction

This section discusses how curve fitting is used to fit mathematical equations to the degradation data and how the degradation rates quantified. The organised data contained 150,952 poskeys which had recordings of track vertical geometry SDs which will be used to categorise the degradation and calculate the degradation rate. Of these 81,965 had known full renewal data, where the ballast, sleepers and rails were renewed at the same time, which was consistent across the whole poskey. With all the data related to the poskeys, including track geometry records (expressed as both days and usage as EMGT since renewal), maintenance histories (expressed as both days and usage as EMGT since renewal), asset information, geology and section traffic types, it was now possible to find the degradation rate for the different phases of the poskey's life-cycle's, and classify these by their maintenance history. It was important to split up the degradation rates by their maintenance history as this ensures that the combined datasets of degradation all relate to railway track in similar conditions. This is required when analysing the datasets to identify the main factors, as it helps ensures conclusions are based on causation not correlation. If a comparison between two datasets, such as concrete sleepers and steel sleepers, is undertaken but the steel sleeper dataset consists of degradation rates from newer track that has only been maintained a couple of times, then the concrete sleeper dataset would probably contain higher degradation values, but it could not be said that this is because of the sleeper choice. The difference would probably be due to the older track in the concrete sleeper dataset, which has already undertaken multiple maintenance activities and has high ballast fouling, degrading faster than the newer track in the steel sleeper dataset because of the loss of resilience caused by fouling and not the sleeper choice.

3.3.2 Finding the Degradation Rate

Once the poskeys had a renewal date and all vertical geometry recordings and major maintenance occurrences recorded as both time and usage since full renewal, the outliers, which were removed in Section 3.2.4, were returned and then a degradation line could be fitted for vertical geometry SD against time in days and against usage in EMGT. For Data A this started from the full renewal to the first maintenance activity, seen in red on Figure 3.4. Fits were found for between each maintenance activity, with the date or usage being normalised to zero at the time/usage of the latest previous maintenance activity. This can be seen in blue in Figure 3.4. If a drop in the vertical geometry SD occurred of more than 0.4 mm, ignoring the outliers, it was presumed that an unknown maintenance activity had occurred and further calculations of fits in the poskey were ignored. This was done as the further fits of data for the poskey would have had a previous unknown maintenance activity, and hence could not be categorised by maintenance history. The value of 0.4 mm was chosen after a discussion with engineers at NR as this would allow for small general fluctuations within the recordings, where small improvements can occur on their own due to factors such as environmental conditions but improvements greater than 0.4 mm would probably be due to a maintenance activity occurring. This could include manual track works, which are unknown from the acquired data and do not encompass the whole poskey. Some manual work is only a few meters, making modelling them difficult when using poskeys as the section size. This is due to the improvement being hard to identify and categorise with most of the manual works improvements not being visible in the calculated vertical geometry SDs, due to improving only a small part of the poskey.

The calculated fits were then recorded and characterised by their poskey and the amount of previous work. Data B, could not be started at full renewal as there were no track geometry recordings around of time of full renewals, due to this it was decided to start after the first known maintenance activity recorded after track recordings started. Then similarly to Data A, fits were found between each sub-sequential maintenance activity. The fits were then characterised by their poskey and the amount of previous works including the ones known from the maintenance records and those estimated to have occurred between full renewal and the start of the maintenance records. The use of the estimated amount of previous works is the primary reason why Data A and Data B have been kept separate as errors in these estimates will result in the found degradation rate being classified with dissimilarly degraded track sections. This may cause issues during the analysis of the primary factors of track degradation, as discussed in Section 3.3.1.

It was decided to firstly use a linear fit, as this gives a good trend of the degradation and the gradient gives a quantifiable rate which can be compared between track sections. Other fits were also explored including power, exponential and polynomial, which equations can be found in Table 3.4. These were chosen as some were similar to equations of previous research explored in the literature review, and they allow for changes in the rate

of degradation, unlike a linear fit. Unlike other common fits such as Fourier, Gaussian, sum of sine and rational, the chosen fits give sweeping curves with consistent gradient signs (except polynomial) and can also fit linear data if required. This was desirable as the fit was not expected to be a wave. Looking at the mean average r^2 and Root Mean Squared Error (RMSE) values obtained from all the poskeys and maintenance history possibilities it can be seen that polynomial gave the best fit, closely followed by power. These values for time and usage can be seen in Tables 3.4 and 3.5. Despite the polynomial equation giving the best fit statistically, its ability to have turning points where the gradient changes sign, results in fits that do not actually represent the underling degradation trend within the track recordings, which can be seen in Figures 3.5, 3.6 and 3.7. These figures also demonstrate how the linear fit, despite not always being the best fit statistically, does fit the general trend of degradation in the poskeys. By examining the fits it is possible to see that the power curve represents the data best, due to the power equation allowing for rapid changes in gradient and an almost linear part in one line. The exponential curve follows closely to the linear fit, and has no real benefit in its use. The GOF measures in Tables 3.4 and 3.5 are an average of many fits, with varying strengths, with many results with r^2 values above 0.9, but also many lower. The data for the poskeys in Figure 3.5 are very linear and consistent whereas Figures 3.6 and 3.7 demonstrate poskeys with decreasing and increasing rates of degradation. The r^2 and RMSE values for the poskeys and fits in these figures can be seen in Table 3.6. In the table it can be seen that despite some low r^2 values of between 0.3-0.4 when the fits are visually inspected, it can be seen that the fits do represent the data well. The spread of the geometry readings and their step nature where multiple readings would have the same vertical geometry SD then jump up or down, is the reason for the low Goodness Of Fit (GOF) results. This is demonstrated in Figure 3.7 where the one decimal place accuracy of the SDs leads to stepped data. After analysing the fits visually it was decided that despite the low r^2 and high RMSE values the fits describe the general degradation within each poskey well, with the power curve fitting the degradation trend best.

Despite the power curve fitting the data well, it has to be noted that it is not always ideal for extrapolation, with the increasing gradient fits in Figure 3.7 seeming to be the least acceptable to extrapolate. This is due to the overly steep gradient that the power curve tends to have when there are some high geometry recordings at the end of the data, which, if extrapolated would quantify the degradation as much faster than is likely. A similar problem could arise for the decreasing gradients as it is possible for the power line to be nearly flat and showing no degradation by the end of the data, which if extrapolated would result in unrealistic high times/usage to higher SD levels.

Table 3.4: GOFs of vertical top [nm] against time [days]

		Data A		Data B	
Fit	Equation	r^2	RMSE	r^2	RMSE
Linear	$Y = a * X + c$	0.5795	108.8	0.5384	124.5
Polynomial	$Y = a * X^2 + b * X + c$	0.6699	98.40	0.6487	115.4
Power	$Y = a * X^b + c$	0.6600	99.19	0.6317	117.7
Exponential	$Y = a * exp(b * X)$	0.5775	109.1	0.5376	124.8

Table 3.5: GOFs of vertical top [nm] against usage [EMGT]

		Data A		Data B	
Fit	Equation	r^2	RMSE	r^2	RMSE
Linear	$Y = a * X + c$	0.5804	108.4	0.5369	124.2
Polynomial	$Y = a * X^2 + b * X + c$	0.6707	98.07	0.6484	115.2
Power	$Y = a * X^b + c$	0.6551	99.97	0.6224	119.2
Exponential	$Y = a * exp(b * X)$	0.5779	109.0	0.5369	124.9

Table 3.6: GOFs of vertical top [nm] against time [days] for Figures TGD - fig: High R2 Degredation Fits, 3.6 and 3.7

	Linear		Power		Exponential		Polynomial	
Figure	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE
3.5 Top Left	0.9851	118.6	0.9924	86.00	0.9482	221.1	0.9926	85.24
3.5 Top Right	0.9808	35.29	0.9823	34.37	0.9803	35.71	0.9829	33.77
3.5 Bottom Left	0.9916	78.82	0.9944	65.02	0.9828	112.8	0.9947	63.55
3.5 Bottom Right	0.9917	49.93	0.9940	43.03	0.9823	72.91	0.9941	42.51
3.6 Top Left	0.3527	201.7	0.7081	136.6	0.3346	204.5	0.4962	179.5
3.6 Top Right	0.3601	82.72	0.9850	13.23	0.3484	83.47	0.6083	67.60
3.6 Bottom Left	0.3138	92.90	0.8017	51.23	0.3000	93.83	0.4895	82.20
3.6 Bottom Right	0.3596	106.4	0.6348	82.14	0.3471	107.4	0.5241	93.77
3.7 Top Left	0.3473	90.43	0.6344	68.54	0.3653	89.18	0.6259	69.34
3.7 Top Right	0.3565	95.33	0.7149	64.71	0.3757	93.89	0.6345	73.27
3.7 Bottom Left	0.3757	241.9	0.5373	211.2	0.4141	234.3	0.4876	222.2
3.7 Bottom Right	0.3670	78.94	0.4669	74.07	0.3807	78.07	0.4763	73.41

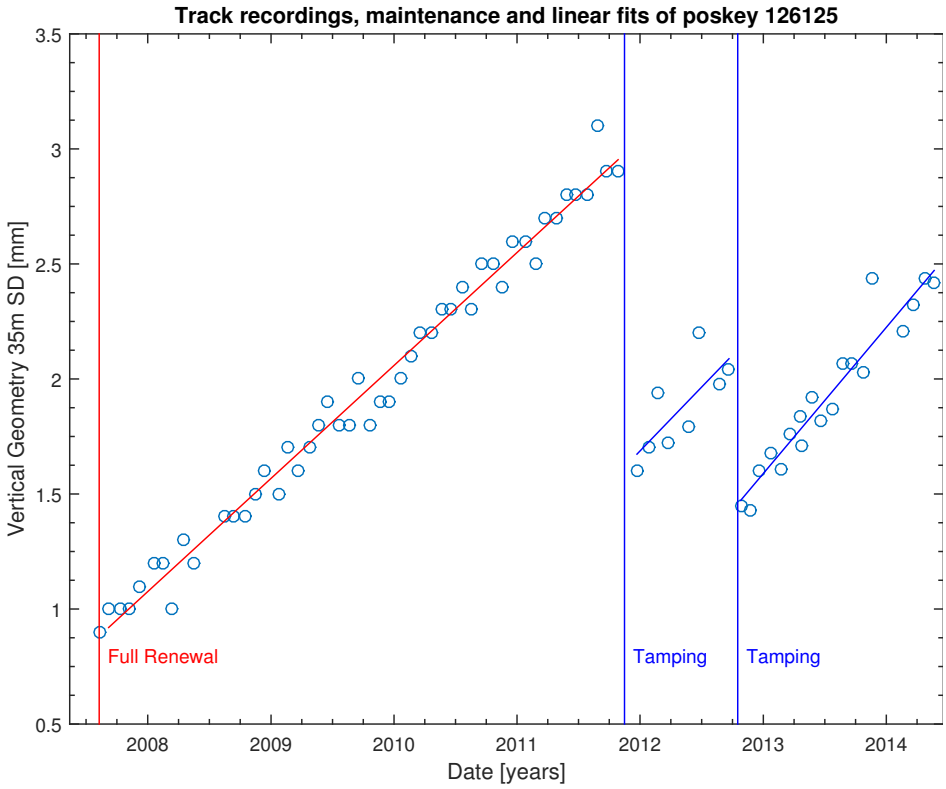


Figure 3.4: Track Recordings With Linear Fits

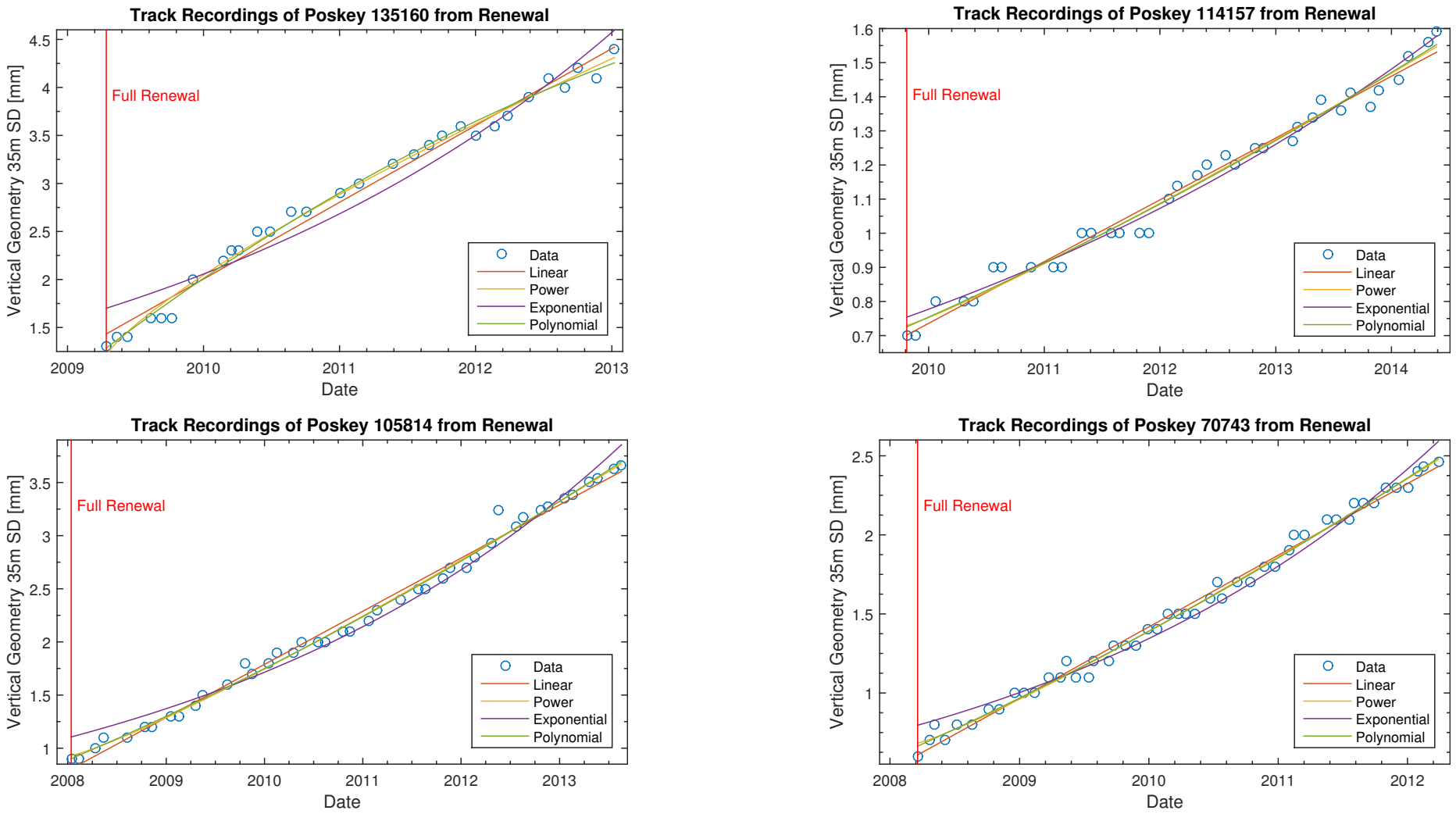


Figure 3.5: Comparison of Fits of Vertical Geometry Recordings against time (Constant Gradient)

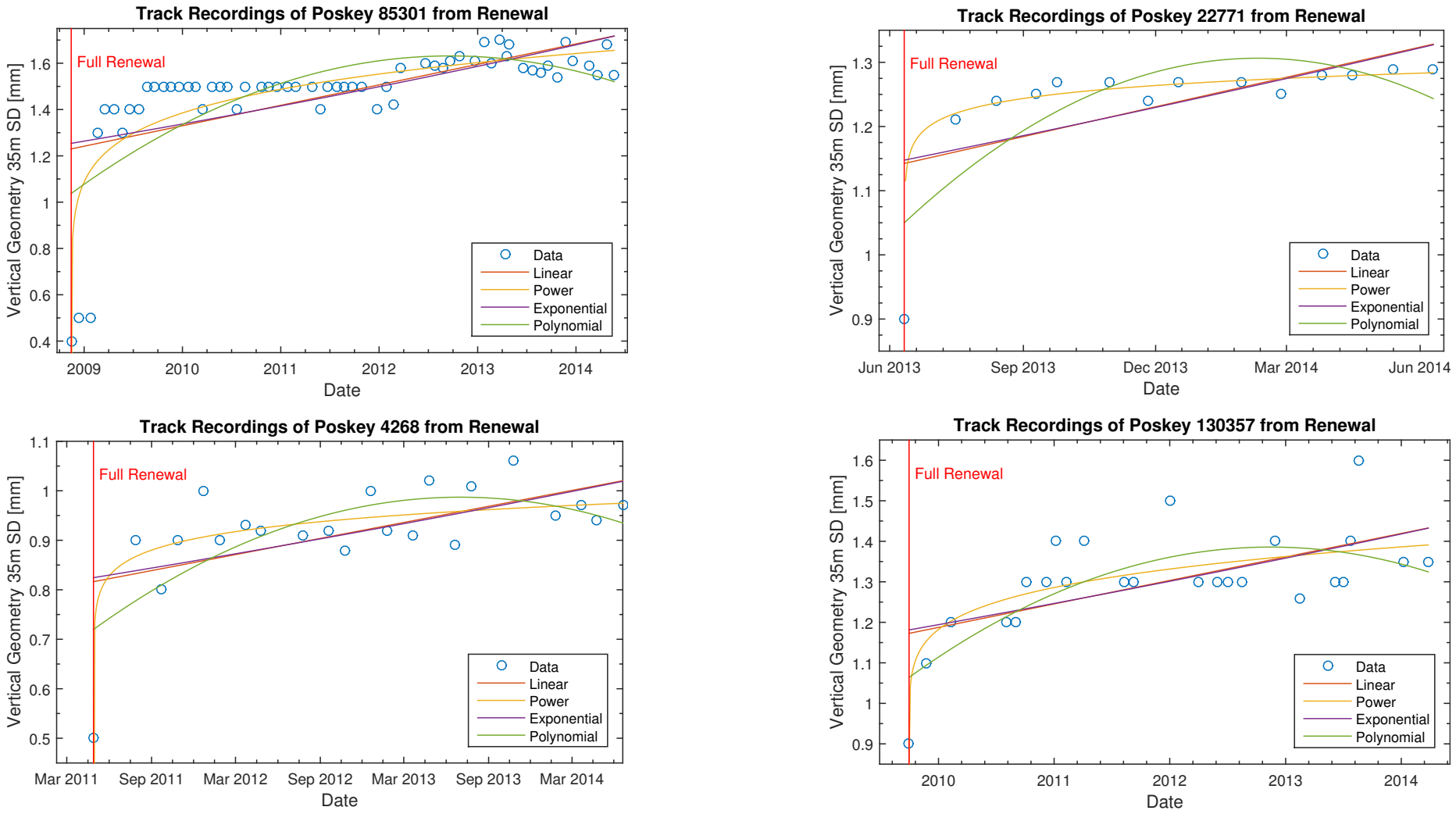


Figure 3.6: Comparison of Fits of Vertical Geometry Recordings against time (Reducing Gradient)

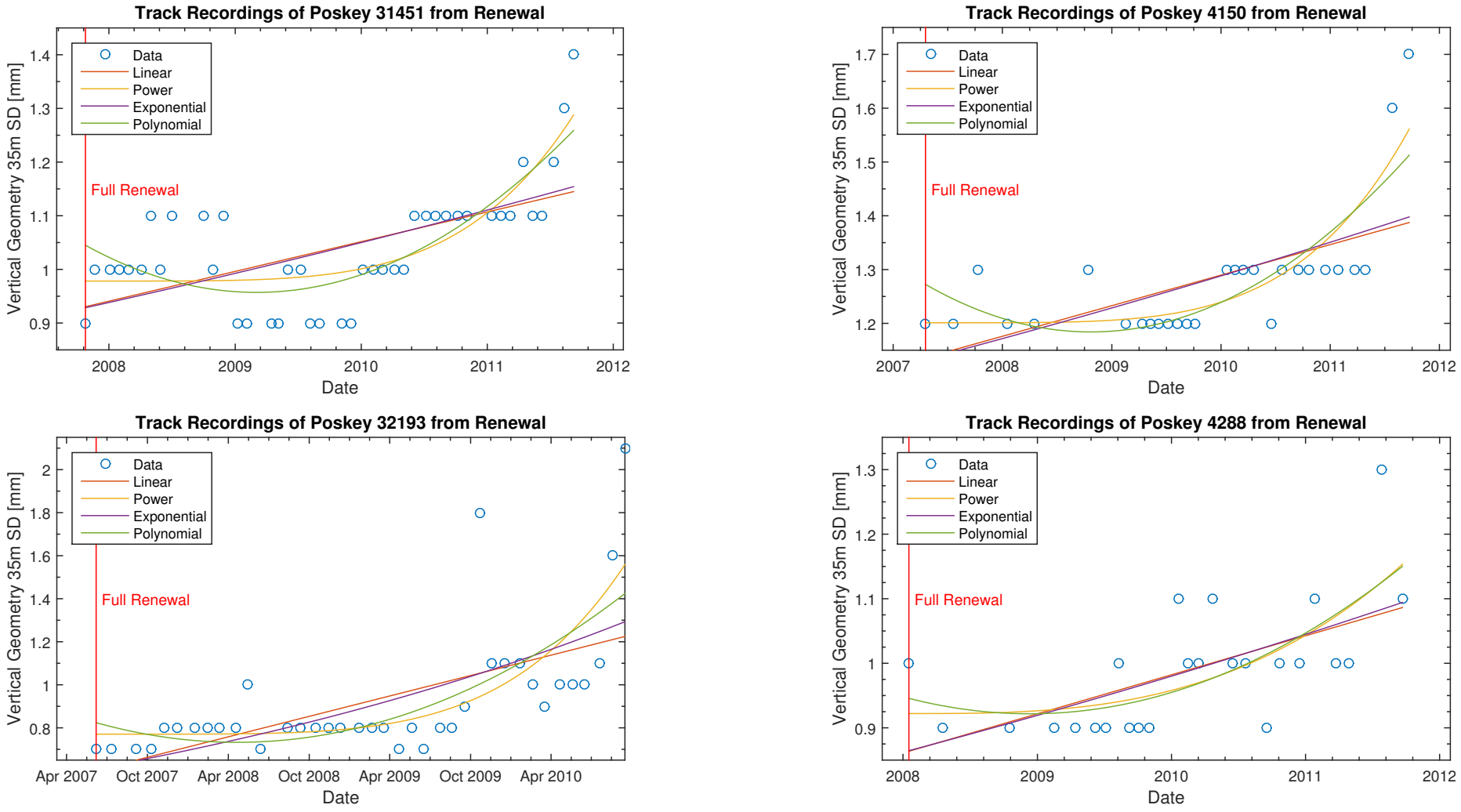


Figure 3.7: Comparison of Fits of Vertical Geometry Recordings against time (Increasing Gradient)

3.4 Degradation Analysis

3.4.1 Introduction

To understand which factors affect the rate of vertical alignment degradation, such as track speed and sleeper type, an analysis of the linear degradation rates has been undertaken. These were chosen over the statistically stronger fitting power equation as the linear fit succeeds in quantifying the general rate of degradation, as seen in Figures 3.5, 3.6 and 3.7 and discussed in Section 3.3.2, as a single number with dimensions which relate directly to the degradation. The dimensions of the gradient of the linear fits for time and usage are; nm/Day and $nm/EMGT$. The analysis was also preformed to see which options within the factors result in similar degradation, and which lead to faster/slower degradation. For example how do G44 concrete sleepers perform compared to W560 steel sleepers, at reducing vertical alignment degradation.

This section is important to help to understand the main factors that can affect the rate of vertical alignment degradation, and hence allows these to be incorporated into a model. This enables the model to include all the important factors without including additional options, that lead to a more complex model that is more difficult and time-consuming to use with no additional benefits.

As part of the analysis both the time and usage degradation rates were used, but the results obtained using the datasets of time generally showed little difference between the factors. This maybe due to track degradation being primarily caused by usage and not time, hence any significant differences within factors, such as track speed, may be masked by the disparity of usage that occurs on different track sections around the UK railway network. Due to this, and that the Petri Net (PN) model to be developed will be based on usage to enable traffic changes to be modelled, it was decided that the analysis outlined in this section will use the usage based degradation rates.

3.4.2 Method

Techniques

The analysis to find the contributing factors involved splitting up the poskeys by a chosen factor then comparing the datasets. The types of techniques that have been used include:

Descriptive Statistics

1. Amount of data.
2. Mean, minimum, maximum and range.
3. Median, quartiles and interquartile range.
4. SD, Coefficient of Variation (COV).

Visual

1. Empirical Cumulative Distribution Function (ECDF) Plots.

2. Box Plots.

Nonparametric Hypothesis Tests

1. **K-S Test (2 Sample):** Evaluates the difference between the ECDF's of the two samples, where the null hypothesis is that both samples come from a population with the same distribution. As part of this the maximum difference between the ECDF's of the two samples is used, as described in Equation 3.3 where F and G are the ECDF's of the two samples, f and g, and x is the degradation rate values, (Stephens, 1970). This is graphically shown in Figure 3.8.

$$D = sup|F(x) - G(x)| \tag{3.3}$$

It is then possible to reject the null hypothesis at a significance level, α, when $D > D_\alpha$, where D_α is based on the size of the two samples, and is calculated by:

$$D_\alpha = c_\alpha \sqrt{\frac{m+n}{m*n}} \tag{3.4}$$

Where m and n are the size of samples f and g, and c_α is based on the chosen significance level. At 0.1%, 1%, 5% and 10% c_α equals; 1.95, 1.63, 1.36 and 1.22. It is also possible to calculate a p-value by using the Inverse Kolmogorov-Smirnov distribution which follows the form of:

$$p = Q_{KS}(z) = 2 * \sum_{j=1}^{\infty} (-1)^{j-1} \exp(-2 * j^2 * z^2) \tag{3.5}$$

Where z can be calculated by Equation 3.6 and involves the found D statistic and the effective number of data points, N_e which is calculated by $N_e = (m * n)/(m + n)$, (Press et al., 2007).

$$z = D * [\sqrt{N_e} + 0.12 + 0.11/\sqrt{N_e}] \tag{3.6}$$

The KS Test involves a couple of assumptions about the data, including;

(a) Observations are independent.
(b) Responses are ordinal.
(c) Data is continuous and does not have to follow any particular distribution.
(d) Both datasets follow a similar shape and hence fit the same distribution, with similar shape parameters.

2. **Mann Whitney U Test:** Tests the null hypothesis that 2 samples come from the same population against an alternative that they come from different populations (2 tailed test) or that one sample comes from a population of larger values (1 tailed test). The test calculates the sum, for each dataset, of the amount of values in one dataset that each value in another dataset is smaller. This is known as the U statistic, with one being calculated for each dataset. The addition of the two U statistics is equal to the product of the two dataset sizes, $U_1 + U_2 = n_1 * n_2$. The U statistic can be calculated by an indirect method, where the datasets are combined and the sum of ranks

taken for each dataset. This is then used in Equation 3.7, where W_i = sum of ranks for dataset i known as the Wilcoxon rank sum statistic, and $i = 1, 2$.

$$U_i = W_i - \frac{n_i(n_i + 1)}{2} \tag{3.7}$$

For larger datasets, more than 10 samples, U, is approximately normally distributed, with $\mu = \frac{n_1 * n_2}{2}$ and $\sigma = \sqrt{\frac{n_1 * n_2 (n_1 + n_2 + 1)}{12}}$. With these it is possible to find the z-score by $z = \frac{U - \mu}{\sigma}$, and hence now possible to obtain a p-value, which at a 5% significance, rejects the null hypothesis if below 0.05. The Mann Whitney U Test has a couple of assumptions about the data, including;

(a) Observations are independent.
(b) Responses are ordinal.
(c) Data is continuous and does not have to follow any particular distribution.
(d) Both datasets follow a similar shape and hence fit the same distribution, with similar shape parameters.
(e) The datasets being compared have similar variances.

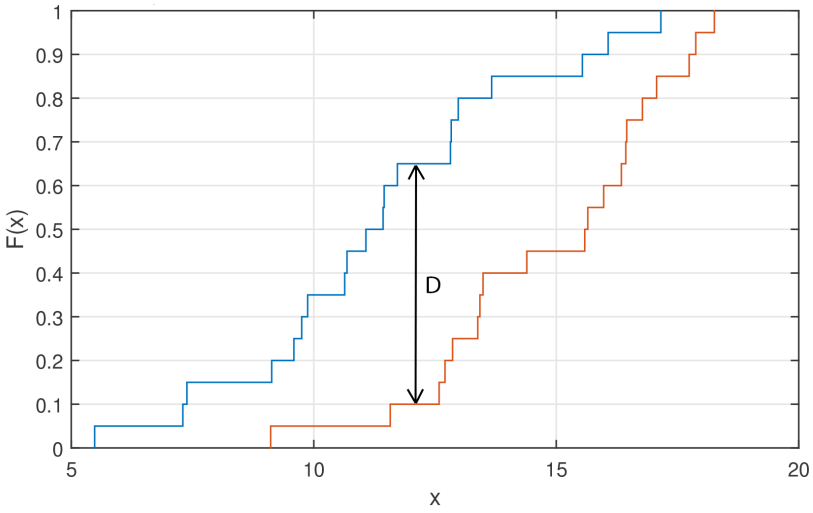


Figure 3.8: Example of the KS-Test D Value

Process

As part of the analysis, many possible factors have been explored with the options within these factors having been grouped if they were statistically similar. This was done as it helped maintain larger datasets, which allow for greater confidence in the statistical results of the used tests and the ability to split the datasets up further to explore more factors. Joining together similar options within a factor will also reduce the complexity of the chosen PN type model by reducing the amount of distributions required for the transitions, without decreasing the model's accuracy. Additionally, the distributions found will be constructed with larger datasets the less the data is split up, increasing

the confidence in the fits found.

A script was written within the statistical software Matlab which split up all the maintenance history datasets within Data A and B by a factor such as track speed, where different speeds create unique datasets. This leads to multiple datasets within each maintenance history for Data A and B. These datasets were then compared using the methods discussed. The non-parametric hypothesis tests consisting of 2 tailed K-S Test and Mann Whitney, where used to compared between the different options of the chosen factor within each maintenance history. The 2-sample K-S and Mann Whitney tests where both used due to the difference in their calculations leading them to have differing shortcomings. The Mann Whitney test, due to being calculated by ranks, is highly dependent on the difference between the medians, whereas the K-S test is based around the largest difference between the datasets which may not be at the medians. Both tests are reasonably poor at differences within the tails, but due to the data being analysed being degradation data, the tails are generally extreme values that are not representative of the normal degradation, so this actually helps ignore outliers. It is important for both tests, especially the Mann-Whitney tests that the compared dataset are similar in shape, so ECDF's were used to visually check that the datasets being compared followed similar shapes without crossovers.

The non-parametric hypothesis tests resulted in K-S Test and Mann Whitney p-values and hypothesis test values, between each option in each maintenance history. An average was then found over all the different maintenance history's, giving one mean p-value and mean hypothesis tests result comparing one factor option and another. The mean was used instead of a median so extreme p-values were accounted for in the resultant number. The mean results from the tests leads to tables such as Tables 3.7 and 3.8 which were obtained using Data A. In the mean p-values table, those with values below 0.05 (coloured red) reject the null hypothesis at a 5% significance level and hence are statistically different datasets. As a statement does not simply become true with p-values below 0.05 and false when above, it is important to look at the p-value Grabowski (2016). A p-value of 0.05 is a significant result, with a low chance of rejecting a true hypothesis (type-1 error). Due to the analysis being undertaken, weaker links between the datasets are still desirable to identify, with higher type-1 errors acceptable for the reduction in type-2 errors (accepting a false hypothesis). It was decided that p-value below 0.2 and above 0.05 (coloured yellow) do not reject the null hypothesis, but they also do not strongly agree that the two datasets are from the same population. The p-values above 0.2 (coloured green) strongly accept the null hypothesis, meaning that it is likely that the datasets have similar distributions and come from the same population.

The mean hypothesis tests results in Table 3.8 can be seen as the percentage of the comparisons which reject the null hypothesis at the chosen significance level, where 0 (coloured in green) means 0% of the comparisons over all the maintenance history's rejected the null hypothesis and 1 (coloured in red) means 100% of the comparisons rejected the null hypothesis. The results between 0-1 (coloured orange and yellow) occur

when some comparisons within the maintenance histories rejected the null hypothesis and some did not, where yellow has been used when the majority accepts the null hypothesis and orange when they reject. It can be noted that the p-values equal 1 on the diagonal as these compare the same dataset against itself and hence accept the null hypothesis.

As From the results seen in Tables 3.7 and 3.8, it can be seen that the poskeys with speeds between 5-60 MPH have similar vertical geometry degradation due to the high p-values and hypothesis tests significance results of zero (no test rejected the null hypothesis). Poskeys with speeds between 75-110 MPH also degrade similarly, but due to the lower p-values than the 5-60 MPH group, and average hypothesis tests result above 0 (but still less than or equal to 0.5 (50% of tests rejecting the null hypothesis)), they are not statistically as similar. It was desirable for the analysis approach to minimise groups, to maintain larger datasets (reducing type-2 errors). Due to this the decision was made to join groups when the majority of tests accepted the null hypothesis and the average p-value was greater than 0.2. The low p-values compared to other speeds of groups 65-70 and 115-125 MPH, show that these groups are distinct. So it can be said that the track with speeds between 115-125 MPH degrades at a different rate than other speed groups. The speed groups have been based on NR policy documents. From these results it was decided to split all the speeds into four groups, consisting of 5-60, 65-70, 75-110 and 115-125 MPH. Again the non-parametric tests were used to see if these groups where similar. These results can be seen in Tables 3.9 and 3.10, which demonstrate that the chosen groups are all statistically unique. Further analysis including the results obtained using Data B can be seen Section 3.4.3.

One tailed tests were then used to see at which speeds poskeys degrade faster. The results are presented in Tables 3.11 and 3.12. For one tailed tests at a 5% significance, p-values below 0.05 (red in the tables) reject the null hypothesis in favour of the alternative that the population mean of the first dataset is higher than the second. Where the two sided test allows for an understanding of which groups populations are significantly different (or similar), one tailed allow us to identify which groups degrade faster or slower. This is read in the tables as column against row, so the p-value of 0.00 where the column 5-60 MPH meets the row 65-70 MPH, means that the poskeys with speeds 5-60 MPH statistically have a distribution consisting of higher values, and hence degrade quicker. It is possible to see from the results in the tables that the faster the track the slower the tracks vertical geometry degrades. This can also be seen visually in Figures 3.10 and 3.11, which show ECDF's and box plots comparing the datasets of different speeds, with different maintenance history's. The figures include both data from Data A and B, which have been kept separate but both used to see if they agree, which they do as seen in the figures. The box plots include notches around the median which are equal in size to:

$$\text{Notch Locations} = q_2 \pm \frac{1.57 * (q_3 - q_1)}{\sqrt{n}} \tag{3.8}$$

Where q_2 is the median, q_1 and q_3 are the 25^{th} and 75^{th} percentiles and n is the size of the dataset. It can be said that the medians of the datasets are significantly different at a 5% significance level if the notches do not overlap. The whiskers of the box plots extend to:

$$\text{Whisker Location Upper } = q_3 + 1.5 * (q_3 - q_1) \tag{3.9}$$

$$\text{Whisker Location Lower } = q_1 - 1.5 * (q_3 - q_1) \tag{3.10}$$

With any data points outside of this being marked with a red x. The whisker locations are equivalent to ±2.7 SDs and 99.3% coverage of the data.

Despite the analysis in Tables 3.11 and 3.12 showing that higher speed track degrades slower, this is not what would be expected. Higher speed trains exert higher dynamic forces which would lead to faster degradation. There are many differences between the different speed tracks not included in the results, which could be the root cause of the results and not the track speed. The higher speed tracks in general have dissimilar types of sleepers and rails to the lower speed track, as well as more frequent maintenance due to lower tolerances of vertical alignment SD before the track is deemed unsafe for the speed. This issue is the same for any first level analysis, and a major problem of many investigations into causes of degradation discussed in the literature review. To reduce the chance of addition factors affecting the results of the chosen factor to explore, the data has to be split up into more levels, such as seen in Figure 3.9. In this Figure the data is first split up by the maintenance history and then speed as before, using the grouped speeds from the earlier analysis discussed above to help maintain larger datasets. These datasets are then split up further by the sleeper type, leading to a comparison of different sleeper types where the maintenance history is the same and track speeds are similar. As before, the non-parametric tests will be used to compare the datasets within the same maintenance history and speed group and then an average found. The analysis would start with sleeper type split into low level groups. The results of the analysis are then used to combine the sleeper types into larger groups, to be analysed again (now that there are larger datasets, with evidence that all the data within each group have similarly performing sleeper types). The approach of splitting data by many factors can quickly lead to very small datasets (less that 10), reducing the confidence in the analytical results. This is why the grouping of similarly performing options in each degradation factor is important.

Table 3.7: Mean 2 Tailed K-S and Mann Whitney U Test p-Values for Maximum Permissible Track Speeds (Data A, Usage)

Track Speed [MPH]	5-30	35-40	45-50	55-60	65-70	75-80	85-95	100-110	115-125
5-30	1.00	0.41	0.31	0.36	0.71	0.06	0.00	0.00	0.00
35-40	0.41	1.00	0.61	0.59	0.22	0.00	0.00	0.00	0.00
45-50	0.31	0.61	1.00	0.78	0.05	0.00	0.00	0.00	0.00
55-60	0.36	0.59	0.78	1.00	0.00	0.00	0.00	0.00	0.00
65-70	0.71	0.22	0.05	0.00	1.00	0.20	0.19	0.31	0.00
75-80	0.06	0.00	0.00	0.00	0.20	1.00	0.30	0.46	0.03
85-95	0.00	0.00	0.00	0.00	0.19	0.30	1.00	0.29	0.05
100-110	0.00	0.00	0.00	0.00	0.31	0.46	0.29	1.00	0.01
115-125	0.00	0.00	0.00	0.00	0.00	0.03	0.05	0.01	1.00

Table 3.8: 2 Tailed K-S and Mann Whitney U Tests null hypothesis rejection decimal percent (alpha = 0.05), for Maximum Permissible Track Speeds (Data A, Usage)

Track Speed [MPH]	5-30	35-40	45-50	55-60	65-70	75-80	85-95	100-110	115-125
25-30	0.00	0.00	0.00	0.00	0.00	0.50	1.00	1.00	1.00
35-40	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00
45-50	0.00	0.00	0.00	0.00	0.50	1.00	1.00	1.00	1.00
55-60	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00
65-70	0.00	0.00	0.50	1.00	0.00	0.33	0.67	0.33	1.00
75-80	0.50	1.00	1.00	1.00	0.33	0.00	0.38	0.25	0.83
85-95	1.00	1.00	1.00	1.00	0.67	0.38	0.00	0.25	0.67
100-110	1.00	1.00	1.00	1.00	0.33	0.25	0.25	0.00	1.00
115-125	1.00	1.00	1.00	1.00	1.00	0.83	0.67	1.00	0.00

Table 3.9: Mean 2 Tailed K-S and Mann Whitney U Test p-Values for Grouped Speeds (Data A, Usage)

Track Speed [MPH]	5-60	65-70	75-110	115-125
5-60	1.00	0.06	0.00	0.00
65-70	0.06	1.00	0.06	0.00
75-110	0.00	0.06	1.00	0.05
115-125	0.00	0.00	0.05	1.00

Table 3.10: 2 Tailed K-S and Mann Whitney U Tests null hypothesis rejection decimal percent (alpha = 0.05), for Grouped Speeds (Data A, Usage)

Track Speed [MPH]	5-60	65-70	75-110	115-125
5-60	0.00	0.83	1.00	1.00
65-70	0.83	0.00	0.80	1.00
75-110	1.00	0.80	0.00	0.82
115-125	1.00	1.00	0.82	0.00

Table 3.11: Mean 1 Tailed K-S and Mann Whitney U Test p-Values for Grouped Speeds (Data A, Usage)

Track Speed [MPH]	5-60	65-70	75-110	115-125
5-60	N/A	1.00	1.00	1.00
65-70	0.00	N/A	0.81	0.99
75-110	0.00	0.09	N/A	0.97
115-125	0.00	0.00	0.05	N/A

Table 3.12: 1 Tailed K-S and Mann Whitney U Tests null hypothesis rejection decimal percent (alpha = 0.05), for Grouped Speeds (Data A, Usage)

Track Speed [MPH]	5-60	65-70	75-110	115-125
5-60	N/A	0.00	0.00	0.00
65-70	1.00	N/A	0.00	0.00
75-110	1.00	0.67	N/A	0.00
115-125	1.00	1.00	0.75	N/A

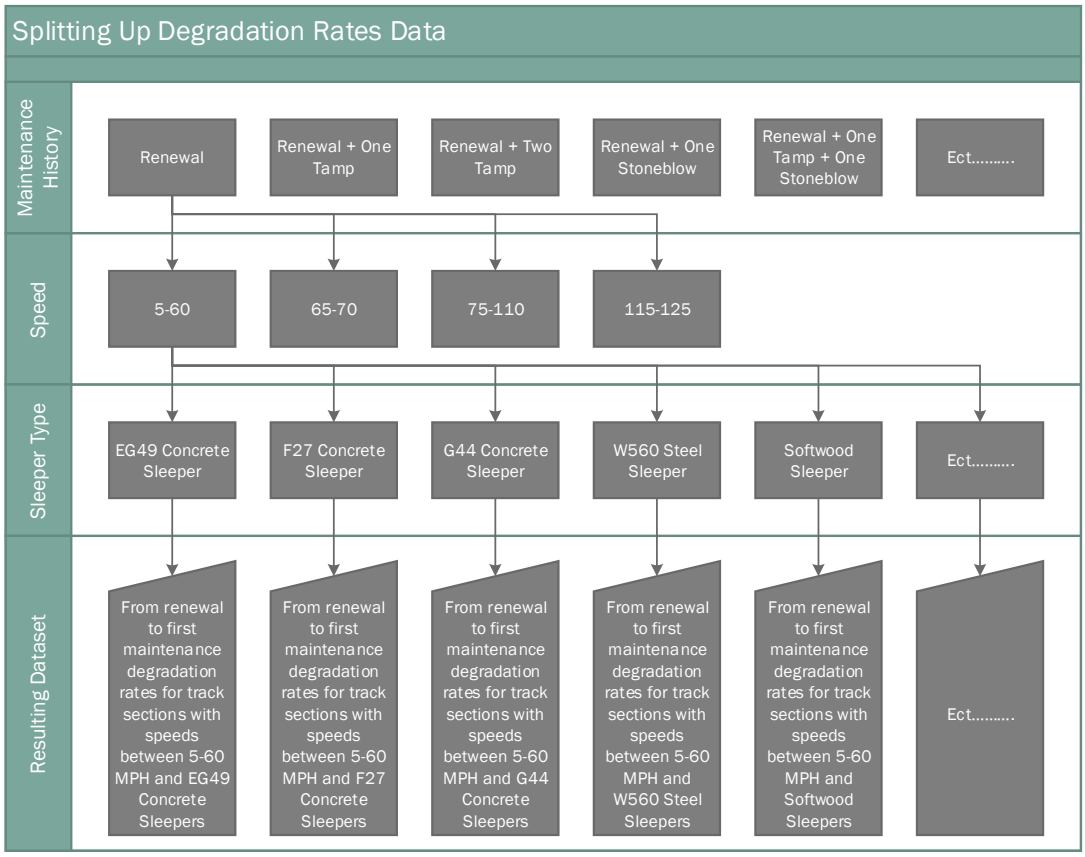


Figure 3.9: Method of Splitting up the Degradation Rates Data

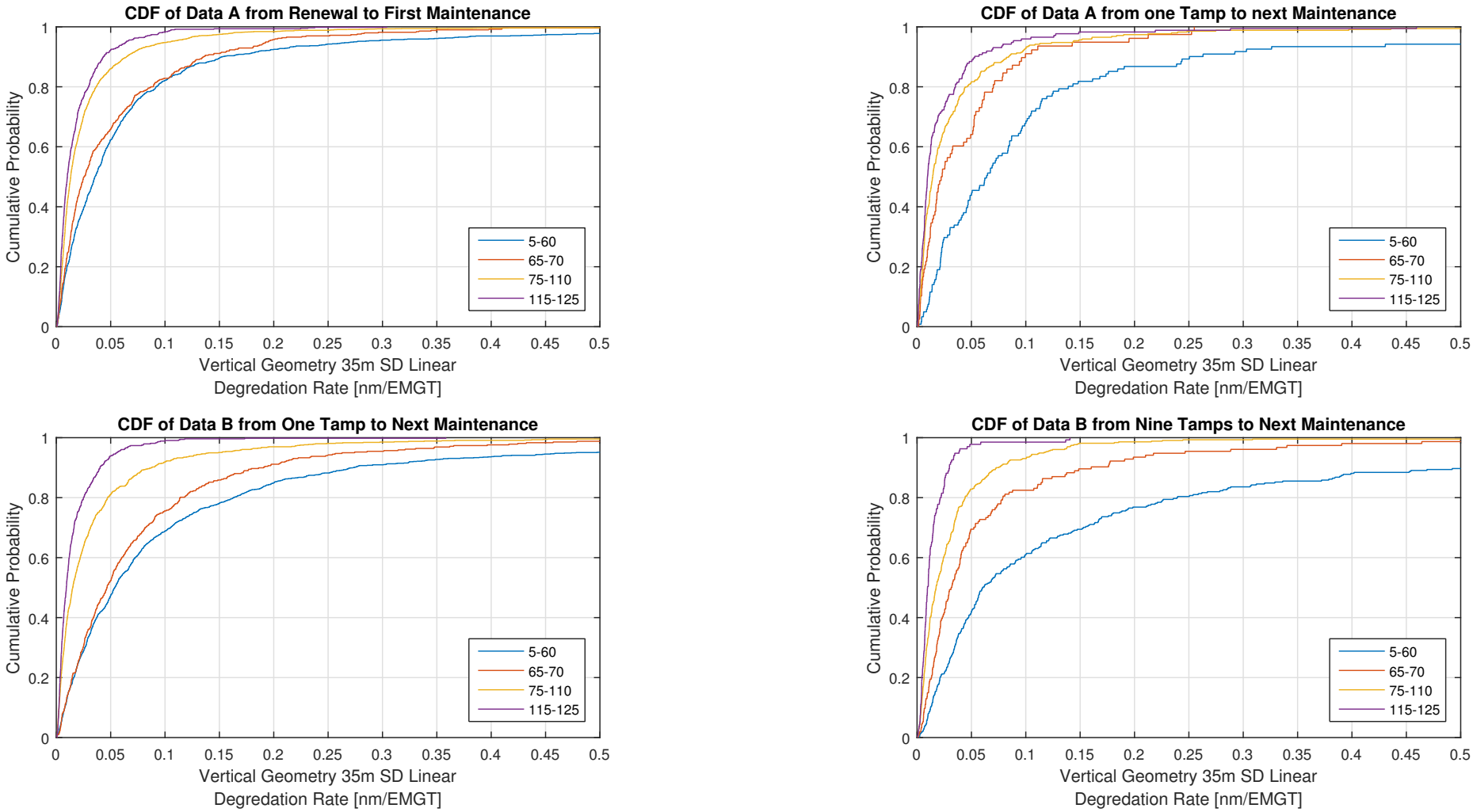


Figure 3.10: Comparison of ECDF's of different speed groups, with differing maintenance history's

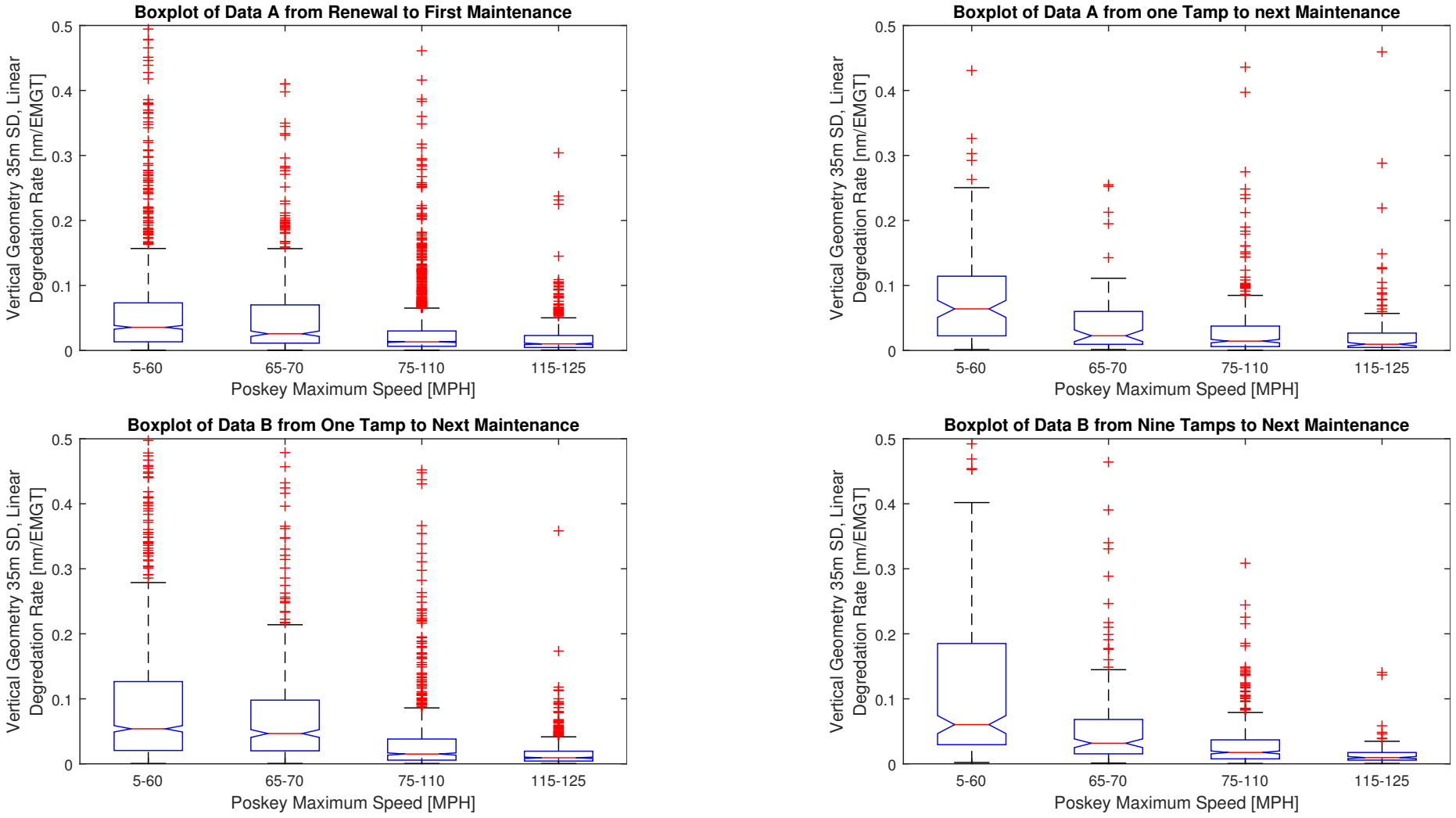


Figure 3.11: Comparison of Boxplots of different speed groups, with differing maintenance history's

3.4.3 Analysis Results

As part of the analysis many factors have been considered, including asset choices, line usage and geology under the track. For each factor a first layer analysis consisting of only splitting up the data once by the one factor, such as with the maximum speed example demonstrated in Section 3.4.2, has been undertaken. This first layer analysis includes the all the data relating to each option in the factor, generally resulting in large datasets and gives a general idea of the factors to concentration on for the further analysis involving more layers. It is important to note that the results from the first layer analysis should not be taken at face value as there may be other reasons for the statistical differences other than the factor being explored. Depending on the results of the first layer analysis, further second and third layer analyses have been undertaken splitting up the datasets by additional factors. These were decided from information researched as part of the literature review, as well as trends within the data, such as the different sleeper types used on varying track speeds. Within this section the resultant statistics, in particular the non-parametric tests, from comparing the split up datasets have been analysed. These results have been summarised in Section 3.4.4 and published in Clarke and Prescott (2017).

General Statistics

To assess the general values and spread found within the datasets, means and Coefficient of Variation (COV) have been found. The COV was used to determine the spread of the data as it is dimensionless making it possible to compare the spread of the time and usage datasets. COV is defined as the ratio between the standard deviation and mean of a dataset, $COV = \sigma/\mu$. The values obtained can be seen in Table 3.13, in which the mean degradation rate values for time and usage have the units of nm/Day and $nm/EMGT$. The mean values show that average degradation rate within Data B is higher for usage, which was expected due to Data B containing older track sections, which are expected to degrade faster. The time data shows the opposite, that the mean rate of degradation is higher in Data A. Looking at the COV values within the table, it can be seen that the spread of the whole time datasets are smaller than the usage datasets (COV of 1.6 and 1.3 vs 2.0 and 2.2). Looking at the Table, when the dataset of vertical geometry degradation rates is split by the maintenance history and then the COV of each remaining dataset found. The mean of the these are lower than the spread in the whole dataset as expected. When split by maintenance history, the usage datasets show a larger decrease in spread than the time datasets. This shows that the high spread found within the whole dataset is due to a larger difference between the maintenance history groups than is found with the time data. By further splitting up the datasets by the track sections speed using the ten groups used in Table 3.7, last column, the spread is reduced further. The mean values obtained show that the usage results from Data A has the smallest spread, with a small spread being desirable for a stochastic model as it reduces the uncertainty in the results of the model. The results show that in general the

spread found in Data B tends to be a lot higher than Data A. This is probably due to the added uncertainty of the estimated amount of previous maintenance actions for these track sections. Additionally, it was expected that older track would have a higher degree of spread. The more distinct maintenance history datasets within the usage results and lower spread when split up, increases the confidence in choosing usage as the driving degradation factor and not time. The much higher spread results obtained using Data B also demonstrate why Data A and Data B have been kept separate.

Table 3.13: General statistics of the degradation rate datasets

Dataset	Mean	COVs Whole Dataset	Mean of COVs of Dataset split by the maintenance history	Mean of COVs of Dataset split by the maintenance history and the track speed
Data A - Time	0.8758	1.6279	0.9630	0.6396
Data B - Time	0.7389	1.3789	1.1638	0.8063
Data A - Usage	0.0643	1.9942	0.8029	0.5862
Data B - Usage	0.0972	2.2212	1.6665	1.4100

Stations and Tunnels

Track sections installed at stations undergo different loading than other sections, due to the trains decelerating, waiting and accelerating. To test if this affects the degradation rate, the poskeys were split into groups, one containing no stations and the other containing poskeys that are entirely or partly stations. Including poskeys which were only partly stations was decided as leading up to and straight after a station where the main decelerating and accelerating occurs will also have different loads to normal track away from stations. Additionally, as there is only a small amount of poskeys which are completely within a station there would not have been enough data to perform an analysis. The amount of data used can be seen in Table 8.5, where for a dataset to be used it was decided that it should include a minimum of twenty degradation rate values. Each dataset is created by first splitting up the whole degradation rate dataset by the maintenance history and then by the existence of a station. Each found degradation rate is its own data point, so each poskey can have many data points for between different maintenance histories (rate between renewal and first tamp, first tamp and second etc.). It was decided that more than twenty degradation rate values were required in each of the made datasets as the 2 sample KS-Test compares between the ECDF's, and datasets that contain less than twenty pieces of data do not tend to create a distribution that would be representative of a population as there is not enough points to form a shape which is not highly swayed by extreme values. The total amount of degradation rate values that made up these datasets are also noted within the table. Furthermore, it

was decided to include tunnels into the analysis by splitting up the non-station dataset into tunnels and neither a station nor tunnel. Tunnels were classified in the same way as stations with poskeys that contained part of a tunnel being recorded in the tunnel group. This was done due to the lack of tunnel data, as seen in Table 8.5, where even with partly tunnel poskeys being included the amount of data available was still low. It was thought that tunnels would degrade differently as the track encounters dissimilar environmental condition, where within a tunnel, there is no precipitation or direct sunshine. Additionally, the temperature in tunnels do not fluctuate as much across the year compared to outside track, resulting in very low chances of frost. Due to these it would be expected that track in tunnels degrades slower but there are also issues with track inside tunnel. Fines from the tunnel roof will enter the ballast in higher quantities than general outside environmental fines, causing faster ballast fouling. Also, due to height restrictions the ballast bed might be thinner, but there is no data on the ballast thickness's to check this. Maintenance activities are also harder to complete in tunnels due to space restrictions and lack of natural light. The last factor which may cause track in tunnels to degrade faster is the drainage, which tends to be less effective within tunnels and hence will lead to the subgrade and ballast having higher levels of saturation reducing their resilience to deformation, increasing settlement and geometry degradation. It was not possible to identify any issue that drainage make cause as there was no data available of the type of or conditions of the drainage around the UK railway system.

To test for any significant difference in the degradation rate of the vertical geometry between track sections installed at stations, within tunnels and neither, 2-sided hypothesis tests were used. The amount of data used in the tests is shown in Table 8.5, where it is apparent that the results from Data A may not give a good representation of any general difference between the groups due to minimal data for track within stations or tunnels. This is because only two datasets contained over twenty values for both; from renewal to the first maintenance and after one tamping operation, causing the resultant mean p-value results and rejection percentages to only be based on two comparisons and contain no information on older track which has undergone more maintenance actions. The results obtained using Data A are recorded in Tables 8.6 and 8.7. Comparing stations to neither, the mean p-value of 0.39 and a null hypothesis rejection percentage of 0%, show that there is no statistically significant difference between track at stations and not at stations. More significant results are seen when analysing track contained within a tunnel. The results from Data A show that there is a significant difference between tunnels and stations or neither with mean p-values of 0.05 and 0.03 and associated rejection percentages of 75%. The results for Data B in Tables 8.8 and 8.9 weakly agree with the results from Data A for the stations with a mean p-value of 0.30 but show much more significance difference with 40% of the tests rejecting the null hypothesis in favour of track within stations degrading significantly different from track between stations. The results for tunnels disagree with the conclusions obtained from Data A showing less significant differences, with mean p-values of 0.35 and 0.26 and rejection percentages of 33%. The lack of data which was available within Data A to perform the tests reduces

the confidence in the results. Due to this the results obtained from Data B, despite the extra uncertainty within Data B, are more reliable for categorising any difference between the groups. Due to the results seen in the tables it cannot be said that there is a significant difference of the rate of vertical geometry gradation between track within stations, tunnels or neither but there are differences as demonstrated by the rejection percentages.

Despite the 2-sided test results not showing a significant difference between the datasets, 1-tailed tests are still useful to give an idea of which group degrades faster even if it is not significantly faster. The results can be seen in Tables 8.10 and 8.11 for Data A and Tables 8.12 and 8.13 for Data B respectively. The mean p-value of 0.20 for Data A gives evidence that track at station degrades faster but not significantly so than track at neither a station nor within a tunnel. However none of the tests undertaken demonstrated a significance at a 5% confidence level (p-values less than 0.05), hence the 0% rejection percentage between neither and stations. Due to this there is not enough evidence to say one degrades faster than the other. The results for track within tunnels show that despite the significant differences found in the two-sided tests, it cannot be said that track within tunnels degrades faster or slower. This is seen by the mean p-values all being close to 0.5 and that the rejection percentages show that tunnels degrade significantly faster in 50% of the tests but the percentages showing that neither nor stations degrade significantly faster than tunnels was 25% and 50%. So some tests showed that tunnels degraded significantly faster and other results demonstrated that tunnels actually degrade significantly slower.

For additional evidence Data B was used. The results obtained from the tests using Data B gives more evidence to track at stations degrading significantly faster than track not at stations, with 40% of the tests rejecting the null hypothesis. The results for tunnels give evidence to suggest that track in tunnels tends to degrade slower, with 25% of the tests rejecting the null hypothesis in favour of neither having higher degradation rates and 42% in favour of stations. Only 17% of the test results returned showed tunnels degrading faster than neither and none of the tests showed that tunnels degraded significantly faster than stations. It must be noted that out of all the hypothesis tests performed across all the investigated factors in this chapter that the ECDF's shapes for tunnels varied the most from the comparable ECDF's of stations and neither. This is due to the extra variability in the degradation rates for track within a tunnel, with higher variance values found. This can cause issues in the results obtained from the hypothesis tests as noted in section 3.4.2 and hence the results were not entirely trustworthy. Due to this and the results of the tests it can be said that there are differences in the rate of track degradation within stations, tunnels or neither but these in general are not that significant at a 5% confidence level. It can also be said that track at stations tend to degrade slightly faster than track between stations.

The results of the 1-sided tests for tunnels are not conclusive with some evidence supporting the conclusion that track in tunnels tends to degrade slower. The more reliable

conclusion is that the degradation rate of track within tunnels tends to vary more than track not within tunnels, sometimes being higher sometimes being lower, depending on the tunnel. This makes sense due to the factors discussed within this section like more variable depth of ballast and drainage quality but due to a lack of data these cannot be tested. The length of the tunnel may also play a factor, due to the differences in internal environment further underground compared to the entrances of the tunnels, but by splitting the tunnel data further depending on the distance from the tunnel entrances created datasets too small for a detailed analysis to be undertaken. It would have been desirable to perform a second layer analysis to help identify the possible reasons for the disparity of degradation rates within the tunnel data, and to remove uncertainty in the results being due to another underling factor but there was not enough data to split it up further. Moreover, as tunnels and stations occur all over the network it can be assumed that there would also be an even amount of asset choice and subgrade geology type spread in the three factor options, reducing the chance these would be affecting the results.

Track Type

The track type describes the method used to join the rail lengths together. This consists of either jointed track which uses plates and bolts or Continuous Welded Rail (CWR) where long sections are welded together on site. Jointed track is a legacy design which is no longer used, hence in Table 8.14, it can be seen that there is no data for jointed track within the recently renewed poskeys in Data A. Due to this only Data B was used to analyse the affect of the joint type on the vertical geometry degradation. The results from the 2-tailed hypothesis tests, Tables 8.15 and 8.16, show that there is strong statistical evidence to suggest that jointed track degrades differently from CWR, with all the tests returning p-values of zero hence rejecting the null hypothesis that they are similar.

Looking at the 1-tailed tests in Tables 8.17 and 8.18, it can be seen that all the tests rejected the null hypothesis, showing that jointed track's vertical geometry degrades faster than CWR. This result maybe due to additional aspects though such as the track speed as jointed rail tends to only be on the slower speed lines, with the faster higher priority lines all having been converted to CWR. Due to this the results might be due to the line speed and not the join type.

To further investigate the effect of the joint type on the geometry degradation a second layer analysis had to be performed to take into account the track speed first and then the joint type. The track speed was chosen as the extra factor as track with similar speeds tends to be built with comparable designs with akin asset choices. Additionally, they tend to have interchangeable maintenance policies and also similar dynamic loading caused by the speed of the trains. The hypothesis tests were performed again by first splitting the degradation rates up first by the poskeys maintenance history and then the speed, in the four groups identified in Section 3.4.3, then either CWR or jointed. As

part of this the tests were only performed between datasets within the same maintenance history and speed group. This ensures that other than the track type the groups of poskey degradation rates would have minimal other differences. The results from these tests showed that no track over 60 MPH still had jointed rail installed and there was not enough data within Data A to perform the analysis due to jointed rail not being installed in newly renewed track any more. So the comparison tests were only performed on track with speeds less than 60 MPH using Data B. The results from the tests again showed a very significant difference between CWR and jointed track, with the 1-sided tests showing that 100% of those performed rejected the null hypothesis in favour of jointed track degrading significantly faster with a mean p-value of 0.00. An additional check was performed to see if rail the rail type was affecting the results in any way. This was performed by first splitting the datasets up by speed, then rail type and then track type. The results from the hypothesis tests again showed a highly significant difference between jointed and CWR track with a mean p-value of 0.00.

Track Construction

Due to there being many track criticality's on the UK rail network not all track it built to the same standards. This enables more critical track to be built to a higher standard and the less to a lower standard, efficiently using the available capital for track construction. The quality levels are defined as A, B, C and D, where A is the highest standard used on higher priority track. Due to better working practices and that most recent renewals have occurred on the high priority lines there are no track sections within Data A in band D and minimal in C, as seen in Table 8.19. Due to there only being one dataset for construction type C in Data A, this type was not included in the analysis. The results from the hypothesis tests can be seen in Tables 8.20, 8.21, 8.22 and 8.23. In these it can be seen that the mean p-value of 0.21 for Data A suggests that construction standards A and B perform similarly but when the pass/reject results are taken into account it can be seen that 75% of the tests rejected the null hypothesis. Using Data B as more evidence, a mean p-value of 0.08 suggests that there is a difference between the construction bands, and similarly to Data A, 67% of the tests rejected the null hypothesis. The other p-values within Table 8.22 show a significant difference between the rest of the construction bands, which is back up with between 90-100% of the tests rejecting the null hypothesis in Table 8.23. Due to the obtained results it can be said that there is a statistically significant difference of the degradation rates between all the different track construction standards, with A and B being closest but still statistically dissimilar.

The 1-tailed tests add to the evidence that the construction groups degrade differently. Tables 8.24 and 8.25 show the results from Data A. In these tables it can be seen by the mean p-value of 0.11 and 75% rejection that there is evidence to suggest that construction quality band B degrades faster than A. This is further cemented by Data B's results, seen in Tables 8.26 and 8.27, with a mean p-value of 0.06 and rejection percentage of 80%.

Data B's mean p-values of zero in column D show that track constructed to quality band D degrades significantly faster than A, B and C, also there is evidence to suggest band C degrades significantly faster than A and B. The results indicate that higher construction bands degrade slower, which was the expected results as these would have been built to a higher standard.

Despite the strong evidence, it was thought that there maybe an underlying factor that is the reason for the results. As the higher construction bands are primarily used for high priority fast track, the results might have been affected by a large amount of high speed track sections in the higher band and primarily slower track in the lower bands. Due to this a further analysis was undertaken to first take into account the track speed and then compare the construction bands. By taking into account the track speed as well as the maintenance history, comparisons between the construction groups will only occur between track sections that contain similar sleeper and rail types, maintenance policies and dynamic loading caused by the speed. The speed was grouped into the four groups of similar degrading speeds discussed in Section 3.4.3. The results from this showed that the 5-60, 65-70 and 75-110 MPH groups contained mostly construction band B and track speeds between 115-125 MPH are mostly built on construction band A. Construction band B is the most common but a higher percentage of newer track is built to band A standard. There is no track of band C with track speeds above 110 MPH and none of band D above 60 MPH. The results of the hypothesis tests showed similar differences to the first layer tests in Tables 8.20, 8.21, 8.22, 8.23, 8.24, 8.25, 8.26 and 8.27. The results from Data A showed an even more significant difference between bands A and B with B always being larger. Data B showed less significant results but still showed that the better quality construction bands encounter slower degradation with the 1-tailed test results never rejected the null hypothesis in favour the better quality construction bands performing worse. The results were less significant as the rejection percentages of the lower construction groups in the one-tailed tests varied between 0.55 and 1.00 whereas the first layer analysis returned results between 0.80 and 1.00. Due to the results obtained from both the first layer and second layer analysis it can be said with confidence that the better the quality of the track construction the slower the track will degrade, with similar increases in the degradation rates occurring as you step down the construction bands.

Track Category

The track category is related to the usage and speed of the track, with 1A being fast heavily traversed lines and 6 being slow with minimal traffic, as seen in Figure 3.1. The amount of available data for the categories is shown in Table 8.28 where it is possible to see that there is not enough data in category 6 to incorporate it into the analysis. It is also possible to see that categories 2 and 3 are the most common. Looking at the 2-tailed tests for Data A and B, Tables 8.29, 8.30, 8.31 and 8.32, it is possible to see that in general the track categories are all statistically separate from each other. Data A

shows a degree of overlap between categories 2 and 3 as well as 3 and 4, with 50% of the test rejecting the null hypothesis. The results obtained from Data B disagree with Data A by demonstrating a degree of similarity between 1A and 1 with only 42% of the tests rejecting the null hypothesis. Data B also does not show the same similarity between categories 2 and 3 or 3 and 4 as Data A with significant rejection percentages of 81% and 86%. The results for Data A and B demonstrate how poskeys within the different track categories all degrade significantly differently from each other.

To see which track categories experienced the slowest vertical geometry degradation, 1-sided hypothesis tests were undertaken and the results noted in Tables 8.33, 8.34, 8.35 and 8.36. Looking at the results it is obvious to see that they suggest that lower track category performs best, with 1A performing best with the lowest degradation values, followed by 1 then 2, 3, 4 and 5. This is shown by none of the tests rejecting the null hypothesis in favour of a lower category containing higher degradation values but between 50-100% of the tests rejecting the other way around.

As the track category is based on the track speed and usage, to see any differences caused by the varying maintenance policies used on the different track categories, the speed and usage need to be taken into account. As the analysis has been undertaken using degradation data related to the usage, this is already taken into account. To take into account the different speeds it was decided to undertake a second layer analysis first splitting the data up by the maintenance history as before and then by the speed groups identified in Section 3.4.3. Over the data there was no speed band which contained both categories 1 and 5 or 6, or 1A and 2, 3, 4, 5 or 6. Due to this there were no results comparing between them but it can be presumed that if category 1 degrades faster than 1A and 2 faster than 1 then it can be inferred that 2 degrades faster than 1A. The results of the hypothesis tests taking into account the track speed returned similar values to the first layer test described above, demonstrating again that lower category track degrades slower with significant differences between the categories. The results of the second layer tests actually showed slightly more significant differences between the track categories which was unexpected as it was thought that some significance between the bands would be due to the different speeds. Due to these results it seems likely that it is the difference in asset policy used for different track categories that causes the low category track to degrade slower.

Track

A railway track consists of Plain Line (PL) and S&Cs. Due to the design of S&Cs, maintenance activities are harder to complete around them and quite often manual maintenance has to be used instead. With maintenance activities harder to complete and the extra dynamic forces encountered as trains traverse S&Cs it was decided this would be an important factor to analysis. Due to the geometry SD data obtained being related to 220yd poskeys, sections containing S&Cs also contain varying amounts of PL track. This means poor quality S&Cs could be hidden within a section with good quality

PL track, leading to a low SD recording for the whole poskey. As there is no way of calculating the SD of just the S&Cs, the analysis could only be completed by comparing poskeys that contained a S&C with purely PL track. Due to the rarity of S&Cs compared to PL there was not enough data within Data A to perform hypothesis testing, as seen in Table 8.37. Using Data B it was possible to perform the hypothesis tests, with the 2 tailed results recorded in Tables 8.38 and 8.39. The mean p-value of 0.21 and with only 33% of the test rejecting the null hypothesis there is not enough evidence to say that S&Cs degrade differently to PL.

Moving on to the 1-tailed test in Tables 8.40 and 8.40, it can be seen that 67% of the tests rejected the null hypothesis test in favour for PL degrading faster than S&Cs. Despite these results it is not possible to say that PL track degrades faster as there are many factors. First of all there is only three datasets of S&Cs used in the tests, increasing the uncertainty in the results. Additionally, there are more issues with the data analysed such as a lack of information of manual maintenance activities which are common on S&Cs, hence the S&Cs could be undergoing more maintenance than the PL track. The time spent on the maintenance is also higher on S&Cs as they are more complicated to maintain and a higher risk if left in poor quality. Due to the small quantity of poskeys which contained S&Cs, it was not possible to undertake a second layer analysis as completed for other factors, as the datasets when split up by an additional factor as well as S&Cs became to small to take worthwhile statistics from.

Route Criticality

The route criticality is a categorisation of track used by NR to organise track sections by incident costs, with 1 having the highest costs and 5 the lowest. Incident costs are related to the track usage as an incident on a line with minimal traffic will not affect many passengers but on a busy line more passengers will be affected hence there will be greater costs. The data amounts used for the comparison are summarised in Table 8.42, which shows that there is an even spread of poskeys for each route criticality with 3 and 4 being slightly more common and 5 being the least. The hypothesis test results for Data A and B is shown in Tables 8.43, 8.44, 8.45 and 8.46. The obtained mean values strongly suggest that criticality 1 and 2 degrade very similarly and 5 is dissimilar from all other criticality's. Data A results for criticality 3 show a significant difference to 1 and 2 with mean p-values under the significant 0.05 and with 88% and 100% rejections of the null hypothesis. Data B collaborates but does not demonstrate the same degree of significance with mean p-values of 0.14 and 0.10, which are about the 5% significance level, but 65 and 68% of the tests rejected the null hypothesis in favour of the alternative that they are significantly different. The difference between 3 and 4 were more significant in Data B, where within Data A 67% of the tests rejected the null hypothesis and in Data B 76% of the tests rejected the null hypothesis. As both the results from Data A and B have rejection percentages above 67% between criticality 3 and 4, it can be said that they significantly different. This can be said for all the criticality bands except 1

and 2, where there is evidence to suggest they are similar. This evidence consists of low rejection percentages of 0% of Data A and 21% of Data B and high mean p-values of 0.39 and 0.25.

By performing 1-tailed tests, recorded in Tables 8.47, 8.48, 8.49 and 8.50, we can see that in general lower criticality track degrades faster except 1 and 2 where there were no tests that rejected the null hypothesis. The columns for criticality 5, show that all the mean p-values are below 0.05 with 85-100% reject rates. These results show that poskeys in criticality band 5 degrade significantly faster than the other bands. This is similar for band 4 which has strong evidence to suggest that poskeys in this band degrade faster than those in bands 1, 2 and 3. The 1-tailed results for criticality 3 show significantly larger degradation rates than 1 and 2, with mean p-values of 0.01, 0.00, 0.09 and 0.07 and rejection percentages of 88%, 100%, 63% and 73% respectively. So it can be seen from the tables that higher criticality track tends to degrade slower except bands 1 and 2 which are too similar to say one degrades faster. The results are similar to the maximum speed analysis, possible due to higher criticality lines tending to have higher track speeds.

To see if the results obtained from the analysis were due to the differing track speeds in the criticality bands, it was decided to perform a second layer analysis first taking into account the speed and then the criticality. This was undertaken by first splitting up the datasets related to the maintenance history by the speed bands obtained in Section 3.4.3, and then by the criticality. The results obtained using Data A showed less significance between the bands expect between bands 1 and 2, where the rejection percentage increased from 0% when speed was not taken into account to 29%. In general though the mean rejection percentage and mean p-value (ignoring the values obtained comparing datasets against themselves (diagonal of the tables)) for the first layer analysis was 86% and 0.07 whereas the second layer analysis returned values of 71% and 0.11. The same can be said with Data B with the mean values for the first layer analysis being 80% and 0.07 whereas after taking into account the speed the tests returned mean values of 61% and 0.15. The lower rejection percentages and higher p-values obtained when speed was taken into account proves that some differences found in the first layer analysis were due to the track speed differences, and all the other differences between the different track speeds such as assets choices. Even after taking speed into account, the results still showed that poskeys of different criticality bands degrade significantly differently from each other except 1 and 2 where there is some evidence of similarities. In general, it can be said that higher route criticality track (1 and 2) tends to degrade slower than less critical track.

Embankments, Soil Cuttings and Rock Cuttings

To test whether installing track on geotechnical infrastructure affects the rate of vertical geometry degradation the poskeys were split into four groups. The first group consisted of poskeys that contained no embankments, soil or rock cuttings and the second, third and

forth groups consisted of poskeys that contained embankments, soil cuttings and rock cuttings respectively. Any poskey that contained more than one type of geotechnical infrastructure was removed as these poskeys might skew the results and hence reduce the accuracy of any comparisons. It can be seen from Table 8.51 that there was not enough data within Data A to include rock cuttings. The table also demonstrates the large amount of track build on geotechnical infrastructures, with more track occurring on them than not. To see if the different types of geotechnical infrastructure contribute to differing rates of vertical geometry degradation hypothesis tests were undertaken, with the mean p-value and rejection percentages noted in Tables 8.52 and 8.53 for Data A and 8.54, 8.55 for Data B. The mean p-values obtained using Data A show no significant differences between the groups but the rejected percentage table shows that 50% of soil cuttings compared to no infrastructure rejected the null hypothesis and 38% were rejected at a 5% confidence when comparing embankments and no infrastructure. Data B agrees with Data A as all the mean p-values are well above the significant level, additionally the highest percentage of rejected null hypothesis occurred between soil cuttings and no infrastructure but this was only 26% and had an associated mean p-value of 0.42. Looking at the results of rock cutting 0% of the tests rejected the null hypothesis against any other group, meaning there is no significant statistical difference between rock cuttings and the other groups. There was only three datasets usable for rock cutting, which does not give a general trend over all ages of track and hence further data would be required to be more confident in this result.

The 1-sided results in Tables 8.56, 8.57,8.58 and 8.59 again show little difference between the groups with the majority of mean p-values greater than 0.4. The most evidential results are comparing soil cutting to no infrastructure using Data A, in which 50% of soil cuttings reject the null hypothesis in favour of poskeys on soil cutting degrading faster than poskeys not on geotechnical infrastructure, but a high mean p-value of 0.25 is associated with this and Data B disagrees with a rejection percentage of only 14%. During the comparison between soil cuttings and embankments a low mean p-value of 0.06 was obtained giving evidence of soil cuttings degrading faster but to obtain this value only 38% of the tests were significant and also Data B only reported 8% of the compared datasets rejecting the null hypothesis. Due to these results it cannot be said that geotechnical infrastructure has a significant effect on the vertical geometry degradation, with all groups performing similar, hence there is no evidence to say that one performs better than another.

Curvature

Curved sections of track undergo different dynamic loads than PL track due to a centrifugal force pushing the train outward as it traverses the curved track section, increasing the load on the outer rail. As cant is normally introduced to curved sections of track to allow higher speeds without the risk of derailment there are additional loads that can occur on the inside rail. This is due to some trains not travelling fast enough for the

centrifugal force to reach equilibrium with the cant, causing forces to shift from the outside rail to the inside rail increasing loading. Due to the different loading that occurs at curved track sections it was decided that it would be an important factor to check. The curvature is measured as 1/curve radius in meters and from the data amounts recorded in Table 8.60 it can be seen that there is a large amount of curved track section on the UK rail network. It was decided to split the degree of curvature into 3 bands; 0-0.0005, 0.0005-0.001 and 0.001-0.002 m^{-1}. This was done to see if different degrees of curvature have varying affects on the vertical geometry degradation. The results from the hypothesis tests are recorded in Tables 8.61, 8.61, 8.61 and 8.61. There was not enough data within Data A to include the 0.001-0.002 band within the analysis, but comparing the others it is possible to see in general that the curvature has little effect on the track vertical geometry degradation. There is evidence to suggest that there is some difference between no curvature and 0-0.0005 with a null hypothesis rejection percentage of 50% but a high mean p-value of 0.39 suggests otherwise. This is backed up by Data B results where a rejection percent of 23% can be seen with a mean p-value of 0.31. Data B also shows little difference between the curvatures with the most different being the highly curved track between 0.001-0.002, which had 5% significance rejection percentages of 44%, 50% and 31% but mean p-values of 0.32, 0.10 and 0.27. The 2-sided hypothesis test results for Data A and B suggest there is a difference between the different curvatures with 0-0.0005 and 0.0005-0.001 being closest and 0.001-0.002 being most different from the rest but none are significantly different.

Looking at the one-tailed results for Data A in Tables 8.65 and 8.66, they suggest that track with no curvature performs worse (experiences higher levels of vertical geometry degradation), with mean p-values of 0.21 and 0.18 and rejections of the null hypothesis of 50% and 33% when comparing against curvatures between 0-0.0005 and 0.0005-0.001. Curvature of between 0-0.0005 results in datasets of poskeys with lower degradation rates than the other curvature groups, with no test rejecting the null hypothesis in favour of 0-0.0005 is larger. Data B results in Tables 8.67 and 8.68 shows less significance than Data A. Comparing between no curvature, 0-0.0005 and 0.0005-0.001 the highest rejection percentage is 28% between no curvature and 0-0.0005, which also had the highest rejection percentage between curvatures in Data A. Similarly to Data A 0-0.0005 seems to perform best with no tests rejecting the null hypothesis. The highly curved track between 0.001-0.002 seems to degrade fastest, with mean p-values of 0.22, 0.05 and 0.20 and rejection percentages of 50%, 69% and 44%. The results show no highly significant differences between the curvatures with the most significant being between 0-0.0005 and 0.001-0.002, in which 0.001-0.002 degrades faster with 69% of the tests rejecting the null hypothesis. In general track with a curvature between 0-0.0005 performs best followed by 0 and 0.0005-0.001, which are very similar and then 0.001-0.002 which degrades fastest. Despite the results suggesting this, the results in general show that curvature has minimal effect on the rate of vertical geometry degradation other than highly curved track which tends to degrade faster.

To check if the results obtained were being skewed by an underlying reason and as there was enough data available it was decided to perform a second layer analysis. For this speed was chosen as the additional factor as used in some other factors. After splitting the data up by speed and then curvature there was not enough data for the highly curved track between 0.001-0.002 to be included. Results obtained for the 2-tailed tests showed fewer differences between the curvatures, with Data A's rejection percentages dropping from 50%, 17% and 17% to 17%, 40% and 0%. The same can be said about Data B's results which show reductions from 23%, 8% and 15% to 9%, 7% and 6%. The results from the 1-tailed tests also saw reductions in rejection percentages. After performing the second layer analysis, it could be seen that curvature below 0.001 has little to no effect on the rate of vertical geometry degradation.

Cant

Cant is the angle of the track and occurs on curved track section, with the measurement recorded being the height difference between the low and high rail in cm, and can lead to differences in the loads on the rails as discussed in Section 3.4.3. The results from the two tailed hypothesis tests can be seen in Tables 8.70, 8.71, 8.72 and 8.73. The results obtained using Data A show a strong similarity of track with no cant and cant between 0-40, with a 0% rejection of the null hypothesis and a mean p-value of 0.54. Similarly, cant between 40-80 and 80-210 show no significant difference but with a lower mean p-value of 0.11. There is evidence of a significant difference between track with a cant between 80-210 and track with no cant or cant between 0-40 with mean p-values of 0.04 and 0.02 and null hypothesis rejection percentages of 75% and 100%. Unlike Data A, Data B shows very little difference between the four groups with the lowest mean p-value of 0.36 and highest rejection percentage of 18% neither of which is remotely significant.

Assessing the results for the one sided tests for Data A it suggests that track with a cant between 80-210 performs best with none of the tests results rejecting the null hypothesis in favour of 80-210 degrading faster. This is backed up by the mean p-values results of 0.02, 0.00 and 0.05 and rejection percentages of 75%, 100% and 75% that were found when comparing no cant, 0-40 and 40-80 against 80-210. The results also suggest that having a cant between 40-80 performs the next best with no cant and 0-40 performing too similar to say if one degrades faster. Like with the 2-sided results Data B shows no significance difference between the different degrees of cant. As Data A and B disagree it is difficult to say if there is any significant difference between the cant degree or which degrades faster as even though Data A has fewer assumptions and hence can be trusted more, Data B contains a lot more data, with Data A only using 2 datasets for 0-40 and 80-210, as seen in Table 8.69. Due to the lack of data in Data A, we cannot be confident in its results especially when Data B disagrees. Hence, it is more probable that Data B results are true which shows little difference between the different cants, which agrees with the analysis on curvature, in Section 3.4.3, which makes sense as canted track occurs on curved track, with higher curves having greater cants.

Similarly to with curvature, it was decided to perform a second layer analysis taking into account the track speed first but the returned results were very similar to the first layer analysis showing very little difference between the different cant amounts.

Maximum Axle Load

The maximum permissible axle load in tonnes for the track sections were split into three groups; 0-22, 23-25 and 26 tonnes. These groups were chosen to try to maintain large datasets whilst selecting three groups for more detail than just two. The amount of data for each group can be seen in Table 8.78, which shows how the chosen groups maximise the amount of datasets used to a minimum of three for Data A and twenty for Data B. The results from the 2-sided hypothesis tests for Data A can be seen in Tables 8.79 and 8.80 and show that 0% of the tests rejected the null hypothesis when comparing between 0-22 and 23-25. The low mean p-value of 0.11 does show that, even though none of the tests showed a significantly different at a 5% significance level, nearly all the tests were close to showing a significant difference. So there is evidence to say that the degradation rates experienced on groups of 0-22 and 23-25 tonnes are different. The rejection percentages of 67% and 75% when comparing between 26 and 0-22 or 23-25 give evidence to support that track with maximum axle loads of 26 tonnes degrades differently to other track with lower axle loads. The results from Data B, seen in Tables 8.81 and 8.82, show more significant differences between the groups. Unlike Data A where 0% of the tests rejected the null hypothesis in favour of a statistical significant difference between 0-22 and 23-25, Data B shows a 75% rejection rate from the tests and a mean p-value of 0.15. The low mean p-value of 0.11 for Data A and 75% rejection from Data B gives evidence to support that there is a statistical difference between the axle loads. As with Data A, Data B shows a difference between axle loads of 26 and 0-22 with a mean p-value of 0.07 and with an 85% rejection from the tests. Data B also shows some difference between 26 and 23-25 with 45% of the tests rejecting the null hypothesis. Looking at the results it can be said that in general the three axle load groups come from populations with differing distributions.

The 1-sided results have been recorded in Tables 8.83, 8.84, 8.85 and 8.86. The results from Data A show that track with permissible axle loads of 26 tonnes tend to have the lowest degradation rates with 67% of tests returning significant p-values in favour of 0-22 degrading faster and 75% for 23-25, with no tests returning significant values for track with axle loads of 26 degrading faster. The results show a rejection percentage of 33% in favour of track with axle loads between 23-25 degrading faster than 0-22, but this with a mean p-value of 0.28 is not significant enough to say with confidence that 23-25 degrades faster. The results for axle loads of 26 tonnes are similar in Data B, with 0% and 4% rejection percentages when analysing if track with axle loads of 26 degrade faster than 0-22 or 23-25. The tests performed the other way round to see if either 0-22 or 23-25 degrade faster than 26 agrees that 26 degrades slowest with rejection percentages of 90% and 46%. Due to the fact that both Data A and Data B strongly suggest that track

sections with axle loads of 26 tonnes degrade slower than track with lower axle loads, gives strong evidence that this is true. The results from Data B when comparing between 0-22 and 23-25 disagree with Data A, by showing that 80% of the tests resulted in significant p-values in favour of rejecting the null hypothesis in preference of the alternative that 0-22 degrades faster than 23-25. Due to the disparity between the results from Data A and B it is hard to say which should be believed, as Data B has more assumptions but a higher quantity of data, hence we cannot say with confidence which degrades faster 0-22 or 23-25 but the 80% rejection percentage obtained using Data B does tend to point to 0-22 degrading faster.

It was expected that higher axle loads would do more damage to the track and hence cause it to degrade faster but this is not what the results show. This could be due to many reasons such as higher axle load track sections tending to be constructed with better sleeper and rail designs. To reduce the impact of such factors an analysis has been performed by first splitting up the datasets by speed groups and then by axle load, as similar speed track tends to have assets installed which are alike in performance. The results obtained from Data A showed more significant results between 0-22 and 23-25 or 26 tonnes but less significance between 23-25 and 26. Performing the hypothesis tests on Data B again yielded results which showed less significance between the axle load groups. The mean rejection percentage and p-values obtained from the first layer analysis were 68% and 0.16 whereas the second layer analysis returned results of 46% and 0.23. The reduction in rejection percentages and increase in mean p-values gives evidence to support that some significant differences that could be seen in the initial results found from the first layer analysis were caused by differences between the poskeys track speeds and not axle load. There was still evidence in the results when the speed was taken into account to say that the three axle load groups are dissimilar, containing varying rates of vertical geometry degradation.

The 1-sided tests performed using Data A did not demonstrate many significant differences with the largest occurring between 0-22 and 26 tonnes. This result was far from significant though as 64% of the tests rejected the null hypothesis in favour of the group 0-22 degrading faster and 29% rejected in favour of track with axle loads of 26 tonnes degrading faster. Data B also showed more significant evidence to suggest that the group containing poskeys of axle loads between 0-22 contained significantly higher degradation rates than 26 tonnes. Between these groups 69% of the tests returned significant p-values in favour of 0-22 containing larger degradation values but only 5% the other way. Additionally, Data B suggests that in general track with lower axle loads degrades faster, but the results were not as significant as with the first layer analysis. This again was unexpected but there may be reasons for the differences which are not due to the axle load. The analysis undertaken was performed using degradation rates which were related to the usage expressed as EMGT. This meant that a degree of the damage caused by the axle loads had already been taken into account, as it is included in the calculation of EMGT obtained from NR and discussed in Section 3.2.6. It is generally known in

physics and engineering that higher loads lead to faster settlement and degradation, so the results from the tests do not make much sense unless the calculation for the EMGT overcompensates for the damage dealt from the different axle loads attaching too high of a factor to the higher axle loads or too small of a factor to the lower axle loads.

To test to see if the use of EMGT for the degradation rates was effecting the results it was decided to explore the affect of the axle load using the time related degradation rates. The results from these showed smaller differences between the groups with lower rejection percentages in the 2-tailed tests. The 1-sided test results disagreed with the usage based analysis, with some evidence to suggest that poskeys with higher axle loads tend to degrade faster, which was expected due to the higher loads.

Electrification

There are three types of electrification on the UK railway network: No Electrification, Overhead Line Equipment (OLE) and 3^{rd} or 4^{th} rail. The different electrification systems use contrasting methods to power the trains; no electrification using heavy diesel locomotives, $3^{rd}/4^{th}$ rail collect electricity via a shoe running along the electrified rail to power an electric train, and OLE uses pantographs to collect electricity from the suspended overhead lines. Despite the different trains having contrasting weights and power associated with them, most of this is taken into account in the usage conversion to EMGT as discussed in 3.2.6, but it was decided to still see if the electrification may still have a large impact on the rate of vertical geometry degradation. Table 8.87 shows the amount of data for each type of electrification which was used as part of the hypothesis tests. By looking at the table it can be seen that $3^{rd}/4^{th}$ rail is the least common with more than twice as much available data for OLE and more than twice again for no electrification. The results for the 2-sided hypothesis tests are noted in Tables 8.88, 8.89, 8.90 and 8.91. Data A shows that in 100% of the tests rejected the null hypothesis at a 5% significance level with all the mean p-values being 0.00. This shows a very strong significant statistical difference between the electrification types but the results for $3^{rd}/4^{th}$ rail were only based on the use of two datasets, due to this further evidence from Data B was used. Data B shows a similar significant different between no electrification and OLE or $3^{rd}/4^{th}$ rail with rejection percentages of 93% and 95% as well as mean p-values of 0.02 and 0.04. Data B does not show as much of a difference between OLE and $3^{rd}/4^{th}$ rail as Data A, with an obtained mean p-value of 0.18 and with 43% of the tests rejecting the null hypothesis at a 5% significance level. Looking at the results from both Data A and B it can be said that there is a highly significant difference in the vertical geometry degradation rates between no electrification and both OLE and $3^{rd}/4^{th}$ rail track. There is also evidence to suggest that OLE and $3^{rd}/4^{th}$ rail track tend to degrade differently as well.

To test which method of electrification degrades faster or slower, 1-sided hypothesis tests were used with the results placed in Tables 8.92, 8.93, 8.94 and 8.95. The results from Data A show that track with no electrification degrades significantly faster than $3^{rd}/4^{th}$

rail which in turn degrades significantly faster than OLE. This is shown by the mean p-values of 0.00 in the no electrification column and the 0.00 in the $3^{rd}/4^{th}$ rail column when it was compared to OLE. These p-values are backed up with 100% of the tests rejecting the null hypothesis in favour of them degrading faster. As with Data A, B shows that track sections with no electrification degrade fastest with mean p-values of 0.01 and 0.02 and associated rejection percentages of 96% and 95% when comparing against OLE and $3^{rd}/4^{th}$ rail. Again Data B also demonstrates that $3^{rd}/4^{th}$ rail degrades faster that OLE, but not as significantly as Data A, with 57% of the tests rejecting the null hypothesis with a mean p-value of 0.16. By looking at both Data A and B, and their agreement with each other, it can be said with confidence that the track with different types of electrification experience significantly different vertical geometry degradation rates, with track installed with no electrification degrading fastest followed by $3^{rd}/4^{th}$ rail and then OLE. Despite these significant results with the use of EMGT to take into account the differences in usage and train weights (with diesel trains are much heavier), there maybe another reason for the disparity between them. Most fast line over 100 MPH in the country have OLE installed and higher speed lines tend to use the larger more resilient sleepers and rail, which would reduce the degradation rate. Due to this a second layer analysis was performed first taking into account the speed then the electrification type.

Using the same statistical techniques an analysis was undertaken taking into account speed first and then the electrification type. The four speed groups found in Section 3.4.3 have been used to split up the datasets first and then split by the electrification type and compared only within the same maintenance history and speed group. Results obtained from the 2-sided hypothesis tests for Data A showed less significant results than the first layer analysis. The results agreed that track sections with no electrification degrade significantly different to OLE and $3^{rd}/4^{th}$ rail, but with less significance than seen in Table 8.89 with rejection percentages of 88% and 90% instead of 100%. The second layer analysis disagrees that OLE and $3^{rd}/4^{th}$ rail are significantly different with only 25% of the tests rejecting the null hypothesis, this is unlike the first layer analysis where 100% rejected the null hypothesis. This shows that some significant difference seen as part of the first layer analysis was actually due to the track speed differences. Performing the tests using Data B gave results that mirrored those obtained using Data A with large significant differences between no electrification and OLE or $3^{rd}/4^{th}$ rail. The associated rejection percentages obtained were 83% and 73%, which were lower than the 93% and 95% obtained as part of the first layer analyses using Data B, as seen in Table 8.91. Data B also returned results which showed similarities between OLE and $3^{rd}/4^{th}$ rail, with a mean p-value of 0.40 and with only 13% of the tests rejecting the null hypothesis. Again the results for the second layer analysis of Data B taking into account the different track speeds showed less significant differences between the electrification types than the first layer analysis, proving that some differences seen were due to the track speed. Assessing the results it is possible to say that track sections with no electrification degrade significantly differently but there are similarity's between OLE and $3^{rd}/4^{th}$ rail.

Due to the results returned from the 2-sided tests showing differences to the first layer analysis it was decided to also re-perform the 1-sided tests. Data A showed that sections with no electrification degrade significantly faster with 90-94% of the tests rejecting the null hypothesis in favour of this. Unlike the first layer analysis where 100% of the tests rejected the null hypothesis in favour of the $3^{rd}/4^{th}$ rail dataset containing higher degradation rate values than OLE, only 50% rejected the null hypothesis when speed was taken into account. Data B agrees that no electrification degrades fastest with rejection percentages between 84-85%. The results from Data B also show little evidence of either OLE or $3^{rd}/4^{th}$ rail degrading faster with associated rejection percentages of 7% and 12% and similar mean p-values of 0.51 and 0.46.

Rail Type

A wide array of rail types are used on the UK railway system, from older Bullhead (BH) designs to newer Flatbottom (FB). As seen from Table 8.96, there are no BH rails in Data A as this is the post-2006 data, in which no BH rails were installed as they had been superseded by FB designs. Due to this BH rails are only found on older, low priority track which has not been upgraded yet. It is also possible to see from the table that 113lb FB rail is the most common followed by UIC and 110lb FB rail which not used any more. There was only enough data in Data A for an analysis to be undertaken on 113 lb FB and UIC 60 rail with the results of the hypothesis tests noted in Tables 8.97 and 8.98. The 75% rejection of the null hypothesis within the analysis shows that the vertical geometry of track installed with 113 lb FB and UIC 60 rail degrade at significantly different rates. Data B results, Tables 8.99 and 8.100, also show a significant difference between the rail types with a mean p-value of 0.06 and with 83% of the tests rejecting the null hypothesis. From the table it can be seen that UIC 60 degrades significantly different to the other rail types with mean p-values of 0.00, 0.00, 0.00, 0.02 and 0.06 with associated percentages of 100%, 100%, 100%, 86% and 83%. The same can be said for 95lb BH rail. In general there is a significant difference between all the rail types with the closest overlap being between 98lb and 109lb FB with mean p-values of 0.16 and with only 30% of the test rejecting the null hypothesis.

To see which rail type leads to faster track geometry degradation, 1-tailed tests were used, with the results noted in Tables 8.101, 8.102, 8.103 and 8.104. These results show that there is significant statistical evidence to say that the heavier the rail the better the track performs. There is also evidence to suggest that the FB design is better than BH for reducing vertical geometry degradation. Despite the minimal 3lb difference between the 95lb BH and 98lb FB rail, BH rail degrades significantly faster with a mean p-value of 0.03 and with 80% of the tests rejecting the null hypothesis. These tables show strong evidence that heavier rail performs significantly better but, the heavier rail types tend to be installed on track with higher speeds and with larger sleepers, which might be the reason for the results instead of the actual affect of the rail type on the vertical geometry degradation. To be able to say with more certainty that the rail type is a significant

factor it was decided to first split up the data by speed and then rail type. Taking into account speed, greatly reduces the differences between rail types. The rejection percentage between 11lb FB and UIC 60 using Data A reduced from 75% to 57%. Data B also demonstrated large reduction in rejection percentages with 95lb BH and 98lb FB dropping from 0.70 to 0.00, 109lb and 110lb FB from 0.44 to 0.30 and 110lb and 113lb FB from 0.61 to 0.28. The reduced rejection values show that the track speed was greatly affecting the difference between rail types. Using the results of the hypothesis tests when speed was also taken into account, three groups could be created by joining together similarly performing rail types. These groups consist of: Group 1; 95lb BH, 98lb FB, Group2; 109, 110lb, 113lb FB and Group 3; UIC 60.

To test that the groups were significantly different 2-sided hypothesis tests were performed with the results recorded in Tables 8.106, 8.107, 8.108 and 8.109. As seen in Table 8.105 there was not enough data with Data A to include Group 1. This is due to Group 1 containing the oldest type of rails installed on the UK railway network and hence they do not occur on track installed post 2006. Group 2 is the most common group followed by group 3 and then there is minimal data from group 1. In results obtained using Data A 57% of the tests rejected the null hypothesis when comparing group 2 and 3. This significant difference is mirrored by Data B where 69% of the tests rejected hypothesis. Data B show a highly significant difference between Group 1 and Group 2 with a mean p-value of 0.00 and 100% of the tests rejecting the null hypothesis. There were no speed groups where there were sleepers from Group 1 and Group 3 so an analysis was not possible between these. To assess which types of sleepers help reduce the vertical geometry degradation of track one-sided hypothesis tests were used, with the results noted in Tables 8.110, 8.111, 8.112 and 8.113. It was expected that the larger rail would perform better at reducing the rate of vertical geometry degradation. This is backed up by Data A in which 57% of the tests rejected the null hypothesis in favour of Group 2 coming from a population with higher rates of degradation, with 0% of the tests rejecting the other way round. This is backed up by Data B with 56% rejection rate but 22% of the hypothesis tests undertaken reject the null hypothesis in favour of Group 3 degrading faster. When comparing Group 1 and Group 2, Data B shows a very significant result that Group 1 contains larger degradation rates. Looking at the tables it can be said that Group 3 performs best, followed by Group 2 and then Group 3 with Group 3 performing much worse than Group 2. From this it can also be inferred that track sections with rail types contained in Group 1 experience much higher levels of degradation than Group 3.

The rail types installed on the UK railway are directly related to the sleeper choices. Due to this it was decided that it would be important to carry out a further analysis. To accomplish this it was decided that the data would be split by maintenance history, speed, sleeper type and then rail type, using the speed groups discussed in Section 3.4.3, and the sleeper groups in Section 3.4.3. This meant that the hypothesis tests were comparing rail types that had similar sleeper types and similar speeds. Speed was still included as

similar speed lines have interchangeable maintenance procedures and thresholds. It was required to use the grouped speed and sleeper types to give large enough datasets to perform an analysis on. The results from this third layer analysis are outlined in Tables 8.115, 8.116, 8.117 and 8.118 with the data amounts in Table 8.114. Looking at the results from Data A it can be seen that including the sleeper type greatly reduces the significance of the result between Group 2 and 3, with the rejection percentage dropping from 0.57 to 0.39. Data B shows even larger reductions with the rejection percentage between Group 1 and 2 reducing from 1.00 to 0.30 and the mean p-value increasing from 0.00 to 0.31. Data B also showed a reduction between Group 2 and 3 with the rejection percentage going from 0.69 to 0.60 and the mean p-value from 0.13 to 0.22. The results show that a degree of the differences found in the first and second layer analysis were due to the sleeper choices. With the sleeper types, speed and maintenance history taken into account it can be seen that there is an effect of the rail choice on the tracks vertical geometry degradation, but it does not seem to be a large difference.

Passenger % Usage

It was decided to see if track sections that undergo different ratios of passenger and freight traffic degrade at varying rates as the different trains types cause differing forces on the track. The percentage of passenger traffic was used, being calculated in two ways, either the amount or the usage as seen in Section 3.2.7. The results when the passenger percentage was split by usage demonstrated more significant differences between the five groups, each encompassing 20%, than when split by amounts. Additionally, as higher usages and axle loads are known to damage the track more, it makes sense to categorise the traffic by usage and not the amounts. Hence, the results for usage are shown and discussed here, with the amounts in each group shown in Table 8.119. The disparity of passenger and freight traffic on the UK rail network can be seen by the amounts in the table, with all five groups containing a substantial amount of data. The group with 80-100% is the most populated and 20-40% the least. There is a large amount of both mainly freight or passenger lines but also highly mixed lines. Tables 8.120 and 8.121 show the hypothesis test results for Data A. In these tables it can be seen by the mean p-values of 0.00, 0.00, 0.01 and 0.01 in column 0-20 that the highly freight based traffic lines with less than 20% passenger trains by EMGT degrade significantly different from the other groups with a greater amount of passenger trains. Another significant result can be seen when comparing 20-40% to 40-60% with a mean p-value of 0.04 and a 75% rejection over the tests. The results from Data A do not show any more significant differences between the groups but show that track sections consisting of 60-80 and 80-100% passenger trains degrade at similar speeds as seen by none of the tests rejecting the null hypothesis at a 5% significance level. Looking at Data B results in Tables 8.122 and 8.123, there is less evidence of significant differences between the groups with the most significant result occurring between 0-20% to 60-80% with a mean p-value of 0.06 and a 93% rejection percentage. Data B shows that again there is evidence that 0-20% degrades differently with the highest rejection percentages of 43%, 55%, 93% and 60%.

Due to the results it cannot be said that there are significant differences between the percentage of passenger trains that make up the traffic, except between 0-20 and the rest, but there are some differences as seen in Table 8.123 for Data B, where there are no comparisons in which no results rejected the null hypothesis, with all the rejection percentages returned occurring between 23-48%.

The one sided test results in Tables 8.124, 8.125, 8.126 and 8.127, show that the 0-20% group degrades fastest. Data A results returned 100% rejection of the null hypothesis in favour of the group 0-20% containing larger degradation rates. Additional the highest rejections occur in 0-20% when using Data B as well as the lowest p-values. The results never rejected the null hypothesis that 0-20% degrades slower in either Data A or B. Data A shows that 40-60% tends to perform best whereas the results from Data B give evidence to suggest that the group with 60-80% passenger trains performs best, with the lowest degradation rates. By looking at all the results it can be said that there is not much evidence to suggest that the percentage of passenger trains has a large affect on the rate of vertical geometry degradation, other than when the usage is more than 80% freight, which causes significantly faster vertical geometry degradation. The reason for a lack of evidence to suggest any significant difference maybe due to the use of EMGT which takes into account the different damage that freight and passenger trains do to the track. To see if this was the case a similar analysis was performed on the degradation data related to time but this showed no significant differences either other than 0-20% containing the largest degradation rates similar to the results when usage was used.

Axle Load $>$ 50 % Usage

It was decided to categorise the traffic similarly to the percentage of passenger trains by looking at the axle load of the trains as this meant that empty freight trains would not be classified as freight but in the less than 50 tonnes categories, so the groups should be more related to the forces exerted on the trains than just passenger and freight. The results were calculated for both the amount of trains with combined axle loads per carriage more or less than 50 tonnes and the tonnage of trains. Results obtained from using the usage datasets instead of the amount of trains showed more significant differences between the percentage groups, so these results have been shown and analysed here. The data amounts used are shown in Table 8.128, in which it is possible to see that the groups are very similar in sizes. Results from the hypothesis tests are noted in Tables 8.129, 8.130, 8.131 and 8.132. The mean p-value results for Data A show that groups 0-20% and 80-100% degrade different from the other groups with mean p-values between 0.00 and 0.08 and rejection percentages of 67-100%. The other groups seem to degrade similarly with a slight difference between 40-60% and 60-80% in which 50% of the tests rejected the null hypothesis. Data B also shows a significant difference between 80-100% and the others with mean p-values of 0.12, 0.11, 0.04 and 0.05. Unlike Data A, Data B shows a degree of similarity between 0-20% and 20-40% as well as 40-60% and 60-80% with mean p-values of 0.32 and 0.33, with only 24% and 25% rejecting the null hypothesis. The

groups 0-20% and 20-40% are then dissimilar from the rest with mean p-values from 0.08 to 0.20 and rejection percentages between 55-70%. In general the groups are dissimilar from each other with the closest two groups being 40-60% and 60-80% in which 50% of the tests rejected the null hypothesis in Data A but only 25% did in Data B. The group with the highest percentage of trains with carriage combined axle loads over 50 tonnes, 80-100%, is shown to differ the most from the others. This is similar to the 0-20% group in the percentage of passenger trains, which makes sense as the 80-100% group would be primarily heavy freight trains.

To see which groups degrade faster, 1-tailed tests were used. The results for these can be seen in Tables 8.133, 8.134, 8.135 and 8.136. The results from Data A and B both agree that poskeys with traffic consisting of more than 80%, by usage, having axle loads above 50 tonnes degrade faster. This can be seen by the mean p-values which vary from 0.00-0.10 and the rejection percentages between 60% and 100%, with Data A showing more significant results. The rest of the results from Data A do not show significantly that any group other that 80-100% degrades faster or slower. Data B shows that 20-40% degrades significantly quicker than 40-60% with a mean p-value of 0.04 and a rejection percentage of 80% and also shows that 40-60% tends to have the lower degradation rates. In general the results show that the ratios of trains with different combined axle loads have little effect on the vertical geometry degradation, other than track with 80-100% traffic above 50 tonnes combined axle load, where this track degrades significantly faster than the others. These results are similar to the passenger percentage results in Section 3.4.3, which was expected, but by using the axle loads of the trains to classify them instead of just passenger or freight, more significant results were found especially for the non-mixed traffic lines.

Dirty Traffic % Usage

NR classify all traffic as either dirty or not depending on the amount of fines which come off the train where dirty traffic tends to be freight that is carry dirty materials which drop of train and infiltrate into the ballast. This causes the ballast to become fouled at an increased rate, and hence possible lead to faster geometry degradation, as discussed in Section 2.3.4. The amount of dirty traffic on any track section in the data never was higher than 40%, so it was decided to split the data up by into four groups of 0-10%, 10-20%, 20-30% and 30-40% dirty traffic. Dirty traffic was calculated as the percentage of the amount of dirty trains as well as the percentage of usage in EMGT. The results from the hypothesis tests showed that using the usage percentages, similarly to the percentage of freight trains discussed above, resulted in more significant results so these have been shown and discussed further here. Tables 8.138, 8.139, 8.140 and 8.141 show the 2-sided hypothesis results obtained from the data. There was not enough data within Data A to include 30-40% as seen from Table 8.137. There is also minimal data within Data A for 10-20% and 20-30% with only 2 datasets for each. The same occurs in Data B where there are only four datasets for 30-40%. Table 8.137 also shows how

the vast majority of the railway track is used by traffic which consists of less than 10% dirty traffic. The results from Data A show many significant differences with 100% of tests rejecting the null hypothesis when comparing 0-10% and 10-20% as well as 30-40% against the other three groups. There is also evidence to suggest a statistical difference between the groups 0-10% and 20-30% with a mean p-value of 0.10 and with 50% of the tests rejecting the null hypothesis. The closest two groups are 10-20% and 20-30% where only 25% of the tests rejecting the null hypothesis, despite these significant results the lack of data for the groups with more than 10% dirty traffic reduces the confidence in the mean results so Data B has been examined as well. Data B disagrees with Data A, showing minimal differences between the groups. The most significant result occurred between 0-10% and 30-40% with a 75% rejection of the null hypothesis and a mean p-value of 0.20. The results between 20-30% and 30-40% also returned a mean p-value of 0.20 but with 50% of the tests rejecting the null hypothesis instead. Despite the extra assumptions associated with Data B, the inclusion of a greater amount of datasets in the analysis makes the results more trustworthy and include more information such as comparisons between older track which has undergone many maintenance activities. Due to the results obtained from Data B it cannot be said that the percentage of dirty traffic has a significant effect on the rate of vertical geometry degradation.

Even though there are not many significant differences between the groups, one-sided tests have been used to discover which groups tend to degrade faster even if it is not significantly so. The results from these tests can be seen in Tables 8.142, 8.143, 8.144 and 8.145. The results from Data A show clearly that the group 30-40% degrades significantly faster than the groups with less dirty traffic. 10-20% degrades the next fastest and then 20-30%. The group with the minimal amount of dirty traffic between 0-10% seems to perform best. Data B, as with the 2-tailed hypothesis tests, shows less significance than Data A. The only significant results occurred between 30-40% and 0-10% in which 75% of the tests rejected the null hypothesis in favour of 30-40% degrading faster, with 0% rejecting the other way round. In general by looking at the mean p-values, which are lowest in column 30-40 and highest in column 0-10, it can be said that the group 30-40% degrades fastest and 0-10% slowest but the results are not statistically significant enough to be confident in this conclusion. The groups 10-20% and 20-30% are reasonably similar as seen from the 2-tailed tests and backed up by the results in Table 8.145 where 22% of the tests rejected the null hypothesis in favour of 20-30% containing higher degradation values and 11% the other way round.

Superficial Geology

When installing a railway track if the bedrock layer is not suitable, either in resistance or drainage, then a superficial layer is placed on top. The five main types of materials used as well as no superficial layer are noted in Table 8.146 with the associated data quantities. There is only enough data for the analysis using Data A to include clay, diamicton and sand, but Data B also includes enough to analyse gravel and peat as well. Table 8.146

shows that more than 50% of the railway track in the UK is built on a superficial layer, with clay and diamicton being the most common followed closely by sand, with gravel and peat barely being used. To see if the different superficial layer materials affects the rate of vertical geometry degradation two-sided hypothesis tests have been implemented, with the results from Data A noted in Tables 8.147 and 8.148. Looking at the values in the tables it can be seen that clay degrades significantly differently to both no superficial layer and diamicton, with mean p-values of 0.05 and 0.02, but has similarity's with sand as only 13% of the tests rejected the null hypothesis. Sand also performs significantly different from both no layer and diamicton with rejection percentages of 75% and 63%. Diamicton and no layer show evidence of having similar performance with only 25% of the tests rejecting the null hypothesis with a mean p-value of 0.29. The results obtained using Data B, as seen in Tables 8.149 and 8.150, show that no two types of superficial layer are very similar, with the lowest rejection percentage being 13% between gravel and clay or diamicton. The most significant results occur between sand and diamicton, peat and none and sand and peat, with these having rejection percentages of 70%, 67% and 58% and mean p-values of 0.15, 0.17 and 0.12. Data A results showed a significant difference between clay and none or diamicton but Data B results do not demonstrate such a significance with mean p-values of 0.33 and 0.20 and null hypothesis rejection percentages of 33% and 45%. Generally the results show no two superficial layer materials performing very similarly but the results also show that there is not enough evidence to be able to conclude with confidence that they perform significantly differently either.

As the results from the 2-tailed tests show some differences between the layer materials, 1-sided tests were performed to see which materials perform best at reducing vertical geometry degradation. The results from these tests for Data A are noted in Tables 8.151 and 8.152, and Data B in Tables 8.153 and 8.154. The results from Data A show that diamicton performs worst with mean p-values of 0.17, 0.01 and 0.04 when comparing against no layer, clay and sand. Additionally, no tests undertaken rejected the null hypothesis in favour of another material performing worse than diamicton, as seen by the zeros in the diamicton row of Table 8.152. No superficial layer being installed performs the second worse with 88% and 75% of the tests returning significant p-values (less than 0.05) and hence rejecting the null hypothesis in favour of no layer performing worse than clay and sand. Clay and sand are quite similar, as seen from the 2-sided test on Data A, and hence neither degrades significantly faster than the other, with 25% of the test results in favour of clay degrading faster and 13% of tests in favour of sand. Data B agrees with Data A that the use of diamicton for the superficial layer instead of no layer, clay and sand leads to faster degradation with mean p-values of 0.09, 0.14 and 0.11. Sand degrades slowest with 0% of the tests returning a significant p-value in favour of sand degrading faster. Peat is the worst material to use for the superficial layer with between 58-75% of the tests against the other material choices rejecting the null hypothesis, and only between 0-8% the other way round. The results from Data B do not show any significance between none, clay and gravel. Looking at all the results together it is possible to say that there are differences between the superficial layer types

but in general they are not significant. Of the materials used there is evidence to suggest that the use of sand is the most effective, followed by gravel and clay, then diamicton and then peat, with these two degrading significantly faster than the rest. Despite diamicton and peat returning results which show that they degrade significantly faster than no superficial layer, it cannot be said that using these materials for a superficial layer instead of nothing actually increases the rate of degradation. This is because the layer would have been installed on poor bedrock, whereas the results for no superficial layer had the track installed directly on top of a good quality bedrock layer.

To take into account the different bedrock and artificial layers that are installed with the superficial layer it was decided to undertake a three layer analysis, first splitting the data up by the maintenance history, bedrock and then artificial. This meant that comparisons were only undertaken between different superficial layer geologies, which occurred on track with similar bedrock and artificial layers as well as maintenance history. The bedrock was grouped using the similarities found in Section 3.4.3. Once this was completed it could be seen that there were only four datasets larger than 20 with an artificial layer but lots of data without, so it was decided to remove all sections that contained an artificial layer and instead split the data up by maintenance history, bedrock and then speed, with speed being chosen as track with similar speeds tend to have interchangeable assets and maintenance procedures. To maintain larger dataset the track speeds were split into the four groups of similarly degrading speeds as discussed in Section 3.4.3. The decision on how to split up the data resulted in comparisons only being undertaken between different superficial layer geologies, which occurred on track with similar bedrock and speeds, no superficial layer as well as maintenance history. Splitting the datasets by multiple factors leads to smaller datasets, which for Data A and B meant that gravel and peat could not be included in the analysis. The amount of data usable for the tests was still large for no superficial layer, clay, diamicton and sand, which can be seen in Table 8.155. Results of 2-sided hypothesis tests on Data A, which can be seen in Table 8.156 and 8.156, returned quite different results to the first layer analysis. This shows that some results were being affected but an underlying difference in speeds or bedrock type and not just the difference in the superficial layer geologies. Data A shows some differences between all the superficial layer types, with all the returned rejection percentages being between 31-56%. Tables 8.158 and 8.158, show the mean results of the hypothesis tests for Data B. These tables again show that there are differences between the superficial layers but the difference is not highly significant with rejection percentages between 24-53%. From the 2-sided hypothesis tests it could be seen that none of the layer types are completely similar with a minimum rejection percentage of 24% seen over the results from Data A and B. Additionally none of the superficial layer types where significantly different with the highest mean rejection percentage of 56% with a high associated mean p-value of 0.28. It was expected that there would be differences between the choices of superficial layer geology as different materials have dissimilar stiffness and drainage speeds but it was also expected that the superficial layer choices would not make a significant difference due to the majority of the vertical top

degradation occurring within the ballast above the superficial layer.

As there had been changes in the results of the 2-sided tests when the bedrock layer and maximum speed were also taken into account and statistical differences between the superficial layer geologies were shown by the 2-sided tests, it was decided that 1-sided tests would also be undertaken again. The results from the 1-sided tests, seen in Tables 8.160, 8.160, 8.162 and 8.162, show which superficial layers perform better. The results from Data A give evidence to support sand being the best choice for the superficial layer. This can be seen by the higher p-values and the low mean rejection percentages in the sand columns. Additionally, the sand row shows the highest rejection percentages with between 28-44% of the tests rejecting the null hypothesis in favour of the layer type containing significantly higher degradation rates at a 5% confidence level compared to sand. The other results between none, clay and diamicton are less clear cut such as clay and none where 20% of the tests saying clay degrades faster and 25% saying none, hence there is not enough evidence to say that one degrades faster than another. Looking at Data B, there is even more evidence to suggest that sand performs best at reducing vertical geometry degradation with only 4%, 0% and 3% of tests rejecting the null hypothesis in favour of sand degrading faster than none, clay and diamicton whereas 43%, 43% and 58% rejected the null hypothesis in favour of sand degrading significantly slower at a 5% significance level. Again, as with Data A, Data B show minimal evidence of either none, clay or diamicton degrading significantly faster than each other, with all the mean p-values calculated being between 41-54%. Looking at both Data A and B it can be said that track built on a sand superficial layer performs significantly better than other track. This was expected as sand is known as a good formation layer due to its high drainage speed. The results also show that no layer, clay and diamicton seem to perform similarly.

Artificial Geology

Sometimes even with a superficial layer installed a further artificial layer is used to increase the quality of the formation below the ballast. The material for this layer is most often a combination of slag, ash and other waste materials and is known as made ground. These are used as they are cheap waste products but produce a layer with high stiffness and permeability. There is only a small amount of track sections that use a superficial layer in the UK, as seen in Table 8.164, as the layer is only installed at locations of particularly poor formation. In this table it is possible to see that under 4% of the railway track installed in the UK was constructed with a superficial layer. To see if track sections with artificial layers perform differently hypothesis tests have been used, with the results for the 2-sided tests being noted in Tables 8.165, 8.166, 8.167 and 8.168. For the results obtain from Data A it can be seen that there is a significant difference between the two groups, with 100% of the tests rejecting the null hypothesis with a mean p-value of 0.00. It has to be noted that these means were created with only two datasets for the artificial layer, hence greater evidence would be recommended, for which Data

B was used. Unlike Data A, Tables 8.167 and 8.168 do not show a significant difference with mean p-values of 0.21 and with only 27% of the tests rejecting the null hypothesis, that they come from populations with similar distributions. Due to these results it was decided that there was not enough evidence to say that the vertical geometry of track sections that are built on top of an artificial layer degrade at different speeds.

The 2-sided test results do show some difference between the groups, so 1-sided tests were performed to see if track sections installed on an artificial layer degrade faster or slow. These results can be seen in Tables 8.169, 8.171, 8.170, 8.171 and 8.172. Data A gives strong evidence to suggest that track built on an artificial layer has slower vertical geometry degradation, with 100% of the tests rejecting the null hypothesis in favour of no artificial layer having higher degradation rate values. It must be noted again that these results are a mean based on only two datsets. Data B results, similarly to the 2-sided results, show less significance than Data A but do agree that an artificial layer reduces degradation. These results are not significant though as only 29% of the tests rejected the null hypothesis, additionally 7% of the tests found that the sections with an artificial layer degrade significantly faster.

Bedrock Geology

Every track section has a bedrock layer as this is the initial on-site geology. Depending on the bedrock and the track usage and design, superficial and artificial layers are sometimes placed on top of the bedrock layer to increase the tracks resilience, reducing the forces that are experienced in the bedrock layer, reducing geometry degradation. There is a vast array of different bedrock materials under the UK railway network, with the thirteen most prevalent being analysed. The amounts of data obtained for each type is outlined in Table 8.184, in which it can be seen that mudstone is the most common followed by chalk, clay and then sandstone. There was only enough data within Data A to include chalk, clay, limestone, mudstone and sandstone in the analysis, with only two datasets for limestone containing more than twenty calculated degradation rates. The 2-side hypothesis test results for Data A can be seen in Tables 8.183 and 8.184. In these tables it can be seen that limestone is significantly different from the other bedrock materials with 100% of the tests involving limestone rejecting the null hypothesis. There are other significant results, with tests between mudstone and chalk returning a mean p-value of 0.05 with 67% of the tests returning a p-value less than 0.05. There is also evidence to suggest that chalk and clay have similarities with a rejection percentage of 33%, mudstone and sandstone are also similar with a low rejection percentage of only 25%, but the two groups are dissimilar from each other with rejection percentages between 63-67%. Moving on to the results obtained from Data B in Tables 8.185 and 8.186 which include thirteen different bedrock geology types. First looking at if the results obtained from Data B agree with those from Data A. Data B agrees that track installed on a limestone bedrock degrades at a significantly different rate to chalk and clay with mean p-values of 0.09 and 0.25 and rejection percentages of 79% and 54%. Data B disagrees

that limestone is significantly different to mudstone and sandstone but still shows some difference with mean p-values of 0.23 and 0.17 and rejection percentages of 29% and 46%. Data A also showed that chalk and clay are similar and mudstone and sandstone are similar. Looking at the results in Data B there is evidence to support both, with only 18% of the tests rejecting the null hypothesis between chalk and clay and only 20% between mudstone and sandstone. Additionally, the results obtained from Data A showed a difference between the two groups, which Data B also supports. This is demonstrated by results of the tests between them which contained low mean p-values between 0.04-0.25 and high rejection percentages between 54-88%. Looking at the results it was decided to group the similarly performing geology materials together, to increase the dataset sizes. To accomplish this, groups of high mean p-values and low rejection percentages where made making sure that geology types in which more than 50% of the tests rejected the null hypothesis were kept in separate groups. This involved a lot of shuffling about but there are two distinct groups; one containing argillaceous, dolomitic, limestone, psammite, siltstone and the other containing sand, chalk and clay. Out of the remaining bedrock geology types it is possible from the results to see that halite and pebbly sand are similar and mudstone, sandstone, and slate also contained similar degradation values. Using these four groups it was possible to perform the hypothesis tests again to confirm that the groups are statistically dissimilar and to see which groups degrade faster.

The results for the 2-side hypothesis tests from the grouped bedrock geology using Data A can be seen in Tables 8.174 and 8.175. The amount of data in each group has been noted in Table 8.173, in which it can be seen that there is not enough data within Data A for group 3 to be analysed and there are only two datasets for group 1 but in general the amount of data that made up each dataset for each bedrock geology group was high. The 2-sided hypothesis test results from Data A show a significant difference between groups 1, 2 and 4. It can be seen that group 1 is significantly different with 100% of tests rejecting the null hypothesis that it is similar to either group 2 or 3 with mean p-values of 0.00. Looking between groups 2 and 4, a significant difference can still be seen with 75% of the tests returning p-values less than the 5% significance level with a low mean p-value of 0.07. For additional evidence and to see about group 3 the analysis has also been conducted using Data B and the results placed in Tables 8.176 and 8.177. These tables again show some significant differences between the groups. Group 1 is significantly different from group 2 with a 95% rejection percentage and group 2 is significantly different from group 4 with a rejection percentage of 86%. The results also show some differences between the other groups with a highest mean p-value of 0.27 and minimum rejection percentage of 38%. The results show that there are differences between the groups, with group 1 being the most significantly different from the other groups. There is not enough evidence within the results to say that any group is similar enough to join together.

To assess which bedrock reduces the track geometry degradation 1-tailed tests were used.

The results of these tests can be seen in Tables 8.178, 8.179, 8.180 and 8.181. There was strong evidence within the results from Data A to suggest that Group 2 has the lowest vertical geometry degradation followed by 4 and then 1 degrades fastest. The results from Data B also show that group 1 degrades significantly faster with rejection percentages between 53-95%. The next fastest group to degrade is group 4, closely followed by group 3 and then group 2 degrades the slowest. This is similar to the conclusions made from the results of Data A.

Maximum Speed

In the research undertaken within the literature review it could be seen that many pieces of previous research show that the speed trains traverse the track has a great effect on the rate of vertical geometry degradation. This is expected as an increase in the speed of a train increases the dynamic forces that are transferred into the track, leading to the ballast particles to break and move causing settlement and hence faster degradation of the vertical geometry. Assessing the types of assets which are installed on different speed tracks shows that in general the faster lines tend to use larger concrete sleeper designs and heavier rail designs than the slower speed track. The larger sleeper and rail designs are used to counteract the additional forces that are experienced on the faster track by giving the track added resilience. In addition, higher speed track tends to be a higher priority and hence keeping the track in a serviceable condition is of a greater importance. The maximum permissible track speeds were grouped using NR's grouping. This was done as these groups are used by NR to decide what quality the track is based on the vertical geometry SD i.e. a SD of 4 mm would be classified as good condition for track speeds 10-20, 25-30 and 35-40 MPH whereas for track speeds between 100-110 and 115-125 MPH the track would be classified as being in a very poor condition. This is due to the safety requirements, as faster tracks need a smoother running surface to be safely traverse. Due to this the track quality level that maintenance is undertaken at differs for each of the ten speed bands NR use. There is also a large disparity between the usage of the higher speed track compared to the slower track but this has been taken account of by using degradation rates calculated with EMGT instead of time. A comparison of the degradation rates of different speed track was undertaken to first see which track speeds degrade significantly differently to each other and then to see which tend to degrade faster.

To identify significant differences between the groups 2-sided significance tests were used as discussed in Section 3.4.2. First both Data A and B were split up by the maintenance history, so comparisons were only undertaken between track which had undergone the same maintenance activities and hence will be of similar ages with akin asset ages and ballast condition. These datasets were then split up by the different speed groups, with the amount of datasets remaining for each speed larger than 20 noted in Table 8.187. The total amount of degradation rate values that made up these datasets are also noted within the table. The table shows that there is a large spread of speeds across the UK

railway network with 75-80 MPH being the most common. Due to only one dataset within Data A being constructed for speeds 0-30, 35-40 and 45-50 MPH these have not been included within the hypothesis tests as the results of using one dataset and not a mean obtained from many contains more uncertainty as a few extreme values in the one dataset would skew the results. Additionally, with only one dataset, which contained degradation values from a renewal to the first maintenance action, the tests would only give an idea of any difference when the track is new and not over its life cycle. Looking at the mean results obtained from the 2-sided hypothesis tests for Data A, which are shown in Tables 8.188 and 8.188, it can be seen that there is evidence to suggest that the groups 55-60 and 115-125 MPH degrade significantly differently from the other speeds. This is shown for 55-60 MPH with all the obtained mean p-values being 0.00 and 100% of the tests rejecting the null hypothesis in favour of the datasets being significantly different at a 5% significant level. The same is true for track with speeds between 115-125 MPH where the hypothesis tests comparing against the other speeds rejected the null hypothesis in 67-100% of the tests with mean p-values between 0.00-0.05 all less than the chosen 5% significance level. The results obtained from the comparisons between the other speeds show a degree of similarity between 75-80, 85-95 and 100-110 MPH track with high mean p-values between 0.29 and 0.46, well above the significant 0.05 value. Looking at the rejection percentages it can also be seen that only 25-38% of the tests rejected the null hypothesis. There is also evidence to suggest that there are similarity's between track sections with speeds 65-70 MPH and 75-80, 85-95 and 100-110 MPH, with mean p-values between 0.19 and 0.31 but rejection percentages between 33-67%. Due to this it cannot be said that there is enough evidence to say that the speed group 65-70 MPH degrades significantly different but there is enough evidence to suggest that there is some difference.

For more evidence and for results for speeds 0-30, 35-40 and 45-50 MPH similar tests were performed using Data B, with the results noted in Tables 8.190 and 8.190. There is strong evidence to suggest that the vertical geometry of track sections with maximum speeds of 0-30, 35-40, 45-50 and 55-60 MPH degrade similarly, with mean p-values between 0.20-0.41 and low rejection percentages between 5-45%. The results also show that these speeds are significantly different from the other faster speeds with mean rejection percentages between 76-100% and mean p-values 0.00-0.07. The results for speeds between 115-125 MPH using Data B echo those obtained from Data A, with rejection percentages between 59-100% suggesting that track with speeds 115-125 MPH degrade significantly differently to the other speeds. Again similarly to Data A the results obtained from Data B show a degree of similarity between track sections with speeds 75-80, 85-95 and 100-110 MPH, with mean p-values between 0.16-0.38 and rejection percentages between 32-41%. The results from Data B show more significant differences than Data A with higher rejection percentages but the majority of the tests still failed to reject the null hypothesis at a 5% significance level and hence there is evidence to suggest that track sections of these speeds degrade at a similarly rate. The results also show similarities between speeds 65-70 and 75-80 MPH with only 35% of the tests rejecting the null

hypothesis but 65-70 MPH shows significant differences to the rest of the speed groups. Assessing both the results from Data A and Data B it can be said that there are four significant groups of speeds, within which there are similarities between the degradation rates. These groups consist of 0-60, 65-70, 75-110 and 115-125 MPH. Additionally, it was thought that the 65-70 MPH group could be a transition group where 65 MPH would fit into 0-60 and 70 into 75-110 MPH but when tested this was not the case, with 65 and 70 MPH being shown as statistically similar.

To check that these groups are significantly different from each other hypothesis tests were used again, with the data amounts displayed in Table 8.192. Due to reducing the amount of groups the size of the datasets has increased, leading to more confidence in the results of the tests which for the two sided tests can be seen the Tables 8.193, 8.194, 8.195 and 8.196. These tables show that the chosen four groups from before are all significantly different with the closest two groups being 65-70 and 75-110 MPH with mean p-values and rejection percentages for Data A and B of 0.19, 0.06, 67% and 80%. To see which track speeds degrade faster 1-sided hypothesis tests were undertaken and the results noted in Tables 8.197, 8.198, 8.199 and 8.200. The results from Data A show that slower track degrades faster. This is seen by 0% of the tests rejecting the null hypothesis in favour of faster track degrading at a greater rate than slower track, and between 67-100% of the tests rejecting the null hypothesis in favour of the slower track degrading faster. Similar results are demonstrated when Data B was used but with higher rejection percentages between 84-100%. The significant results in the tables and higher amount of data used, mean that it can be said with a high degree of confidence that slower track degrades faster than higher speed track. This was not the expected result of higher speed track degrading faster due to the higher dynamic loads caused by the passing trains. As noted before the faster track in general uses larger sleepers and rails, which could reduce the loads and hence damage caused by the trains by more than the increase caused by the higher speed trains leading to an overall decrease in degradation rate.

Sleeper Type

There are many sleeper designs and materials used in the UK railway network. The majority of these are made of concrete, but steel and timber are also used. Concrete sleepers tend to be used on faster, higher priority track due to their higher resilience whereas the steel and timber tend to occur on the slower track where the extra resistance is not required and hence the cheaper and easier to install materials are used. The amount of data used as part of the hypothesis tests are outlined in Table 8.201, where it is possible to see that there was enough data to include eight concrete, six steel and three timber sleeper designs. Of the designs the G44 and F27 seem to be the most common concrete sleepers and W560 the most frequently used steel. Despite the F27 sleeper currently being installed on large sections of track it can be seen from the table that minimal new track (Data A) has been installed using F27 sleepers, with the G series

sleepers becoming standard. There was only enough data within Data A to include the G44, G47 and G49 concrete designs and the W560 steel design in the analysis. The results from the 2-sided hypothesis tests for Data A can be seen in Tables 8.202 and 8.203. Looking at the tables the first obvious conclusion to draw is that the W560 steel sleeper design degrades significantly differently to the concrete design sleepers with 100% of the tests rejecting the null hypothesis and mean p-values between 0.00-0.01. The three concrete designs show less significant differences between them. The G49 sleeper shows evidence of performing differently to the G44 and G47 sleepers with 50% of the tests returning significant p-values, whereas only 25% of the tests rejected the null hypothesis when the tests were performed between G44 and G47 sleepers. For additional evidence and to include more designs in the analysis Data B was used, with the results of the 2-sided hypothesis tests noted in Tables 8.204 and 8.205. Comparing the results from Data A and B it can be seen that Data B agrees that W560 is significantly different from G44, G47 and G49 designs with mean p-values of 0.00 but it disagrees that G44 and G47 are more similar than G47 and G49, with a 50% rejection percentage between G44 and G47 whereas Data B returns no significant p-values at a 5% significance across the undertaken tests between G47 and G49. Track installed with G47 and G49 seem to behave similarly to F24, F27, F28 and W600 with rejection percentages all below 25%. Softwood and timber sleeper are statistically different from the other sleeper types with rejection percentages between 75-100% and are also different from each other with 50% of the tests rejecting the null hypothesis.

Due to the sleeper choice installed being highly related to the track speed it was decided that it was important to perform the tests first by splitting up the data by the track speeds, into the four groups shown in Section 3.4.3, and then the sleeper type. This was decided as comparing timber sleepers, which are only installed on track with speeds less than 70 MPH, against large concrete sleepers, which are used on high speed track, is comparing the speed differences as well the the sleeper difference. The two-tailed hypothesis test results for Data A and B are recorded in Tables 8.207, 8.208, 8.204 and 8.205 with the amount of available data used in Table 8.206. Using the results in these tables, the types of sleepers were grouped together, making sure the groups contained sleeper types which when compared against each other had low rejection percentages and high p-values. This was undertaken as for the stochastic PN model proposed in this thesis requires large datasets to make the required distributions. Similar sleepers where grouped and the tests re-performed until all the remaining groups were significantly different from each other. This resulted in three groups: Group 1; F23, F24, W500, W600, F27, W560, Group 2; F28, G49, G44, G47, F40 and Group 3; Timber, W402, HH10, Metal, W500. The types of sleepers that make up: Group 1 are smaller concrete sleepers and larger steel sleepers; Group 2 are large concrete sleepers and Group 3 are timber and smaller steel sleepers.

The results from the two-sided hypothesis tests for the grouped sleepers can be seen in Tables 8.212, 8.213, 8.214 and 8.215. Looking at the amounts seen in Table 8.211, it

can be seen that there was not enough data within Group 3 for it to be included in the analysis of Data A. The results of the hypothesis tests show strong differences between the groups with Data A rejecting the null hypothesis that Group 1 and Group 2 come from similar populations in 83% of the tests. Data B agrees with 72%. Looking at Group 3 in Data B, there is a significant differences between Group 2 and itself with a mean p-value of 0.03 and 87% of the tests rejecting the null hypothesis. There is some evidence of similarities between Group 1 and Group 2 but the results were deemed too significant to join the groups together. To assess which sleepers choices reduce the vertical geometry degradation, one-sided hypothesis tests were undertaken, with the results of these recorded in Tables 8.216, 8.217, 8.218 and 8.219. Data A shows Group 1 degrading significantly faster than Group 1 with 83% of the tests rejecting the null hypothesis in favour of Group 1 containing higher degradation vales, and 0% the other way round. The p-values also demonstrate such a strong conclusion. The results from Data B back this up with 74% of the test results returning a significant p-value. Data B also shows that Group 3 degrades significantly faster than Group 2. There is also evidence to indicate that Group 3 degrades faster than Group 1, with 52% of test results suggesting that Group 3 degrades faster and 11% the other was round. These results show that the large concrete sleepers perform best, followed with a large gap by small concrete and large steel sleepers and then timber and small steel sleepers.

Similarly to the analysis of the rail types performed in Section 3.4.3, due to the connection between sleeper type and rail type it was decided that a further third layer analysis of sleepers should be undertaken. This involved splitting the degradation rate data up by the maintenance history, speed and rail type, using the grouped speed and rail types from Sections 3.4.3 and 3.4.3. The resultant tables from the hypothesis tests have been noted in Tables 8.221, 8.222, 8.223 and 8.224. The data amounts are recorded in Table 8.220. Due to the additional splits, more datasets consisting of more than 20 datapoints (degradation rates) were created (datasets with less than 20 datapoints are removed from the analysis). This enabled all sleeper groups to be included in the analysis of Data A. The results from Data A suggest that the rail type was not the reason for the significant differences found as part of the first and second layer analysis with the rejection percentage between Group 1 and 2 actually increasing from 0.83 to 1.00 and the mean p-value reducing from 0.08 to 0.00. Data B also shows little difference in the results with the rejection percentages and mean p-values of the third layer analysis being 0.59-0.83 and 0.09-0.17 compared to the second layer analysis where it was 0.50-0.87 and 0.03-0.19. Due to this it can be said that the significances between the sleeper groups are most likely due to the sleepers and nothing else.

3.4.4 Analysis Summary

Tunnels and Stations

The results from the analysis did not show many significant differences between the vertical geometry degradation of track sections at stations and between stations, but there was enough evidence to suggest that there is a small difference between the two with in general track at stations degrading faster. The results showed a smaller difference than was demonstrated by Andrade and Teixeira (2011) but agree that track at stations degrades faster with Andrade and Teixeira (2011) suggesting that on average track at stations degrades 40% faster than PL track. The results for tunnels were more inconclusive but again were enough to say that track in tunnels tends to degrade at a different rate than track at stations or between them. Some results returned demonstrated track in tunnels degrading faster and other results indicated they degrade slower. What could be seen is that there was a greater variability of degradation rates experienced by track within tunnels, which could be due to them experiencing different environmental factors, drainage, ballast depth or maintenance procedures.

Track Type

The track type describes the method of joining the rails together either by the legacy jointed method, using plates and bolts or the newer, more common, method of welding long rail sections together to make CWR. The results from the analysis strongly demonstrate a large statistically significant difference in the vertical top geometry degradation rates between track installed with either jointed or CWR, with track installed on jointed rail degrading at a much faster rate. This is a common conclusion about the rail joints with Sato (1995).

Track Construction

The UK railway network track is constructed to different quality levels depending on the tracks future purpose. These construction bands vary from A to D with A being the highest quality construction. From the analysis there was strong evidence to suggest that track constructed to different quality bands degrade significantly different from each other, with the closest being bands A and B. The hypothesis test results also demonstrate that the track built to the higher construction bands undergoes slower rates of degradation.

Track Category

The track category is related to the usage and speed of the track, with 1A being fast heavily traversed lines and 6 being slow with minimal traffic, as seen in Figure 3.1. The outcome of the hypothesis tests showed a significant difference in the degradation rates between the track categories, with lower track categories seeming to degrade faster.

Track

Railway track can be either PL or contain S&Cs. As the analysis had been undertaken using 220yd poskeys, all sections categorised as S&C would also contain a large amount of PL track, which if in good condition could hide a fast degrading S&C, when the SD is calculated. Additionally, work on S&Cs is quite often manual work, which was not included in the analysis, hence small work could be being undertaken on the S&Cs reducing there degradation rates, but this work was unknown. Despite these issues it was decided to still include track, either PL or S&Cs, into the analysis as previous research has shown concern about the degradation of track at S&Cs, due to higher loading and the inability to maintain them with large maintenance equipment. The results of the analysis showed some differences between the two, with PL track tending to degrade faster.

Route Criticality

A track sections route criticality is based on the cost of failure of the track, where highly trafficked lines tend to be of a higher criticality as a fault would affect more people. The results of the analysis showed that there are significant differences between the degradation rates of the route criticality bands, except 1 and 2 which are reasonably similar. It can also be noted that higher criticality bands degrade significantly slower. This was probably due to the higher quality assets and more rigorous maintenance procedures used on the higher criticality track reducing the rate of degradation.

Embankments, Soil Cuttings and Rock Cuttings

It was thought that constructing track on geotechnical infrastructure may affect the rate of vertical geometry. This affect could have been positive if the geotechnical structure was of good quality, or negative if the structure had low stiffness and hence high degrees of settlement. The analysis instead showed very little difference in the degradation rates of poskeys installed on embankments, soil and rock cuttings or no geotechnical structure. This does make sense though, as research undertaken by Selig and Waters (1994) showed that most of the settlement occurs within the ballast and not in the layers below, hence changes in the rate of settlement below the ballast would have little effect on the track vertical geometry degradation. Additionally, due to most of the UK railway network being initially constructed many years ago, most of the settlement below the track would already have occurred.

Curvature

A large quantity of track sections installed on the UK railway network incorporate a curve. It was thought that these curves due to causing uneven forces across the rails when traversed would lead to inconsistent settlement. To test this an analysis was performed comparing the vertical geometry degradation rates of track sections of different

curvatures. The results of these test showed minimal differences between levels of curvature between 0-0.001 m^{-1}. They also showed evidence to suggest that highly curved track between 0.001-0.002 tend to degrade slightly faster than the other degrees of curvature.

Cant

Track is designed with a cant to allow trains to travel at high speeds on curved sections of track as the cant counteracts the centrifugal forces. Performing the analysis it could be seen that, similar to the curvature that cant is linked to, that the degree of cant has little effect on the rate of vertical geometry degradation. This is maybe due to the cant design keeping the forces across the rails even as the trains traverse the curved track, or that the additional forces that occur have a minimal impact on the vertical geometry SD. It maybe the case that the cant and curvature effect other aspects of the track such as the gauge more than the vertical geometry.

Axle Load

Increased loading causes faster settlement and hence it was expected that higher axle loads would cause faster vertical geometry degradation but from the analysis this did not seem to be the case. The results showed that there were differences between the three chosen axle load groups of 0-22, 23-25 and 26 tonnes with the groups of lower axle loads tending to contain poskeys with higher rates of degradation. This maybe due to the degradation rates being calculated using the EMGT supplied by NR, which already took into account the axle loads. To check this the time related degradation rates were used instead for the hypothesis tests, with the results showing less significant differences between the groups and with the higher load poskeys experiencing slightly faster degradation, which was the expected result.

Electrification

There are three types of electrification on the UK railway network; no electrification which uses heavy diesel trains, $3^{\text{rd}}/4^{\text{th}}$ rail and OLE. The results from the analysis showed significant differences between no electrification and $3^{\text{rd}}/4^{\text{th}}$ rail and OLE as well as some difference between $3^{\text{rd}}/4^{\text{th}}$ rail and OLE. 1-sided hypothesis tests showed that poskeys which have no electrification, and hence use diesel trains, degrade significantly faster than electrified line, with OLE degrading slightly slower than $3^{\text{rd}}/4^{\text{th}}$ rail. This result was expected due to the additional weight of diesel trains causing increased damage to the track but the use of EMGT should already have taken this into account but this calculation my not be fully taking into account the difference in damage caused by different axle loads.

Rail Type

On the UK railway network there are two main types of rail design; BH and FB, with BH being a legacy design. Within these there are different sizes of rail, recorded as weights. It was thought that larger rail would increase the stiffness of the track and hence reduce vertical geometry degradation. From the results of the two sided hypothesis tests it was possible to group together similarly performing rail types. This created three significantly different groups: Group 1; 95lb BH, 98lb FB, Group 2; 109, 110lb, 113lb FB and Group 3; UIC 60. Further analysis of the rail type was completed taking account of the maintenance history, track speed and sleeper types. When all of these were taken account of the significance between the rails was reduced, showing some differences but these were no longer highly significant. This proved that a large degree of the differences of vertical geometry degradation between the rail types found by the first and second layer analysis was actually due to the sleeper choices and track speed. The results also showed that Group 1 degraded faster than the other groups and Group 2 degraded faster than Group 3. This was expected as Group 3 contained the largest rails and Group 1 the smallest. Another thing to note from the analysis was that BH designs tended to lead to faster degradation than similar weight FB designs.

Passenger Percentage (Usage)

The percentage of passenger trains was calculated in Section 3.2.7 and was thought to be important as freight trains tend to be heavier. This would lead us to expect that lower amounts of passenger trains and hence higher amounts of freight trains would cause faster degradation of the track due to the higher forces. The use of EMGT to calculate the degradation rates takes into account the different weights and tractive units and hence was expected to reduce any differences seen. The results from the tests show that there are differences between the five groups (each encompassing 20%) but the only significant results were between 0-20%, which contains the most freight, and the others. In this there was evidence to suggest that the poskeys with less than 20% passenger trains tend to degrade faster.

Axle Load $>$ 50 % Usage

Similar to the percentage of passenger trains, it was decided to categorise the traffic by the combined axle load of each carriage. This was thought to differ from just looking at the percentage of passenger trains, as full freight trains have axle loads over 50 tonnes whereas empty freight and passenger tend to have lower loads. Hence the groups are created by the actual axle loads and not the train classification which does not distinguish between empty and full freight trains. The results obtained showed more significant differences between the five groups, each encompassing 20%, than just classifying trains as either freight and passenger. The results though still did not show any large differences between the groups, with the most significantly different group being 80-100% traffic over 50 tonnes, which had higher degradation rates. This shows that the amount of high axle

load trains that traverse a line has little affect on the vertical geometry degradation unless they make up the majority of the traffic. Additionally, the more significant differences found compared to just using the passenger percentages shows that it is important to classify traffic by the loads in transfers onto the track and not the use of the train.

Dirty Traffic Percentage (Usage)

The percentage of dirty traffic varied in the data from 0-40% and was calculated as the percentage of dirty trains that traverse each poskey in respect to usage in EMGT, see Section 3.2.7. It was thought that higher levels of dirty traffic would cause faster degradation due to fines from the load infiltrating the ballast, increasing the rate of ballast fouling. This is important as ballast fouling can cause higher levels of ballast saturation and reduce ballast resilience, increasing the rate of settlement. Results from the performed hypothesis tests showed some differences between the groups, 0-10%, 10-20%, 20-30% and 30-40% with 10-20% and 20-30% being the closest together. The 1-sided hypothesis test results showed that in general poskeys with undergo large amounts of dirty traffic, 30-40% tend to encounter faster vertical degradation rates and on the other hand the poskeys with minimal dirty traffic, 0-10%, experienced slower rates of degradation. There is minimal evidence to suggest if 10-20% or 20-30% degrade faster but a higher percentage of the tests were rejected in favour of 20-30% degrading faster. So it can be said that the percentage of dirty traffic does have a mild effect on the rate of vertical geometry degradation, with higher percentages causing faster degradation.

Superficial Layer

When initially building a railway track, if the bedrock layer in-situ is not a suitable formation to build a railway track on, an additional superficial layer is installed. This layer can consist of many types of geology with the most common on the UK railway network being clay and diamicton followed closely by sand with some places using peat and gravel. It was thought that the superficial layer would have an effect on the rate of vertical degradation due to settlement occurring within the layer, lack of stiffness increasing the load on the bedrock layer and drainage which if poor can increase the saturation within the ballast reducing its resilience causing settlement. The results from the hypothesis tests performed show that there are small differences between the layer choices, with sand performing the best and peat the worst. No layer, clay and diamicton tend to behave similarly with no evidence to be able to say which performs better. The superficial layer may have only a small impact on the rate of vertical geometry degradation due to many reasons. The layer is often thin so not much settlement can occur within the layer. Additionally, as most of the superficial layers have been installed a long time ago, with many track renewals occurring above them, most the settlement may already have occurred within the layer. This is backed up by previous research by Selig and Waters (1994) which shows that the majority of track settlement and hence vertical geometry degradation occurs in the ballast and not the layers below.

Artificial Geology

A small proportion of track that makes up the UK railway network are constructed on top of an artificial layer, placed above the superficial layer and bedrock. This artificial layer is typically made from waste products such as slag and ash to create a cheap layer of high stiffness and permeability. Hypothesis tests have been used to help understand the effect that the use of an artificial layer has on the rate of vertical geometry degradation. The tests showed that there is only a small difference between track sections with an artificial layer installed and those that do not, possibly due the primary settlement occurring within the ballast and not in the formation, (Selig and Waters, 1994). This means that a change in the rate of settlement within the formation layers will have a minimal effect on the overall track settlement and hence the vertical geometry degradation. There were some conflicting results within the tests with some tests returning significant results in favour of the geometry of artificial track degrading slower and some tests returning results in favour of the opposite. There were a slightly higher amount of tests returning results that showed track constructed with an artificial layer degrading slower, but this is not strong enough evidence to be able to say with confidence that this is always true.

Bedrock Geology

The bedrock geology is the initial in-situ geology under the track. If the bedrock geology did not have a high enough stiffness and drainage then further superficial and artificial layers are placed on top before track is constructed. It was thought that a degree of the track settlement and hence vertical geometry degradation would occur in the bedrock layer. In the UK there were many types of bedrock materials with the most common being mudstone, chalk, clay and sandstone. The results obtained from the two tailed hypothesis tests showed that certain materials performed similarly, so these were grouped together. This lead to four significantly different groups consisting of: Group 1; Argillaceous, Dolomitic, Limestone, Psammite and Siltstone, Group 2; Sand, Chalk and Clay, Group 3; Halite and Pebbly Sand and Group 4; Mudstone, Sandstone and Slate. The test results using the grouped bedrocks showed that they were all different with Group 3 and 4 being the closest. Performing one tailed hypothesis tests on the grouped geologies showed that Group 2 contained the lowest rates of degradation followed by 4 and 3 and then 1.

Maximum Speed

Previous research examined as part of the literature review agrees that the track speed is one of the most important factors contributing to the rate of railway track degradation. This is due to the increase in dynamic forces transmitted through the track by faster trains. The results from the performed tests echo the previous research by showing significant differences of the vertical geometry degradation rates between the different track speeds. The initial 2-sided hypothesis tests showed that the speeds could be split into four distinct groups due to similarities between the vertical geometry degradation

rates. These groups were 0-60, 65-70, 75-110 and 115-125 MPH. With these groups it could be shown with a high degree of significance that the higher speed groups tend to degrade slower, which was unexpected as it was thought that the increased dynamic loads caused by the faster trains would lead to higher rates of geometry degradation. The result might be due to the larger sleepers and rails installed on the higher speed tracks reducing the rate of geometry degradation by a greater degree than the increase in dynamic loads increases it. Additionally, the different speed track has different maintenance procedures and thresholds.

Sleeper Type

The sleeper transfers forces from the train, distributing them into the ballast, as well as holding the rail in place and reducing vibrations. Due to this it was thought that the sleeper design would have an impact on the vertical geometry degradation. To test this two-sided hypothesis tests were used comparing sleeper types within similar speed groups, with sleepers showing similarities being grouped together. This lead to three distinct groups. The first one contained small concrete sleepers and larger metal sleeper designs (F23, F24, W600, F27, W560), the second group consisted of large concrete sleepers (F28, G49, G44, G47, F40) and the third group contained timber and smaller metal sleepers (Timber, W402, HH10, Metal, W500). Of these it could be shown by one sided hypothesis tests that the large concrete sleepers performed best followed by the smaller concrete and larger steel sleepers and then the smaller steel and timber sleepers. As the rail type and sleeper choice are closely linked it was decided to check if the rail type was affecting the results of the hypothesis tests. So the datasets were split up by maintenance history, speed and then rail types. The results from this showed no reduction in the significance between the groups unlike when a similar test was performed on rail type when taking account of the sleepers greatly reduced the significance. This proves that the sleeper choice has a large impact on the rate of vertical geometry degradation whereas the rail type does not.

3.5 Vertical Geometry Degradation Model

The analysis in Section 3.4 has shown that many factors influence the rate of vertical geometry degradation. The analysis also showed that there are large variances within the rate, even when the datasets were subset into groups with many similar features. Due to the high level of variability seen in the rates, a stochastic model approach would work best for predicting the rate of degradation of a track section.

To model the rates it was decided to fit distributions which could then be used as part of a larger stochastic asset management model, such as the one proposed in Chapter 6. Previous research by Audley (2014) has shown that Weibull distributions give the best fit for times for track to degrade to certain track quality levels. Linear degradation rates

were used by Audley (2014) to produce the distributions. Due to this the shape of the rates would be the same as the shape of the linear rates and gives evidence to Weibull distributions being applicable. The Weibull distribution is a very versatile distribution and is commonly used in reliability engineering, (Peacock, 2011).

Due to a lack of data, Data A and B were combined to obtain stronger more accurate distribution fits. Additionally, there was not enough data to subset it by all the factors that impact the degradation rates which are discussed in Section 3.4. It was decided from the results of the analysis and the size of possible datasets that the data should be split by the maintenance histories. There was not enough data to maintain starting stoneblowing after separate quantities of tamping i.e. keeping tamp, stone separate from tamp, tamp, stone. So the data was split into the amount of previous tamping operations for the degradation rates after tamping and the amount of previous stoneblowing operations (ignoring the previous amount of tamps) for stoneblowing. To maintain larger datasets it was also decided to group the histories, so they go up in groups of two, hence one group is zero and one previous tamp and another is three and four previous tamps. Assessing the impact of the other factors in Section 3.4, it could be seen that the grouped results for sleepers and track speed were the most statistically different from each other, so these have been chosen as additional factors to split the data by. There were three groups of sleepers, which are described in Table 3.14, fours groups of track speeds (0-60, 70-75, 80-110, 115-125 MPH) and eight maintenance histories (Renewal, Tamping after a renewal (zero previous maintenance), up to nine previous tamps and stoneblowing up to three previous stoneblowing operations (in groups of two)). This results in $4 * 3 * 8 = 96$ Weibull distributions of vertical geometry degradation rates. Descriptive statistics of the datasets as well as the fitted Weibull parameters can be found in Tables 8.225, 8.226, 8.227, 8.228, 8.229, 8.230, 8.231, 8.232, 8.233, 8.234, 8.235 and 8.236 for each of the possible combinations of sleepers, track speeds and maintenance histories. The Weibull distribution requires two parameters, scale (η) and shape (β). The scale is recorded in the same unit of measure as the data fitted, in this case, $nm/EMGT$. Higher scale values lead to a greater range of results and higher average values. The shape is dimensionless, and when equal to one makes the Weibull distribution an exponential, with a consistent failure rate. If beta is less than 1, the failure rate is decreasing, and increasing if more than 1. Value for beta around 3, tend to closely resemble the normal distribution. A beta less than 3 is negatively skewed. Looking at the beta values in Tables 8.225, 8.226, 8.227, 8.228, 8.229, 8.230, 8.231, 8.232, 8.233, 8.234, 8.235 and 8.236 it can be seen that most are between 1-2. This shows that the datasets of degradation rates were slightly negatively skewed with increasing failure rates.

The fitted Weibull distributions of track degradation rates (related to usage), can be used is stochastic based models solved using Monte-Carlo. Degradation rates can be sampled from the distributions with more common rates being sampled more often.

Table 3.14: Sleeper Groups

Group	Included Assets
Small concrete and large steel	F23, F24, W600, F27, W560
Large concrete	F28, G49, G44, G47, F40
Timber and small steel	Timber, W402, HH10, Metal, W500

3.6 Summary

This chapter presents an in-depth analysis of the vertical geometry degradation of railway track. As many aspects of a railway track are interlinked (high speed, high priority track has large rails, large sleepers and more maintenance) a multiple layer approach was taken. This removed the impact of linked factors on the results by removing as many factors as possible before comparing. This proved necessary as some results showed a large impact when analyses on their own but had a much small impact when the factor was singled out, i.e. rail joints when sleeper type was removed.

The main factors identified by the analysis were; track speed, sleeper design, maintenance history and rail joint type. Track speed and joint type were expected as they change the dynamic forces experienced by the track. It is expected that the sleeper choice makes a large difference due to the changes in track structure and stiffness. Another highly influential factor identified was the bedrock. This appeared to be related to the geology's permeability and not the stiffness. The sand superficial layer was shown to perform best showing the importance of good drainage in track design.

Other factors identified as important drivers of geometry degradation include: rail type, electrification, dirty traffic and curvature. In literature the rail type is commonly classified as an important factor, which the first-level analysis agreed with but the multiple layer analysis showed a reduced effect.

Grouping the degradation data not only allowed a multiple layer analysis to be performed but is also required to maintain larger distinct datasets, split by the identified factors, for stochastic modelling like the one presented in Section 3.5.

The distribution based model presented is useful as it takes into account the wide variation of vertical geometry degradation rates seen including the changes in probability of different rates occurring.

Chapter 4

Track Geometry Maintenance Effectiveness

4.1 Introduction

Maintenance is an important aspect of managing any type of asset, from infrastructure to vehicles to computer systems. This is due to the extended life an asset is granted by an effective maintenance procedure, increasing the time between costly and often inconvenient renewals. It is important to understand the different types of possible maintenance, their effectiveness and the optimal implementation time to enable efficient asset management choices to be made. Additionally, to be able to produce an asset management model, these have to be quantifiable and predictable.

There are two main types of maintenance techniques employed on the UK railway infrastructure to correct the track geometry as it worsens with use. These are tamping and stoneblowing, which are described in detail in Section 2.3.4. In this chapter the effectiveness of these two methods are explored, as well as the factors which may influence them.

Previous research, discussed in Section 2.4.3, showed many factors which influence the effectiveness of maintenance, with the most common factor being the initial quality before maintenance (Halcrow, 2012; Caetano and Teixeira, 2016; Stephen M Famurewa and Kumar, 2015). Additional aspects discussed include the track speed (Halcrow, 2012; Audley and Andrews, 2013), track age (Caetano and Teixeira, 2016; Velt, 2007) and the amount of previous maintenance actions (Audley and Andrews, 2013) with the latter two being easily quantifiable measures that are related to the level of ballast fouling.

4.2 Available Data and Processing

To understand the geometry improvement caused by maintenance activities, it was required to be able to quantify the track geometry before and after maintenance. For this, the same data utilised in Chapter 3 was employed. Using the track geometry inspection data, recorded as a Standard Deviation (SD) for each 220 yd poskey on the network, and the times of maintenance occurrences, it was possible to quantify the improvement caused by the maintenance actions. Due to the possibility of additional degradation between the last inspection and the maintenance occurrence or the maintenance action and the first inspection, only maintenance activities which occurred within 60 days of the last recorded inspection and 60 days of the next recorded inspection were utilised. This decision was made to reduce the underestimation of the improvement for sections which degraded further before maintenance or after maintenance, with 60 days chosen due to the generally slow rate of geometry degradation. Additionally, it was found that a large proportion of the maintenance activities underwent an inspection within the 60 days before and after maintenance. This meant that large datasets of geometry improvement were still created, which was important for analysis and building a stochastic asset management model. Processing the data in this way resulted in datasets of geometry improvements for tamping and stoneblowing for each poskey, with these being related to their maintenance history and additional information such as track speed.

4.3 Analysis of the Effectiveness of Track Geometry Maintenance Activities

4.3.1 Methodology

In order to identify the effectiveness and variability of maintenance activities it was decided that visual aids, such as scatter graphs, boxplots and Empirical Cumulative Distribution Function's (ECDF's), would best describe the data, as well as help to identify any factors which affect the degree of improvement. Non-parametric tests were also used in a similar way to how they were used when performing the track geometry degradation analysis in Section 3.4.2, as these allowed any significant differences between datasets to be found, taking into account the size and spread.

4.3.2 Results

Initial Track Quality Effect on the Improvement

It was thought that the possible vertical track geometry improvement caused by tamping and stoneblowing would be related to the initial geometry. This is demonstrated in Figure 4.1, where it can be seen that the worse the initial track quality, the greater

the improvement seen. The linear fits show that the effect tends to be greater for stoneblowing than tamping, with a gradient for stoneblowing of 0.3626 compared to a gradient for tamping of 0.3062. However, it must be noted that the linear fits were poor with r^2 values for tamping and stoneblowing of 0.29 and 0.35 due to the high variability in the data. Despite this the fits still show the general trend. The linearity was also demonstrated when a univariate quadratic line was fitted, as the resultant coefficients for the x^2 part were insignificant. The graphs also demonstrate that stoneblowing is more likely to have a negative effect on the track geometry, with the linear fit showing a negative improvement below an initial vertical geometry of 1.5 mm. Stoneblowing's higher chance of reducing the track quality is also shown by the number of maintenance actions which lead to a negative improvement with 33.8% of the actions reducing the track quality compared to 12.2% of tamping actions. The confidence in this significant difference is high due to the large datasets used of 4,920 tamps and 3,221 stoneblowing operations. There is a greater risk of performing maintenance on already good quality track with these percentages being higher for lower initial track geometry SDs as shown in Table 4.1. This means that the lower the track geometry SD which maintenance is performed at, the higher the probability that the maintenance will have a negative impact. The top line which no points are above in the cone shape of points in Figure 4.1 shows that there is a maximum improvement achievable for all initial qualities governed by a minimum possible vertical geometry SD, which is slightly above 0 mm. The cone shapes also show the increase in variability in the effectiveness of maintenance activities as the initial quality increases.

Due to the strong evidence of the effect the initial quality has on the amount of improvement achieved, it was decided to study the relative/percentage improvement. This was calculated as the percentage reduction in the track geometry SD from the last inspection before the maintenance action:

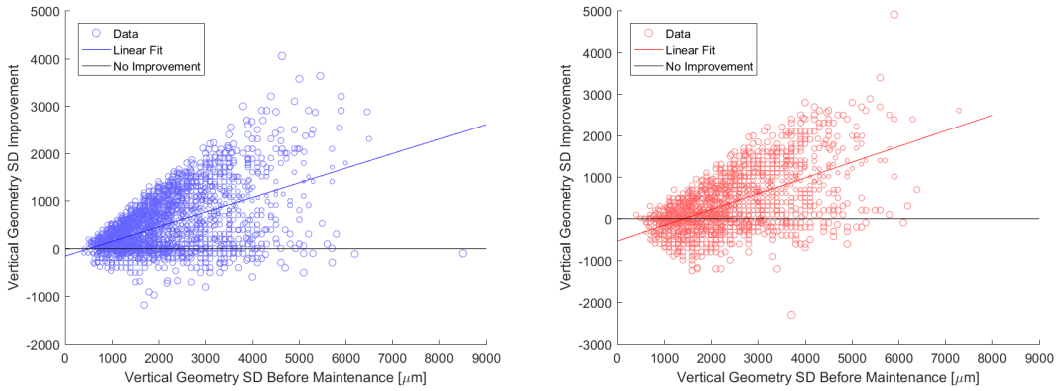
$$\text{Relative SD Improvement} = (\text{Initial SD} - \text{Resultant SD})/\text{Initial SD} \tag{4.1}$$

Plotting the relative improvement against the initial quality removes the cone shape relationship, as seen in Figure 4.2. The non-flat gradient of the linear fits demonstrate that the initial quality still affects the relative improvement, with the steeper gradient of stoneblowing showing that it is more affected than tamping. A main cause of this is the large amount of negative improvements that occur when maintenance is performed on good quality track. To remove the effect the initial quality has on the improvement when analysing further factors, it was decided that the relative improvement should be used instead of the actual improvement. This is important for factors such as track speed, as the speed dictates the track quality at which maintenance is undertaken, with higher track speeds having maintenance performed at lower geometry SDs due to safety and ride comfort. If the relative improvement was not used it is likely that the higher track speeds would show lower improvement, but this would be skewed as a large improvement

is not possible if the track quality is already good.

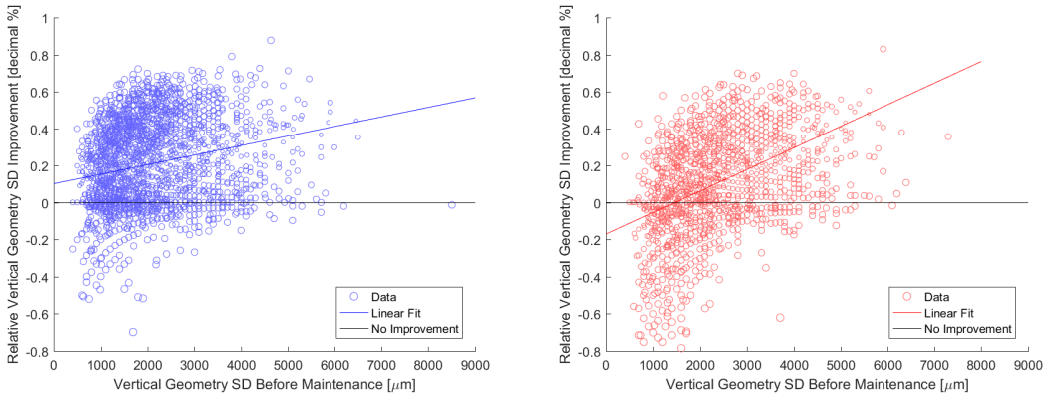
Table 4.1: Percentage of maintenance actions with negative improvement by initial quality

		Initial Track Geometry SD [mm]					
	All Data	0-1	1-2	2-3	3-4	4-5	5-6
Tamping Percentage [%]	12.2	15.6	12.5	10.5	10.0	9.3	11.8
Stoneblowing Percentage [%]	33.8	61.1	45.1	24.1	16.7	14.2	8.6
Tamping Dataset Count	4920	651	2622	1119	381	108	34
Stoneblowing Dataset Count	3221	134	1467	1076	377	127	35



(a) Tamping (b) Stoneblowing

Figure 4.1: Initial quality against improvement from track maintenance



(a) Tamping (b) Stoneblowing

Figure 4.2: Initial quality against relative improvement from track maintenance

Track Speed

Due to previous maintenance effectiveness models including track speed, such as those developed by Audley and Andrews (2013) and Halcrow (2012), it was decided to analyse its affect. This was accomplished by splitting the datasets of relative geometry improvement from tamping and stoneblowing according to different speed groups. The low speeds, below 50 MPH, were grouped to maintain large datasets. Descriptive statistics were used to compare the groups, with these outlined in Table 4.2, as well as boxplots and ECDF's as shown in Figure 4.3. The non-parametric, K-S and Mann Whitney U tests were also utilised to show any significant differences between datasets.

Tamping

Looking at the mean and median values from Table 4.2 as well as the boxplot and ECDF for tamping in Figure 4.3, there seems to be minimal difference between the speed groups 5-60 and 75-95 MPH. This is backed up by the mean 2-tailed K-S and Mann Whitney U test p-values in Table 4.3. The non-parametric tests, show that speeds 65-70 and 115-125 MPH are significantly different from the other groups, and 100-110 MPH is slightly different from the rest. It can be seen from the boxplot in Figure 4.3 that tamping in the speed group 65-70 MPH tends to have a lower relative improvement, whereas 100-110 is higher than the rest and then 115-125 is significantly higher than this. The cause of this disparity is unknown but maybe due to differences in assets or maintenance regimes.

This increase in improvement can be seen when plotting the average for each speed group against the middle speed of the group, seen in Figure 4.4a. This shows an upwards trend, where tamping on higher speed tracks leads to a greater relative improvement. The gradients of the mean and median fit are 0.0005 and 0.0006, with r^2 values of 0.402 and 0.284 respectively. Despite the low r^2 it was felt that the linear fit gave a general trend linking track speed to tamping's relative improvement. The effect was calculated back into an actual improvement to get an idea of speed's effect. For a track section with an initial vertical geometry SD of 3.00 mm, the resultant SD would be 2.56 mm for a track speed of 40 MPH and 2.40 mm for a track speed of 125 MPH, if the median line was used. Looking at the very small difference between these values, taking into account the large variability demonstrated in the boxplots, it can be said that the track speed has little impact on the improvement gained from maintenance within most speed bands. The non-parametric tests did give evidence to suggest that tamping performed on track of speeds 115-125 MPH tends to give a significantly greater relative improvement than on other speeds. This is the highest speed track which tends to have heavier sleepers and rails which maybe impacting the possible quality. Additionally, more speed lines also tend to be higher priority so are given increased levels of maintenance.

Stoneblowing

The effectiveness of stoneblowing does not show any dependency on the track speed,

with the non-parametric test results in Table 4.4 showing no significance between any of the speed groups. This can also be seen in the boxplot in Figure 4.3 where all the bars are very similar and the ECDF, where the lines are very close to each other. The general link between the track speed and stoneblowing effectiveness, using the average relative improvement in each speed group, is shown to be negative in Figure 4.4b. This negative link has gradients of -0.0003 for the mean and -0.0003 for the median, with low r^2 values of 0.201 and 0.404. The low gradient values show how minimal the effect of track speed is on stoneblowing effectiveness. A track section with an initial SD of 3.00 mm for a track speed of 40 MPH, using the median values, would have a resultant SD after stoneblowing of 2.68 mm, whereas if the track speed was 125 MPH it would be 2.77 mm. The small difference between these values again shows that the track speed tends to have no effect on the effectiveness of stoneblowing.

Table 4.2: Descriptive statistics of the relative improvement from maintenance by track speed

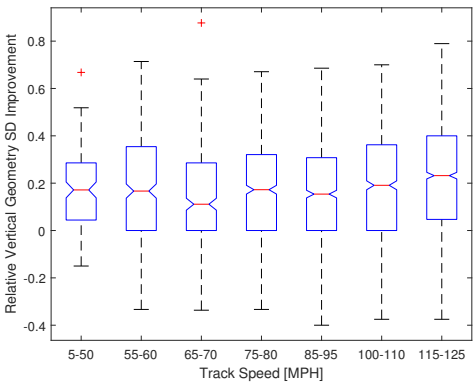
			Track Speed [MPH]						
		All Data	5-50	55-60	65-70	75-80	85-95	100-110	115-125
Tamping	Mean	0.197	0.174	0.189	0.151	0.183	0.171	0.198	0.231
	Median	0.182	0.171	0.167	0.111	0.172	0.154	0.191	0.232
	N	4920	128	343	316	744	903	909	1577
Stoneblowing	Mean	0.099	0.116	0.131	0.097	0.080	0.087	0.106	0.099
	Median	0.091	0.111	0.115	0.098	0.077	0.073	0.091	0.089
	N	3221	200	386	394	680	632	664	265

Table 4.3: Mean 2 Tailed K-S and Mann Whitney U Test p-values for tamping relative improvement by track speed

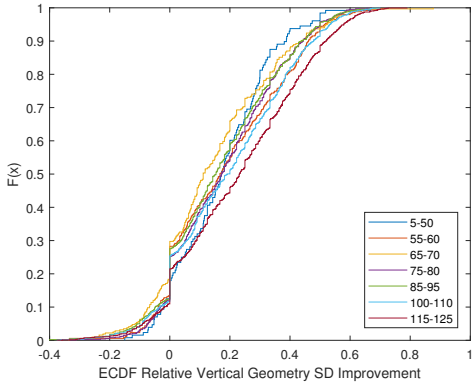
	5-50	55-60	65-70	75-80	85-95	100-110	115-125
5-50	NA	0.40	0.06	0.49	0.46	0.12	0.00
55-60	0.40	NA	0.03	0.57	0.32	0.33	0.00
65-70	0.06	0.03	NA	0.01	0.10	0.00	0.00
75-80	0.49	0.57	0.01	NA	0.35	0.07	0.00
85-96	0.46	0.32	0.10	0.35	NA	0.01	0.00
100-110	0.12	0.33	0.00	0.07	0.01	NA	0.00
115-125	0.00	0.00	0.00	0.00	0.00	0.00	NA

Table 4.4: Mean 2 Tailed K-S and Mann Whitney U Test p-values for stoneblowing relative improvement by track speed

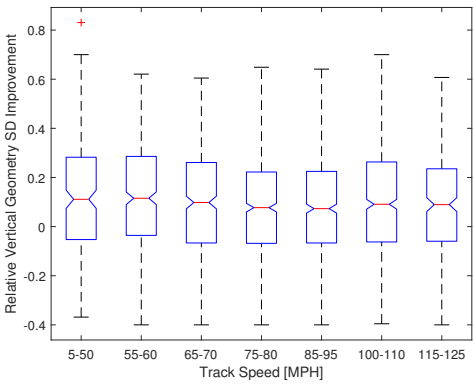
	5-50	55-60	65-70	75-80	85-95	100-110	115-125
5-50	NA	0.72	0.70	0.52	0.56	0.76	0.72
55-60	0.72	NA	0.52	0.50	0.50	0.53	0.54
65-70	0.70	0.52	NA	0.59	0.69	0.86	0.98
75-80	0.52	0.50	0.59	NA	0.81	0.53	0.61
85-95	0.56	0.50	0.69	0.81	NA	0.57	0.68
100-110	0.76	0.53	0.86	0.53	0.57	NA	0.90
115-125	0.72	0.54	0.98	0.61	0.68	0.90	NA



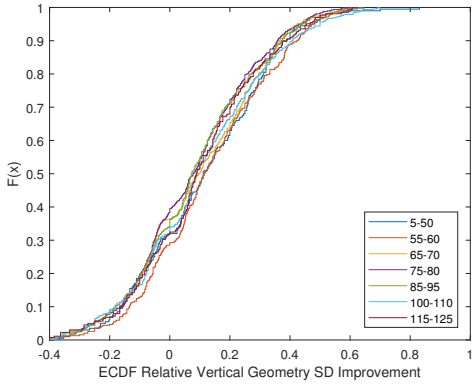
(a) Boxplot of Tamping



(b) ECDF of Tamping

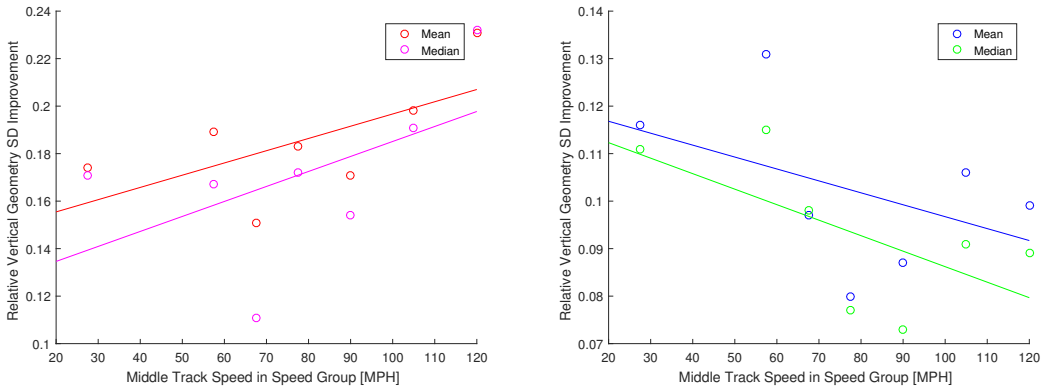


(c) Boxplot of Stoneblowing



(d) ECDF of Stoneblowing

Figure 4.3: Maintenance relative improvement by track speed



(a) Tamping (b) Stoneblowing

Figure 4.4: Average relative improvement of maintenance for different track speeds

Previous number of Tamping and Stoneblowing Activities

Maintenance history is another aspect discussed in the literature with regard to its effect on the improvement following maintenance, (Audley and Andrews, 2013). This is due to the link between maintenance history and the level of ballast fouling, with track which has undergone multiple maintenance activities having higher levels of fouling. Caetano and Teixeira (2016) and Velt (2007) also include the affect of ballast fouling within their proposed models, using the track age instead of the maintenance history. It is thought that the inclusion of fines in the ballast negatively impacts the effectiveness of maintenance.

It was decided that the maintenance history would better represent the amount of ballast fouling than age due to it having a link with track age as well as it being a direct cause of fouling within the ballast due to ballast breakup. As discussed in Chapter 3, the datasets were maintained as either Data A or Data B, with Data A containing recently renewed track with a full recorded maintenance history and Data B containing older track sections with estimated maintenance histories. These have been kept separate so the added assumptions and likelihood of an error in the estimate of the maintenance history for the data in Data B did not impact the high quality of Data A, where the entire history of maintenance and inspections were known.

Tamping

Descriptive statistics were calculated for the different maintenance histories which tamping was performed at within the data, with these being shown in Table 4.9. The relative improvement has been used to reduce the likelihood of the results being skewed by the effect the initial quality has on the level of improvement. The datasets have been shown visually using boxplots and ECDF's in Figure 4.5. Looking at Data A in Table 4.9, there were more than twenty quantified improvements for the maintenance history groups from the first tamp after a renewal to after three previous tamps, allow these datasets to be

used for non-parametric tests with confidence. Non-parametric tests were used to show any significance between these datasets with the results recorded in Table 4.5. The results showed that whilst there is enough evidence to suggest that the effect of tamping after a renewal is significantly different to the effect after one tamp, there are no other significant differences between the datasets. This is backed up by the medians and means of the histories in Table 4.9, with tamping after a renewal showing the smallest relative improvement and tamping after one tamp showing the highest (of the groups with at least 10 data points). The average improvements following two and three tamps were between the two. The boxplot, in Figure 4.5a, shows that as well as having the lowest relative improvement, tamping after a renewal has the largest range.

Due to Data B containing older track sections, it is possible to see the effect of an increased amount of previous maintenance actions, all the way up to ten previous tamps, all with datasets above 200 records. The mean non-parametric p-values are shown in Table 4.6. The results show that from four previous tamps to ten tamps there are no significant differences in the relative improvement. This is echoed by the boxplot, in Figure 4.5c, where the bars are all very similar as are the averages in Table 4.9. The results also show that there are no significant differences in the improvement seen between one tamp and three tamps. The two separate groups, 1T-3T and 4T-10T are shown to have some significant differences by the low average p-values in the results of the non-parametric tests in Table 4.6. The group 1T-3T shows slightly greater relative improvements, but from looking at the boxplots it can be also be seen that 1T-3T have a higher spread. Looking at the averages between Data A and B, it can be seen that for the groups that appear in both (1T-3T) that the averages are very close. Therefore, based on the combination of the descriptive statistics, boxplots, ECDF's and non-parametric hypothesis tests for Data A and B it seems that tamping is most effective for track that has previously undergone 1T-3T and less effective for track that has undergone 4T-10T. This is due to ballast breakdown from previous maintenance increasing ballast fouling which reduces the effectiveness of tamping. The results also show the least effective is tamping performed for the first time following a renewal. This was not expected as the railway track assets are still relatively new, with minimal ballast fouling. It is likely this is due to tamping being performed too earlier after a renewal when the track geometry is still good. Figure 4.2a shows that on average tamping performed on low track geometry SD shows smaller relative improvements.

The reduction in improvement following tamping that is observed as the number of tamps performed increases can be seen in Figure 4.6a, where there is a downward trend in the linear fits. Despite the extra uncertainties in Data B, the much larger datasets give a good degree of confidence in the average results, with Data A lacking the large datasets required to be confident in an average value being representative of a larger population. The downward trend is gradual with Data B median values having a gradient of -0.0041 and the mean values a gradient of -0.0038, with r^2 values of 0.549 and 0.302. Despite the low r^2 values it was felt that the linear fit gave an idea of how the maintenance

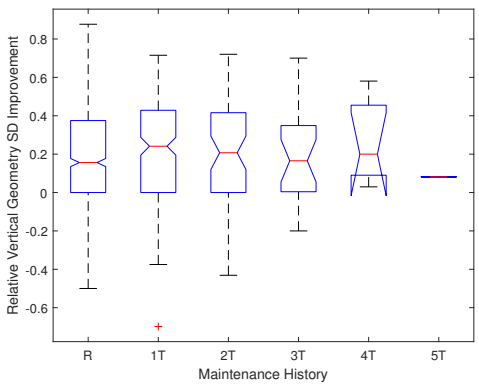
history affected the improvement from tamping, with the results showing that each previous tamping operation only reduced the relative improvement by around 0.004. If a track section was tamped after one previous tamp at an initial vertical geometry SD of 3.00 mm, using the median line this would result in a SD after maintenance of 2.41 mm, whereas after ten previous tamps it would be 2.52 mm. This small difference gives evidence of the minimal effect the maintenance history has on the effectiveness of tamping. This, due to the link between ballast fouling and the amount of maintenance actions, shows that the quality of ballast has a minimal impact on the improvement found from tamping, for the levels of ballast fouling found on UK railway infrastructure. Despite having only a small effect, the non-parametric hypothesis tests showed the existence of three groups of similar performing maintenance histories: tamp directly after a renewal, one to three previous tamps and more than 3 previous tamps.

Table 4.5: Mean 2 Tailed K-S and Mann Whitney U Test p-values for tamping relative improvement by the amount of previous tamps [Data A]

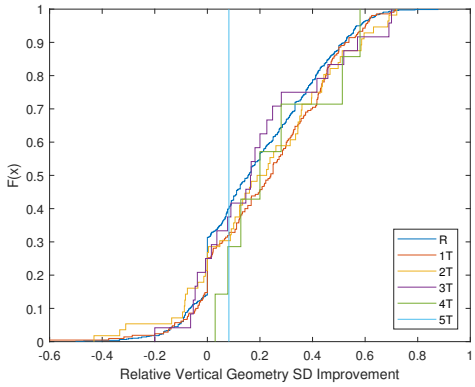
	R	1T	2T	3T
R	NA	0.02	0.46	0.96
1T	0.03	NA	0.79	0.35
2T	0.46	0.79	NA	0.73
3T	0.96	0.35	0.73	NA

Table 4.6: Mean 2 Tailed K-S and Mann Whitney U Test p-values for tamping relative improvement by the amount of previous tamps [Data B]

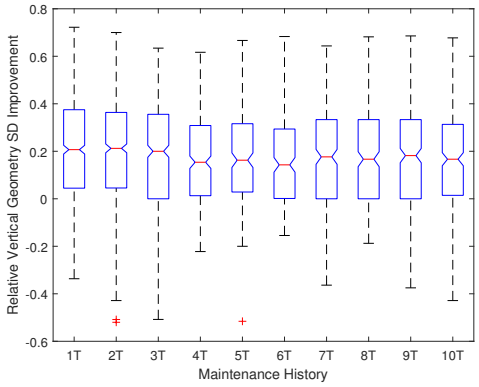
	1T	2T	3T	4T	5T	6T	7T	8T	9T	10T
1T	NA	0.89	0.31	0.00	0.10	0.01	0.18	0.05	0.02	0.03
2T	0.89	NA	0.32	0.00	0.12	0.00	0.12	0.08	0.03	0.01
3T	0.32	0.32	NA	0.05	0.53	0.04	0.62	0.31	0.14	0.11
4T	0.003	0.00	0.05	NA	0.55	0.82	0.39	0.57	0.78	0.77
5T	0.10	0.12	0.53	0.55	NA	0.28	0.94	0.60	0.69	0.63
6T	0.01	0.00	0.04	0.82	0.28	NA	0.23	0.46	0.57	0.61
7T	0.18	0.12	0.62	0.39	0.94	0.23	NA	0.75	0.57	0.49
8T	0.05	0.08	0.31	0.57	0.60	0.46	0.75	NA	0.91	0.90
9T	0.02	0.03	0.14	0.78	0.69	0.57	0.57	0.91	NA	0.98
10T	0.03	0.01	0.11	0.77	0.63	0.61	0.49	0.90	0.98	NA



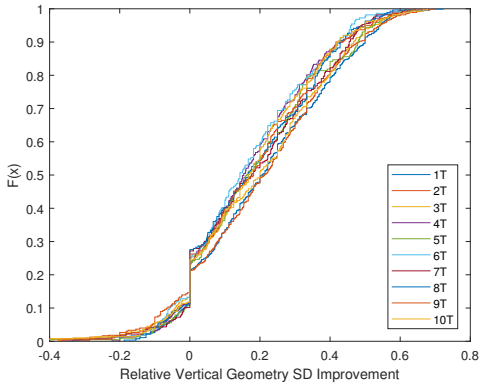
(a) Boxplot of Data A



(b) ECDF of Data A



(c) Boxplot of Data B



(d) ECDF of Data B

Figure 4.5: Maintenance relative improvement the amount of previous tamps before tamping

Stoneblowing

Similarly to tamping, the effect of a track section's maintenance history on the improvement obtained by stoneblowing has been explored. Due to the possibility of stoneblowing occurring after any number of previous tamps there were more maintenance history combinations to consider than when considering tamping alone. The different combinations, dataset sizes and averages are described in Table 4.10. The datasets are visually demonstrated in the boxplot and ECDF in Figure 4.8. It was not possible to use Data A for the analysis of stoneblowing, as there was minimal data due to stoneblowing being primarily used on older track. The use of non-parametric tests, K-S and Mann-Whitney U, produced the p-values in Table 4.11.

Looking first at the effect of the first stoneblowing action after tamping, the results showed no significant difference in its relative effectiveness for maintenance histories 1T-6T and 8T-9T but the results for 7T showed an increase in effectiveness and an even larger increase for 10T. There is no mechanistic reason for these anomalies. Looking at the general trend in Figure 4.6b, it can be seen that the more tamps that have been performed prior to stoneblowing, the greater the effect of the first stoneblow. The linear fits for the median and mean values had gradients of 0.0095 and 0.0115 with r^2 values of

0.456 and 0.476. If a track section had an initial SD of 3.00 mm, using the median line, the resultant SD after stoneblowing would be 2.86 mm if the previous maintenance only consisted of one tamp and 2.55 mm if after 10 tamps. The amount of previous tamping action appears to have the opposite effect on the first stoneblow compared to performing another tamp, additionally the effect of more previous maintenance actions is greater for stoneblowing (higher gradients), but neither is a significant amount, especially when the variability of the datasets (seen in the boxplot in Figure 4.8a) is taken into account.

Looking at the effectiveness of multiple stoneblowing operations it can be seen that there is a large reduction in improvement over multiple actions. This is seen by the averages in Table 4.10 where the mean and median relative improvement for 1S is 0.092 and 0.083 but these drop to 0.014 and 0.059 for 2S. A similar trend is seen from 1T1S to 1T2S, 2T1S to 2T2S, 3T1S to 3T2S, etc. This shows that stoneblowing is only effective when used twice on the same track section, with a third stoneblowing operation, no matter how many tamps were performed previously, leading to very little improvement. It is thought this is due to the amount of small stones beneath the sleepers becoming too numerous to support the load (does not have the skeleton strength of larger ballast stones), resulting in the sleeper quickly moving back to its position prior to stoneblowing. The results in Table 4.10 also show that the second stoneblowing operation is on average slightly more effective than the first with the relative improvement from 1T1S being greater than 1T, 2T1S than 2T, 4T1S than 4T, etc. The average p-values in Table 4.11 show that the number of previous tamping operations has little effect on the improvement from stoneblowing, with very few significant differences between the datasets which had undergone the same number of stoneblowing operations. Due to this it was decided to group together the maintenance histories by the number of previous stoneblowing operations, ignoring the number of previous tamps. This led to the boxplot and ECDF in Figure 4.7 and averages in Table 4.7. Looking at these it is very apparent that past the second stoneblowing actions the improvement seen is greatly reduced, with zero and one previous stoneblowing actions leading to similar improvements. The median relative improvement more than halves between a second and third stoneblowing operation, with the mean being nearly one ninth. The results of the non-parametric hypothesis tests seen in Table 4.8, show a similar conclusion, where there are similarities between 0S and 1S, as well as 2S and 3S but a highly significant difference between these groups.

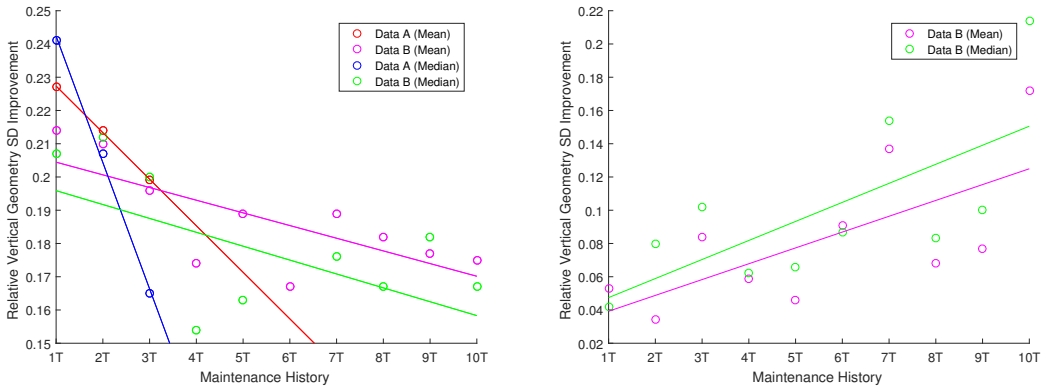
From the results it can be said that the number of previous tamping actions has little effect on the improvement caused by stoneblowing, but the effectiveness of stoneblowing is reduced greatly after the second stoneblowing operation.

Table 4.7: Descriptive statistics of the relative improvement from stoneblowing by the amount of previous stoneblows, ignoring the previous amount of tamps

		No. of previous stoneblows			
	All Data	0S	1S	2S	3S
Mean	0.075	0.082	0.089	0.010	0.007
Median	0.089	0.095	0.105	0.047	0.040
N	3018	1309	1295	335	79

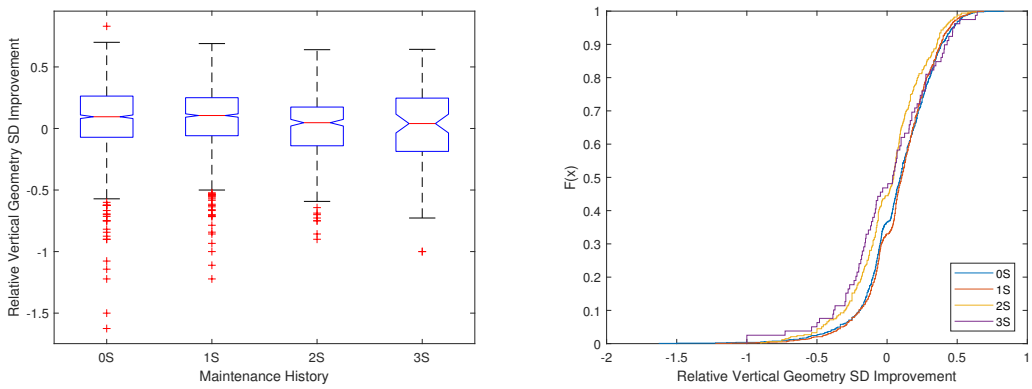
Table 4.8: Mean 2 Tailed K-S and Mann Whitney U Test p-values for stoneblowing relative improvement by the amount of previous stoneblowing actions ignoring the amount of previous tamps [Data B]

	0S	1S	2S	3S
0S	NA	0.27	0.00	0.02
1S	0.27	NA	0.00	0.01
2S	0.00	0.00	NA	0.81
3S	0.02	0.01	0.81	NA



(a) Tamping (b) Stoneblowing

Figure 4.6: Average relative improvement of maintenance after a maintenance history of just tamping



(a) Boxplot

(b) ECDF

Figure 4.7: Maintenance relative improvement by number of previous stoneblows before further stoneblowing

Table 4.9: Descriptive statistics of the relative improvement from tamping by the amount of previous tamps

			Maintenance History Before Tamping									
		All Data	R	1T	2T	3T	4T	5T	6T	7T	8T	9T	10T
Data A	Mean	0.199	0.190	0.227	0.214	0.199	0.258	0.082					
	Median	0.167	0.156	0.241	0.207	0.165	0.200	0.082					
	N	1076	778	210	56	24	7	1					
Data B	Mean	0.194		0.214	0.210	0.196	0.174	0.189	0.167	0.189	0.182	0.177	0.175
	Median	0.182		0.207	0.212	0.200	0.154	0.163	0.143	0.176	0.167	0.182	0.167
	N	3893		801	719	484	317	215	219	276	272	301	289

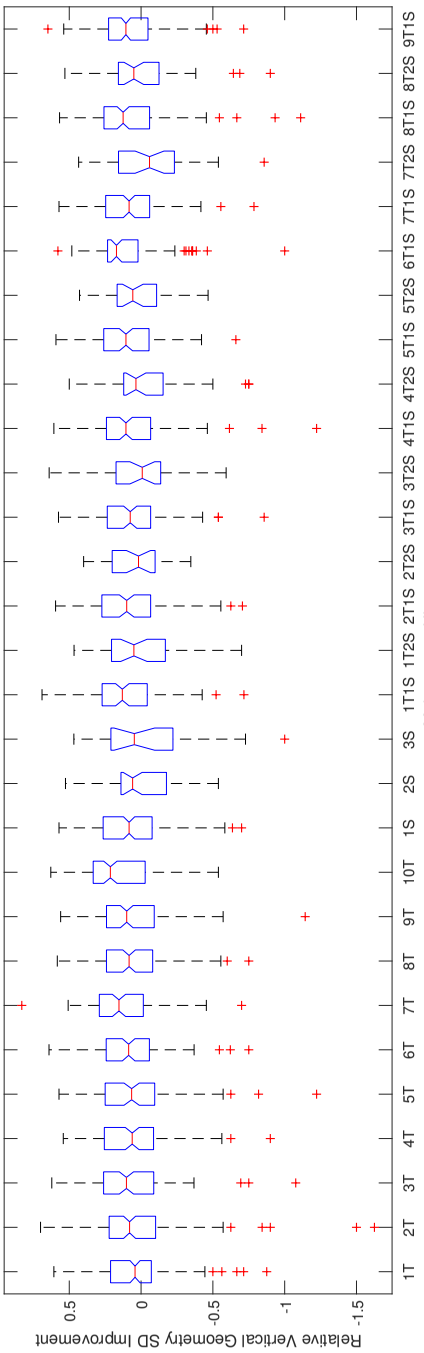
Table 4.10: Descriptive statistics of the relative improvement from stoneblowing by the amount of previous tamps and stoneblows [Data B]

	All Data	1T	2T	3T	4T	5T	6T	7T	8T	9T	10T	1S	2S	3S	1T1S
Mean	0.076	0.053	0.034	0.084	0.059	0.046	0.091	0.137	0.068	0.077	0.172	0.092	0.014	-0.042	0.118
Median	0.090	0.042	0.080	0.102	0.062	0.066	0.087	0.154	0.083	0.100	0.214	0.083	0.059	0.047	0.130
N	2942	162	147	122	126	104	116	109	133	140	150	189	57	23	111

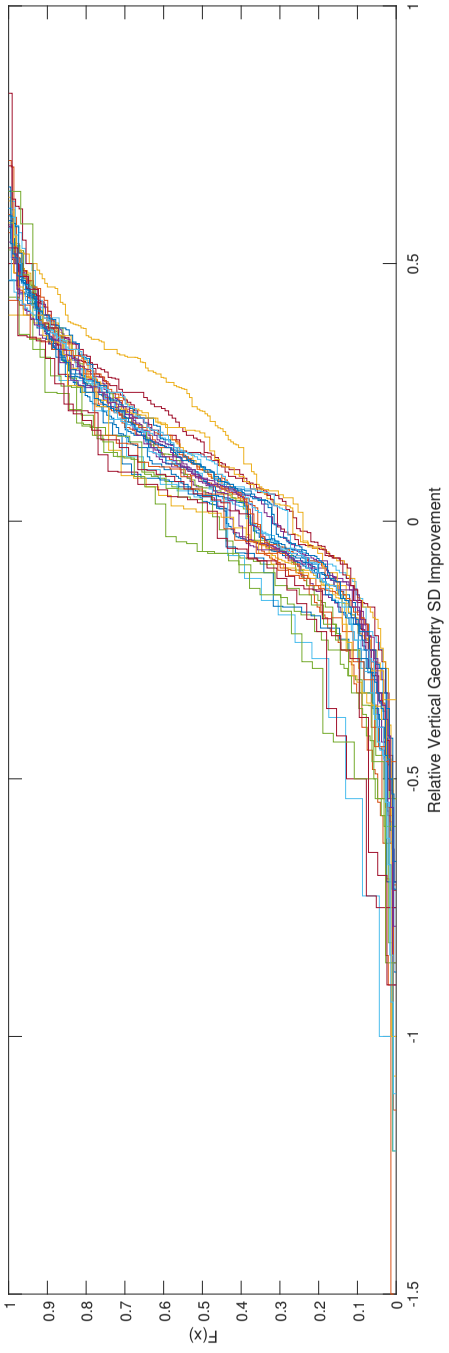
	1T2S	2T1S	2T2S	3T1S	3T2S	4T1S	4T2S	5T1S	5T2S	6T1S	7T1S	7T2S	8T1S	8T2S	9T1S
Mean	0.021	0.089	0.043	0.067	0.012	0.072	-0.033	0.100	0.040	0.094	0.090	-0.062	0.085	-0.001	0.086
Median	0.050	0.100	0.018	0.075	-0.008	0.105	0.036	0.105	0.058	0.171	0.083	-0.059	0.125	0.050	0.106
N	41	109	31	106	32	121	39	115	36	79	141	37	143	42	181

Table 4.11: Mean 2 Tailed K-S and Mann Whitney U Test p-values for tamping relative improvement by the amount of previous tamps and stoneblows [Data B]

	1T	2T	3T	4T	5T	6T	7T	8T	9T	10T	1S	2S	3S	1T1S	1T2S	2T1S	2T2S	3T1S	3T2S	4T1S	4T2S	5T1S	5T2S	6T1S	7T1S	7T2S	8T1S	8T2S	9T1S
1T	NA	0.60	0.22	0.84	0.61	0.45	0.00	0.54	0.27	0.00	0.18	0.19	0.31	0.02	0.26	0.20	0.61	0.67	0.26	0.21	0.26	0.09	0.83	0.02	0.09	0.06	0.03	0.48	0.06
2T	0.60	NA	0.34	0.77	0.91	0.25	0.03	0.60	0.40	0.00	0.21	0.27	0.52	0.04	0.38	0.31	0.36	0.56	0.37	0.53	0.15	0.34	0.76	0.10	0.23	0.06	0.25	0.52	0.22
3T	0.22	0.34	NA	0.51	0.53	0.77	0.10	0.50	0.72	0.00	0.90	0.05	0.24	0.21	0.13	0.78	0.22	0.67	0.18	0.73	0.05	0.79	0.28	0.42	0.64	0.04	0.51	0.13	0.66
4T	0.84	0.77	0.51	NA	0.91	0.36	0.02	0.77	0.51	0.00	0.27	0.29	0.41	0.06	0.31	0.55	0.55	0.75	0.20	0.47	0.18	0.25	0.69	0.09	0.24	0.07	0.11	0.46	0.20
5T	0.61	0.91	0.53	0.91	NA	0.45	0.12	0.86	0.65	0.00	0.39	0.28	0.49	0.11	0.42	0.64	0.42	0.80	0.21	0.70	0.16	0.46	0.64	0.19	0.42	0.06	0.28	0.45	0.47
6T	0.45	0.25	0.77	0.36	0.45	NA	0.08	0.63	0.61	0.00	0.93	0.05	0.13	0.20	0.09	0.83	0.20	0.68	0.07	0.84	0.09	0.57	0.41	0.22	0.66	0.02	0.29	0.25	0.60
7T	0.00	0.03	0.10	0.02	0.12	0.08	NA	0.06	0.08	0.11	0.12	0.01	0.07	0.66	0.01	0.20	0.02	0.04	0.01	0.14	0.00	0.39	0.06	0.40	0.20	0.00	0.44	0.02	0.10
8T	0.54	0.60	0.50	0.77	0.86	0.63	0.06	NA	0.81	0.00	0.66	0.13	0.24	0.15	0.16	0.62	0.29	0.92	0.17	0.63	0.08	0.45	0.72	0.11	0.52	0.04	0.16	0.44	0.48
9T	0.27	0.40	0.72	0.51	0.65	0.61	0.08	0.81	NA	0.00	0.60	0.06	0.21	0.20	0.10	0.76	0.17	0.81	0.18	0.97	0.03	0.73	0.47	0.28	0.67	0.03	0.59	0.22	0.65
10T	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	NA	0.00	0.00	0.02	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00
1S	0.18	0.21	0.90	0.27	0.39	0.93	0.12	0.66	0.60	0.00	NA	0.03	0.12	0.23	0.06	0.88	0.17	0.60	0.12	0.83	0.05	0.72	0.41	0.29	0.89	0.01	0.57	0.25	0.84
2S	0.19	0.27	0.05	0.29	0.28	0.05	0.01	0.13	0.06	0.00	0.03	NA	0.86	0.01	0.94	0.07	0.60	0.09	0.98	0.06	0.75	0.06	0.74	0.01	0.04	0.49	0.03	0.83	0.04
3S	0.31	0.52	0.24	0.41	0.49	0.13	0.07	0.24	0.21	0.02	0.12	0.86	NA	0.07	0.81	0.18	0.61	0.23	0.91	0.31	0.93	0.22	0.70	0.15	0.16	0.66	0.21	0.82	0.18
1T1S	0.02	0.04	0.21	0.06	0.11	0.20	0.66	0.15	0.20	0.03	0.23	0.01	0.07	NA	0.03	0.53	0.04	0.08	0.01	0.38	0.01	0.61	0.12	0.71	0.40	0.00	0.65	0.04	0.47
1T2S	0.26	0.38	0.13	0.31	0.42	0.09	0.01	0.16	0.10	0.00	0.06	0.94	0.81	0.03	NA	0.13	0.70	0.18	0.87	0.12	0.67	0.09	0.76	0.05	0.09	0.46	0.05	0.93	0.08
2T1S	0.20	0.31	0.78	0.55	0.64	0.83	0.20	0.62	0.76	0.01	0.88	0.07	0.18	0.53	0.13	NA	0.16	0.59	0.10	0.88	0.03	0.95	0.33	0.45	0.93	0.03	0.71	0.15	0.84
2T2S	0.61	0.36	0.22	0.55	0.42	0.20	0.02	0.29	0.17	0.00	0.17	0.60	0.61	0.04	0.70	0.16	NA	0.37	0.72	0.18	0.60	0.10	0.77	0.05	0.13	0.19	0.06	0.86	0.10
3T1S	0.67	0.56	0.67	0.75	0.80	0.68	0.04	0.92	0.81	0.00	0.60	0.09	0.23	0.08	0.18	0.59	0.37	NA	0.19	0.71	0.13	0.44	0.71	0.16	0.42	0.05	0.27	0.32	0.36
3T2S	0.26	0.37	0.18	0.20	0.21	0.07	0.01	0.17	0.18	0.00	0.12	0.98	0.91	0.01	0.87	0.10	0.72	0.19	NA	0.13	0.90	0.06	0.71	0.03	0.06	0.66	0.05	0.89	0.05
4T1S	0.21	0.53	0.73	0.47	0.70	0.84	0.14	0.63	0.97	0.00	0.83	0.06	0.31	0.38	0.12	0.88	0.18	0.71	0.13	NA	0.04	0.71	0.36	0.30	0.88	0.01	0.66	0.20	0.95
4T2S	0.26	0.15	0.05	0.18	0.16	0.09	0.00	0.08	0.03	0.00	0.05	0.75	0.93	0.01	0.67	0.03	0.60	0.13	0.90	0.04	NA	0.04	0.43	0.01	0.06	0.71	0.01	0.75	0.02
5T1S	0.09	0.34	0.79	0.25	0.46	0.57	0.39	0.45	0.73	0.01	0.72	0.06	0.22	0.61	0.09	0.95	0.10	0.44	0.06	0.71	0.04	NA	0.29	0.59	0.81	0.01	0.97	0.14	0.79
5T2S	0.83	0.76	0.28	0.69	0.64	0.41	0.06	0.72	0.47	0.00	0.41	0.74	0.70	0.12	0.76	0.33	0.77	0.71	0.71	0.36	0.43	0.29	NA	0.07	0.36	0.35	0.17	0.82	0.42
6T1S	0.02	0.10	0.42	0.09	0.19	0.22	0.40	0.11	0.28	0.01	0.29	0.01	0.15	0.71	0.05	0.45	0.05	0.16	0.03	0.30	0.01	0.59	0.07	NA	0.38	0.00	0.66	0.02	0.22
7T1S	0.09	0.23	0.64	0.24	0.42	0.66	0.20	0.52	0.67	0.00	0.89	0.04	0.16	0.40	0.09	0.93	0.13	0.42	0.06	0.88	0.06	0.81	0.36	0.38	NA	0.01	0.62	0.18	0.93
7T2S	0.06	0.06	0.04	0.07	0.06	0.02	0.00	0.04	0.03	0.00	0.01	0.49	0.66	0.00	0.46	0.03	0.19	0.05	0.66	0.01	0.71	0.01	0.35	0.00	0.01	NA	0.00	0.30	0.01
8T1S	0.03	0.25	0.51	0.11	0.28	0.29	0.44	0.16	0.59	0.00	0.57	0.03	0.21	0.65	0.05	0.71	0.06	0.27	0.05	0.66	0.01	0.97	0.17	0.66	0.62	0.00	NA	0.08	0.61
8T2S	0.48	0.52	0.13	0.46	0.45	0.25	0.02	0.44	0.22	0.00	0.25	0.83	0.82	0.04	0.93	0.15	0.86	0.32	0.89	0.20	0.75	0.14	0.82	0.02	0.18	0.30	0.08	NA	0.13
9T1S	0.06	0.22	0.66	0.20	0.47	0.60	0.10	0.48	0.65	0.00	0.84	0.04	0.18	0.47	0.08	0.84	0.10	0.36	0.05	0.95	0.02	0.79	0.42	0.22	0.93	0.01	0.61	0.13	NA



(a) Boxplot of Data B



(b) ECDF of Data B

Figure 4.8: Maintenance relative improvement by maintenance history before stoneblowing

4.3.3 Summary

The results of the analysis showed that the initial track quality has a large impact on the improvement seen from performing maintenance, with the improvement achieved from stoneblowing being shown to be more dependent on the initial quality than tamping. Stoneblowing was also found to be the riskier maintenance method with a greater amount of historic operations leading to a negative improvement, particularly on track with good geometry. It was found that the worse the initial quality, the greater the improvement but also the spread; by using a relative/percentage improvement this was normalised.

When track speed was analysed it could be seen that it had little effect on the improvement following tamping, with a slight increase in the relative improvement being observed for higher speed track. This result might be due to higher quality maintenance being performed on higher speed sections as these tend to be busier and more critical lines. Track speed had an even smaller effect on the improvement following stoneblowing, with no significant differences between any of the speed groups.

Exploring the effect of the maintenance history showed a small impact on the effectiveness of maintenance. The number of previous tamps caused a slight reduction in the improvement seen from tamping with three distinct groups being seen; directly after a renewal, after one to three tamps and after more than three tamps. One to three tamps showed the greatest relative improvement followed by more than three tamps. Stoneblowing was shown not to be affected by the amount of previous tamping operations. There was strong evidence to suggest that whilst the improvement seen from the first two stoneblowing operations, on any combination of previous tamps, was consistent, the improvement seen by a third and fourth was greatly reduced. This evidence makes it hard to suggest ever performing a third stoneblowing operation on a track section, with a renewal probably being required by this point. The reduction in improvement seen is probably due to a build up of small stones beneath the sleepers from previous stoneblowing activities.

Due to these results it was decided that when modelling the improvement achieved from performing maintenance that the significance of tamping higher speed tracks leading to a greater improvement needs to be taken into account. Additionally, the three distinct maintenance histories for tamping should be separated (first tamp after a renewal, after 1-3 tamps and more than 3 tamps), with it also being important to model the first two stoneblowing operations differently to the third and fourth.

4.4 Modelling the Effectiveness of Track Geometry Maintenance

Due to maintenance actions not returning the track state back to a perfect condition, it is important to be able to predict the quality level after maintenance in order to optimise

the timing and choice of maintenance, or decide whether to renew instead. It has been demonstrated that there is a linear link between the initial vertical track geometry and the resultant geometry, with the variability increasing with the initial quality.

From analysing the improvements from tamping actions it was seen that there were three groups of maintenance actions which behaved significantly different; first tamp after a renewal, tamping after one to three previous tamps and after more than three tamps. Additionally, tamping track with speeds under 115 MPH had significantly different results than track with speeds over and including 115 MPH. This leads to six different groups of tamping. There were also shown to be two significant groups for stoneblowing: stoneblowing after zero and one previous stoneblowing actions or after more than one.

The linear fit between the initial and resultant quality was seen to be a good basis for a model since the linearity was proved when univariate quadratic lines were fitted, where the resultant coefficients of x^2 were seen to be insignificant. A linear model linking initial and resultant quality has been proposed previously by Halcrow (2012), but in addition to the linear fit it was decided to stochastically model the high variability seen in the data. This involved modelling the residuals of the linear fit. The residuals, as expected, were more variable the higher the initial track geometry SD, creating the cone shape seen in Figures 4.9a and 4.10a.

To remove the effect the initial track quality had on the variability of the residuals, it was decided to divide the residuals by the initial quality to obtain relative residuals, Figures 4.9b and 4.10b. In all cases the relative residuals were slightly skewed, hence a normal distribution could not represent the data. To allow more types of distribution fits, all residuals needed to be positive. This was accomplished by adding the largest negative value in the dataset of relative residuals plus 10% to the other residuals. It was then possible to demonstrate the relative residuals using the Weibull distribution, with the PP plots of these fits shown in Figures 4.9c and 4.10c. Weibull was used as its highly flexible allowing both positive and negative skew. Other distributions were tested, but Weibull was found to best represent the data.

For each distinct group, the resulting model involved using the initial quality and a linear fit to obtain an initial resultant geometry SD, as seen in Figure 4.1a and represented by the equation:

$$TG_{After} = a * TG_{Before} + b + \epsilon \tag{4.2}$$

Where TG is the track geometry recorded as a SD of a length of track either before or after maintenance, a and b are the gradient and constant of the linear fit. ϵ represents the residuals.

Variability to account for the linear residuals was included by adding the distributed relative residuals multiplied by the initial track SD to convert it back to a geometry SD.

The transformed relative residual data for the Weibull distributions were found by:

$$Residuals = (a * TG_{Before} + b) - TG_{AfterActual} \tag{4.3}$$

$$RelativeResiduals = Residuals/TG_{Before} \tag{4.4}$$

$$TransformedRelativeResiduals = RelativeResiduals + c \tag{4.5}$$

Where c is the largest negative value in the dataset of relative residuals plus 10%. As the Weibull distribution is being used to represent the transformed relative residuals, combining the equation gives:

$$Weib(\eta, \beta) = (Residuals/TG_{Before}) + c \tag{4.6}$$

$$Residuals = (Weib(\eta, \beta) - c) * TG_{Before} \tag{4.7}$$

Where η and β are the scale and shape parameters for the fitted Weibull distribution of transformed relative residuals. Combining the linear fit and variability gives:

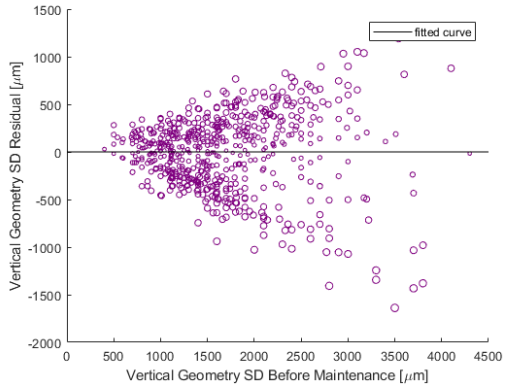
$$TG_{After} = a * TG_{Before} + b + ((Weib(\eta, \beta) - c) * TG_{Before}) \tag{4.8}$$

This equation can be more cleanly written as:

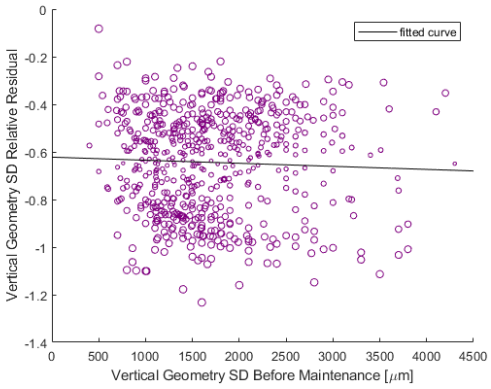
$$TG_{After} = (a + Weib(\eta, \beta) - c) * TG_{Before} + b \tag{4.9}$$

with the values for the variables found in Table 4.12. In this table it can be seen that all of the Weibull distribution fits had β values between 2.9 and 3.9, showing that the results had minimal skewness (very slightly negative skewed).

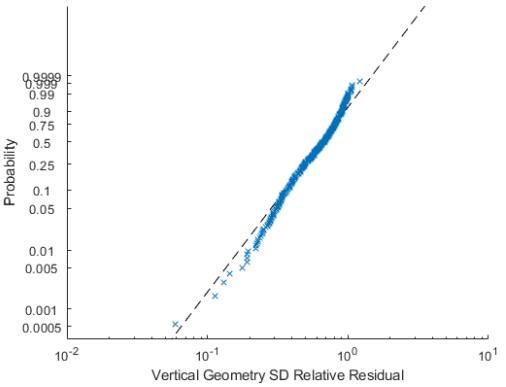
The model could then be simulated using a Monte-Carlo method. To ascertain the shape of the model results and compare against the actual data one simulation of Monte-Carlo was run using the actual data initial qualities (vertical geometry SD) as the input. This resulted in one result per original initial quality data point (sampled dataset same size as actual historical dataset being comparing to). From this it can be seen that the models accurately represent the actual effectiveness of maintenance seen in the data, Figures 4.9d and 4.10d.



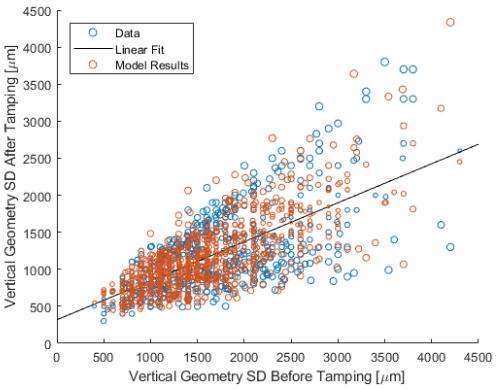
(a) Residuals of linear fit



(b) Relative Residuals of linear fit

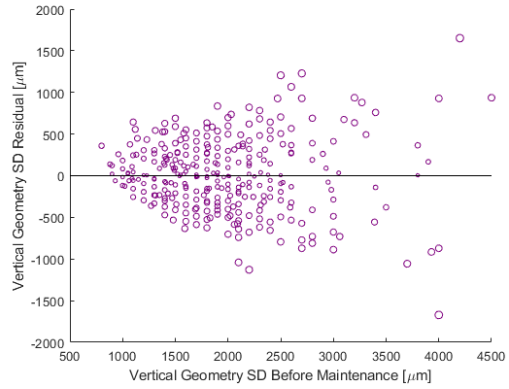


(c) PP plot of relative residuals Weibull distribution fit

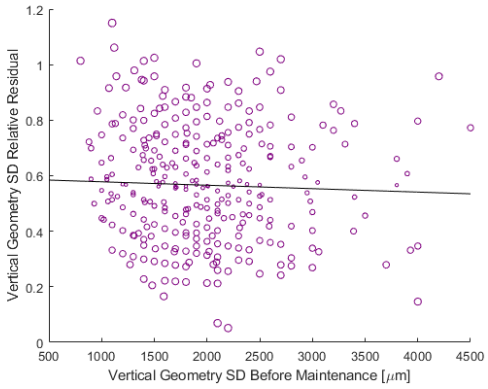


(d) Historical data vs model

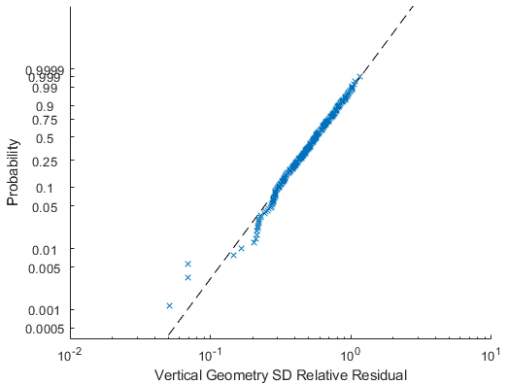
Figure 4.9: Estimating the resultant quality of the first tamp after renewal is performed on track with a speed greater than 115 MPH



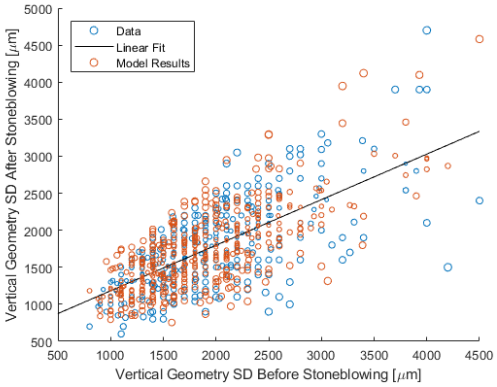
(a) Residuals of linear fit



(b) Relative Residuals of linear fit



(c) PP plot of relative residuals Weibull distribution fit



(d) Historical data vs model

Figure 4.10: Estimating the resultant quality when stoneblowing is performed on track with two or more previous stoneblowing operations

Table 4.12: Maintenance Effectiveness Model Parameters and Fit Statistics

		Stoneblowing			Tamping				
	Maintenance History	0S-1S	>2S	R	1T-3T	>3T	R	1T-3T	>3T
	Track Speed [MPH]			<115	<115	<115	$\geq$ 115	$\geq$ 115	$\geq$ 115
Linear Fit	Gradient (a)	0.641	0.615	0.574	0.694	0.781	0.424	0.527	0.526
	Constant (b)	468.7	567.2	303.4	175.6	83.4	476.8	318.2	348.3
	N	2781	440	555	1372	1416	217	896	464
	r^2	0.63	0.48	0.60	0.72	0.77	0.34	0.46	0.47
	RMSE	469.6	404.0	431.3	441.1	394.8	394.6	362.6	350.5
Relative Residuals Weibull Fit	η	0.654	0.589	0.763	0.661	0.616	0.736	0.720	0.713
	β	3.342	2.890	3.791	3.425	3.477	3.148	3.730	3.912
	Constant (c)	-0.632	-0.565	-0.683	-0.590	-0.557	-0.642	-0.646	-0.641
	chi^2	0.076	0.512	0.527	0.000	0.000	0.065	0.000	0.001
	AD test	0.130	0.965	0.648	0.011	0.000	0.341	0.004	0.055

4.5 Maintenance Output Rate

4.5.1 Introduction and Available Data

Due to railway infrastructure being used for a large part of every day, minimal access time is available for performing maintenance. Due to this, access window lengths play an important role in the ability to efficiently maintain the track. Maintenance data obtained from Network Rail (NR) was recorded as separate jobs (tamping and stoneblowing), with the time to complete the job and the length of track maintained. From this data it was possible to obtain an output per hour for each job.

4.5.2 Analysis

It was decided to subset the data to reduce the impact of outliers by removing maintenance jobs that took less than an hour or more than 24 hours to complete. A summary of the maintenance output rates for tamping and stoneblowing are shown in Table 4.13. The results show that the output rate of tamping is slightly higher than stoneblowing with tamping's median and mean rates being 1% and 5% higher. Tamping jobs tend to be shorter with a mean of 5.3 hours compared to the 6.2 hours for stoneblowing. Despite stoneblowing jobs tending to be longer, and hence possibly obtaining greater output economies of scale, the outputs were slightly lower than tamping. This could mean, if they had similar job lengths the difference between tamping and stoneblowing would be larger.

Table 4.13: Maintenance Output Rates

	Tamping			Stoneblowing		
	Poskeys/Hr	Yds/Hr	Hrs/Job	Poskeys/Hr	Yds/Hr	Hrs/Job
Minimum	0.228	50.1	1.0	0.227	50.0	1.0
Q1	0.576	126.8	2.5	0.578	127.2	3.0
Median	1.008	221.8	4.0	1.000	220.0	4.8
Mean	1.230	270.6	5.3	1.176	258.4	6.2
Q3	1.743	383.4	6.9	1.571	345.5	8.0
Maximum	4.545	1000.0	24.0	4.545	1000.0	24.0

4.5.3 Model

To characterise the output rates distributions were used. Before fitting distributions it was decided to reduce the output rates so the smallest values were closer to zero. This was done as many distributions including Weibull start at zero. A value of 80% of the

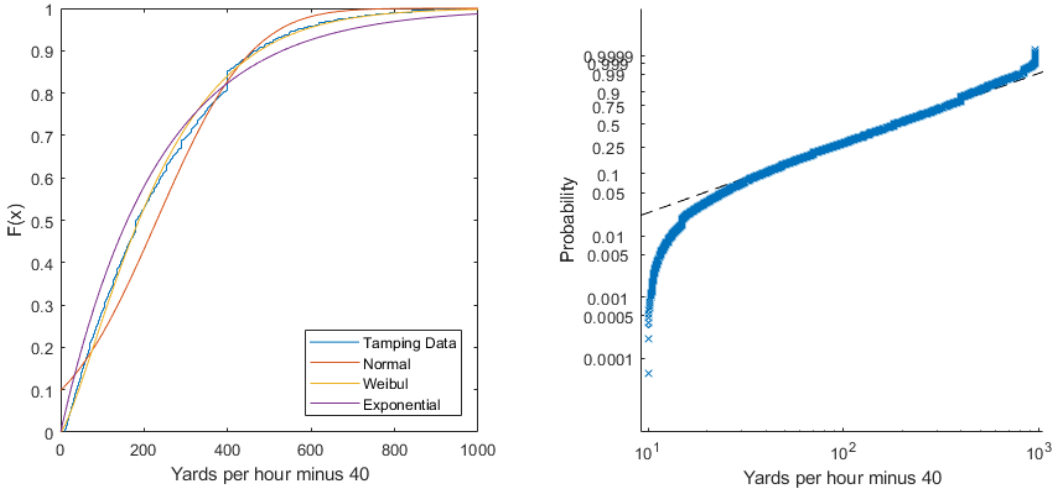
minimum values in Table 4.13 was used. The minimum value was 50 yds/hr, so 80% of this is 40. 50 was not used as it was desirable to have all non-zero positive values for distribution fitting. The Weibull, normal and exponential distributions were tested. Looking at the ECDF's for tamping and stoneblowing in Figures 4.11a and 4.12a it is clear that the Weibull distribution is the best fit for both types of maintenance. The quality of the fit can be seen in the PP-Plots in Figures 4.11b and 4.12b, where apart from some small problems with the lower tails the fits are strong. The fitted Weibull parameters are recorded in Table 4.14, in which it can be seen that the results for tamping and stoneblowing are very similar. This was expected due to the closeness of results in Table 4.13. The fitted distributions can be used to obtain expected maintenance job output rates using the Weibull inverse Cumulative Distribution Function (CDF), as demonstrated by Equations 4.10 for tamping and 4.11 for stoneblowing, where p is a uniform random number between zero and one. To obtain an estimate of amount of yards of maintenance is complete in one work shift, for example 6 hours, the sampled values from Equations 4.10 and 4.11 (via Monte-Carlo), are multiplied by the number of hours of work. This is then divided by 220 to obtain an estimate of the number of Poskeys which could be maintained (Poskeys are 220yds each).

$$\text{Tamping Output Rate [yd/hr]} = 249.278 * (-\ln(1-p))^{1/1.28754} + 40 \tag{4.10}$$

$$\text{Stoneblowing Output Rate [yd/hr]} = 237.267 * (-\ln(1-p))^{1/1.30895} + 40 \tag{4.11}$$

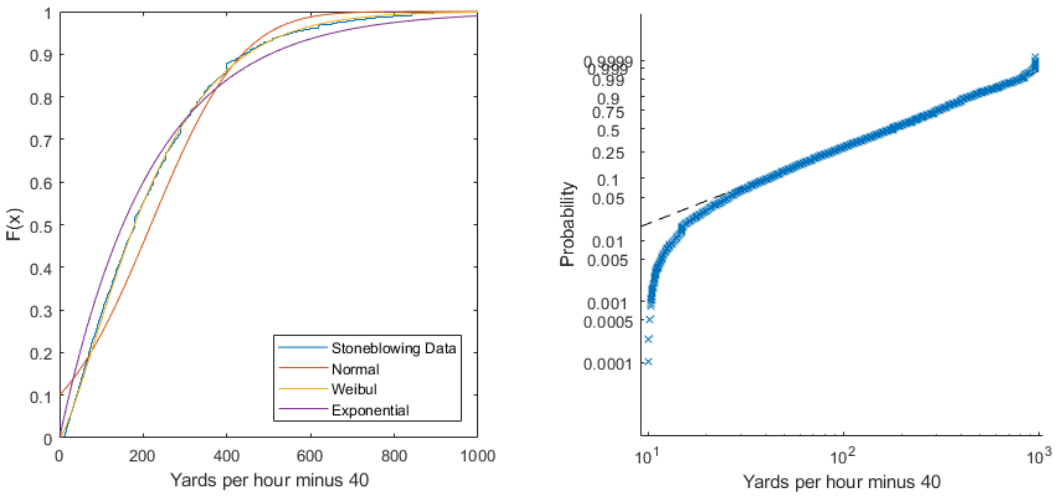
Table 4.14: Maintenance Output Rate [yds/hr] (minus 40) Weibull Distribution Parameters

	Weibull Parameters	
	η	β
Tamping	249.278	1.28754
Stoneblowing	237.267	1.30895



(a) ECDF Comparing Fits (b) PP-Plot of Weibull Fit

Figure 4.11: Distribution fits of tamping output [yards per hour]



(a) ECDF Comparing Fits (b) PP-Plot of Weibull Fit

Figure 4.12: Distribution fits of stoneblowing output [yards per hour]

4.6 Summary

This chapter has explored the factors that influence the effectiveness of maintenance as well as, the output rate of tamping and stoneblowing activities. Models for each of these were also introduced.

The initial track quality was shown to have a large impact on the improvement seen from performing maintenance, more so for stoneblowing compared to tamping. The geometry records before and after maintenance showed that an increase in the track geometry SD can happen (quality reduced by maintenance). This was much more numerous for

Stoneblowing activities, especially if performed on track with already good geometry. The worse the initial quality (when maintenance performed), the greater the average improvement was, but the spread of possible improvements also increased; by using a relative/percentage improvement this was normalised.

Track speed was shown to have little effect on the improvement following tamping, other than track speeds of 115 MPH and over, which had a significantly higher average relative improvement. There is not an obvious link between track speed and maintenance. Higher speed sections tend to be busier and more critical lines, so higher quality maintenance might be being performed. The improvement from stoneblowing had no relationship with track speed.

Maintenance history was showed a small impact on the effectiveness of maintenance. The more tamps previously performed the less effective the maintenance actions became. Three distinct groups being identified; directly after a renewal, after one to three tamps and after more than three tamps. The greatest relative improvement was found in the group one to three tamps followed by more than three tamps. Stoneblowing was shown not to be affected by the amount of previous tamping operations. The number of previous stoneblowing activities was shown to be important. The improvement seen from the first two stoneblowing operations, on any combination of previous tamps, was consistent. The average improvement of stoneblowing in the group 0-1 previous stoneblows was significantly higher than found stoneblowing performed on track which has previously undergone 2 or more previous stoneblowing activities. The significance of the results suggest that after a second stoneblowing activity the next maintenance activity should be a full renewal. As stoneblowing adds small stones into the ballast, these small stones will build up, and appear to impact the ability to successfully undertaken stoneblowing. The reduction in improvement seen is probably due to a build up of small stones beneath the sleepers from previous stoneblowing activities.

Based on the results of the analysis it was decided to split the maintenance effectiveness data into eight groups. Two for stoneblowing activities; performing stoneblowing on track which has previously been stoneblown one or fewer times and more than one time. Six groups for tamping; first tamp after a renewal, after one to three tamps and after more than 3 tamps. These three groups are duplicated for track speeds under 115MPH and those equal to and over.

In Section 4.4, the data and results from the maintenance effectiveness analysis were used to develop a stochastic model. The eight significantly different factors are used to split up the available data. The models consisted of a linear relationship between the initial and resultant vertical track geometry SD with the addition of a stochastic variable to take into account the large variations found in the data. The stochastic variable consisted of transforming the linear equation residuals by dividing by the initial quality, and then represented these by a Weibull distribution. The residuals were transformed to remove the high amounts of heteroscedasticity, making the Weibull distribution represent the

spread at all initial qualities. The final model is shown to give results that are comparable to the actual maintenance effectiveness data.

As well as maintenance effectiveness, the output rate has also been analysed. For a maintenance model it is important to understand how much mainteance can be performed at a time. If ten poskeys need to be maintained, it would be wrong to classify this has ten separate maintenance activities. Railway works activities cost per shift of work. So if the ten poskeys could be completed in one shift, this should only be recorded as one maintenance activity. Whereas if only five poskeys could be maintained in a shift, it is two maintenance activities to complete the work (which would have a higher cost for the same number of poskeys).

The analysis showed that there was a wide range of possible output rates and job lengths (in hours). Tamping and stoneblowing had similar output rates, with tamping being very slightly higher. Tamping activities tended to be shorter than stoneblowing, with a mean job lengths of 5.3 hours and 6.2 hours. The longer jobs tended to show higher output rates, which was expected due to economies of scale and setup being a small proportion of the job (by time). Weibull distributions was found to strongly represent the range of output rates for tamping and stoneblowing. Output rates can be sampled using Monte-Carlo and the inverse Weibull CDF.

In the previous chapters the rate of track geometry degradation has been analysed as well as the effectiveness of the primary maintenance actions to repair the geometry, tamping and stoneblowing. A railway track is made up of many aspects which can degrade. Tamping and stoneblowing are common maintenance activities but so are rail replacements and maintenance. To build a more complete model it was decided to include rail degradation (as faults and breaks), which is discussed in the next chapter.

Chapter 5

Rail Faults

5.1 Introduction

Rails experience large forces (nominal, dynamic and impact wheel loadings) from traversing trains leading to a wide range of defects, which can be common occurrences. The majority of these faults are caused by fatigue from the contact between the trains wheels and the rail. Section 2.3.2 describes the main types of faults, their failure mechanisms and the inspections and maintenance required. In the UK there have been two main types of rail design; the superseded Bullhead (BH) and the currently used Flatbottom (FB), which come in different sizes, described by their weight. Additionally, there are two types of rail joining methods: the obsolete jointed rail where plates bolted either side hold rail sections together, and Continuous Welded Rail (CWR) where the rail ends are welded, these are described in further detail in Section 2.3.2.

Rail faults can cause many problems on a railway track as it is important that there is good contact between the wheels and rail as well as a smooth running surface. Faults such as corrugations alter the rail surface smoothness, impacting ride quality, whereas others such as head wear lead to a reduced wheel/rail interface area which decreases the grip for breaking and accelerating. Rail faults can turn into breaks which lead to increased forces on surrounding sleepers and speed restrictions being required to maintain safety. These are known as Early Speed Restrictions (ESRs) and can be costly to implement due to subsequent delays in the train timetables. Breaks also impact the track electric circuits which are used to know the location of trains on the track.

Rail faults are usually identified through visual checks or by the use of ultrasonic equipment used to identify internal defects. To correct the identified faults, grinding or welding is used to return the rail to the correct profile or the rail is replaced, depending on the fault type and severity.

Due to the impact that rail faults can have on safety and cost, it is important to understand the likelihood of occurrences to be able to predict future costs of maintaining the

track and make informed decisions on asset choices, especially in terms of upgrades.

5.2 Data and Processing

Inspection data has been obtained from Network Rail (NR), in which each fault found has been recorded for the whole UK railway infrastructure from 2007 to 2015. The information included the fault, rail and joint types, whether a rail break occurred or an ESR was required. This data is summarised in Table 5.1, in which it can be seen that squats are the most common fault type. These are followed by other, weld, then profile issues, such as lipping and side wear, and then Rolling Contact Fatique (RCF). The fault group 'other' includes problems such as vertical and horizontal head splits. Looking at the summary results in Table 5.1, within the rail fault dataset 0.538% of faults also had a rail break reported. The probability of a break occurring is highly dependent on the type of fault, with bolt holes leading to breaks nearly 6% of the time. ESRs are more common with 2.807% of faults leading to a speed restriction being required. Bolt hole faults are shown to require an ESR 43% of the time, which, when considered with prevalence of breaks, shows the importance of the move from jointed rail to CWR (which has much fewer joints, with these being Insulated Rail Joints (IRJs) that are required for track circuits). Head wear faults seem to have a greater impact on the running of a railway than other wear-related faults (lipping and side wear), where the other ones almost never require ESRs and have no recorded breaks. Wheelburn faults also do not tend to turn into breaks or require ESRs.

Table 5.1: Available Rail Fault Data

		Breaks		Early Speed Restrictions	
Fault Type	Total Faults	Amount	Percentage	Amount	Percentage
Squat	63722	10	0.016	649	1.018
Tache Ovale	7420	57	0.768	633	8.531
Bolt Hole	1705	102	5.982	735	43.109
Weld	20295	278	1.370	393	1.936
Other	27989	502	1.794	1782	6.367
RCF	10693	6	0.056	424	3.965
Wheelburn	9795	1	0.010	49	0.500
Lipping	11956	0	0.000	2	0.017
Side wear	18735	0	0.000	11	0.059
Head wear	4248	4	0.094	153	3.602
Corrugation	24	0	0.000	0	0.000
Unknown	2117	1	0.047	185	8.739
Overall	178699	961	0.538	5016	2.807

5.2.1 Inspections and Maintenance Types

A number of different inspection and maintenance types are required due to the variation of faults. Visual inspections are primarily used for external faults such as wear, whereas ultrasonic inspections find inner defects such as tache ovale and squats. Figure 5.1 shows the percentages of each fault type found by each inspection method. From this it can be seen that many fault types are found using one primary type of inspection apart from weld, other and RCF defects that are found using more than one type of inspection. The results show that ultrasonic inspections find more faults than visual inspections, which is due to the fact that a large proportion of fault types start internally.

The breakdown of the type of maintenance used for each fault type is shown in Figure 5.2. This shows a similar trend to inspections, that certain types of maintenance are better suited to different faults. Wear-related faults tend to be ground or welded, but corrugations and bolt holes need rail replacements.

Table 5.2: Inspection and maintenance methods for different rail faults

	Inspections			Maintenance			
	Ultrasonic	Other	Visual	ReRail	Weld	Grind	Other
Squat	84.4%	1.4%	14.2%	32.8%	62.6%	1.9%	2.7%
Tache Ovale	97.9%	1.4%	0.7%	72.2%	24.1%	1.2%	2.4%
Bolt Hole	92.3%	6.6%	1.1%	90.0%	1.9%	0.1%	8.0%
Weld	47.4%	7.3%	45.3%	51.9%	38.5%	4.5%	5.2%
Other	57.7%	6.0%	36.3%	64.1%	27.7%	4.3%	3.9%
RCF	63.9%	4.7%	31.4%	50.8%	27.8%	18.8%	2.6%
Wheelburn	94.1%	0.8%	5.0%	26.7%	60.7%	6.3%	6.3%
Lipping	5.5%	5.5%	89.0%	2.9%	10.5%	85.1%	1.4%
Side wear	0.5%	20.3%	79.2%	4.4%	29.4%	63.4%	2.8%
Head wear	11.2%	13.0%	75.7%	21.3%	53.9%	20.9%	3.9%
Corrugation	69.6%	13.0%	17.4%	70.6%	5.9%	11.8%	11.8%
Unknown	92.6%	0.8%	6.6%	46.4%	16.6%	30.8%	6.2%

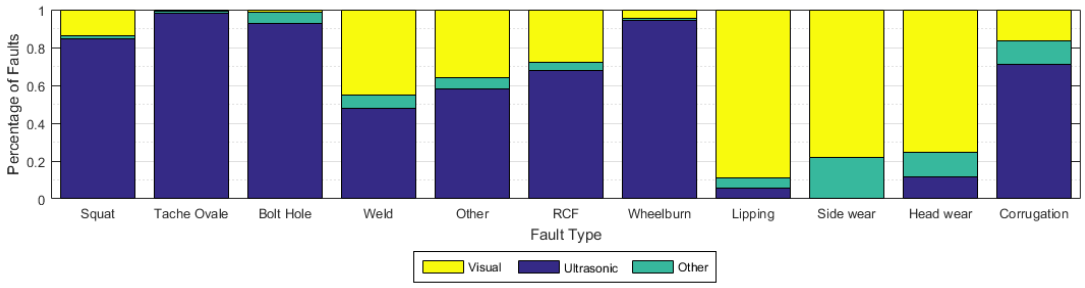


Figure 5.1: Inspection methods for different rail faults

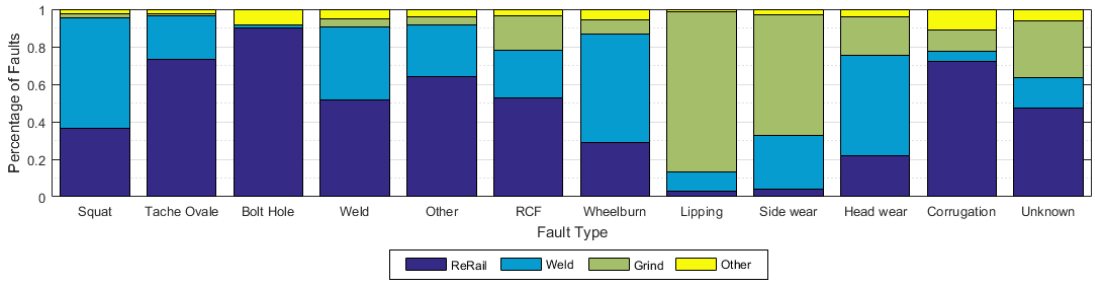


Figure 5.2: Maintenance methods for different rail faults

5.3 Analysis of Rail Fault Occurrences

5.3.1 Methodology

To understand the occurrences of rail faults it is important to interpret the factors which influence them. A fault rate related to the length of track and the amount of traffic was found and used to analyse the impact of different factors which were thought to be related to the rate at which rail faults occur. It was decided to use a track length of 220yd (same as one poskey) and to take account of usage in Equivalent Million Gross Tonnage (EMGT). This is because usage rather than time is the main driver of rail faults. First the rail fault occurrence rate per 220yd poskey per year was found using:

$$Rate_{Faults/Poskey/Year} = NumFaults/TrackLength * 220/8 \tag{5.1}$$

Where $TrackLength$ is in yards, and the results was divided by eight, as the dataset included eight years of faults. This was then converted to fault rate per Poskey per EMGT by:

$$Rate_{Faults/Poskey/EMGT} = Rate_{Faults/Poskey/Year}/Traffic_{EMGT/Year} \tag{5.2}$$

Where $Traffic_{EMGT}$ is the average EMGT which traverses a poskey a year.

The rate of faults per poskey per EMGT was found for the many factors explored. For example, calculating the rate of faults on jointed rail involves taking the total amount of faults found on jointed rail divided by the length of track with jointed rail (in yards) multiplied by 220 to obtain a rate per poskey, divided by 8 to make it yearly.

This is then divided by the average traffic in EMGT recorded traversing jointed rail track per year over the time period of the inspection data (2007-15). Performing this calculation normalises the occurrence rates taking into account the differing length of track within which different asset types occur and variability of traffic across the railway network.

5.3.2 Results

The fault rates per poskey per EMGT were calculated initially with no factors to give an average across the whole network. This is shown in Figure 5.3, where the numbers above the bars are the number of faults. The overall fault rate of a poskey is 0.016/EMGT, meaning a fault occurs every 62 EMGT. Squats are the most numerous and hence have the highest rate with it expected that a squat would occur on a poskey every 167 EMGT of traffic.

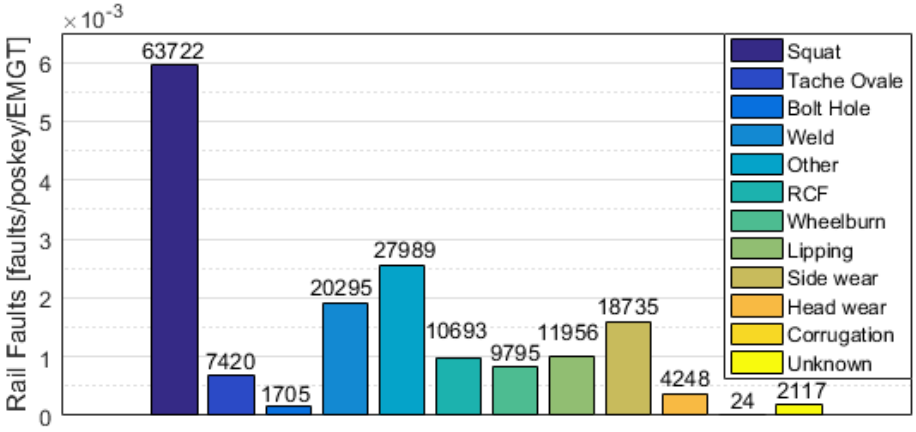


Figure 5.3: Rail fault rates

Rail Age

To assess the impact the rail age has on the fault rate it was decided to split the data into track sections with rails installed in differing decades. To remove the impact of the changing type of rails and joint types across the decades, only 113 lb FB CWR was assessed, as this is the most common rail and joint type combination on the UK railway and has been used since the 1960's. By keeping the rail and joint type consistent, any differences in the fault rates are likely due to the rail age. It was expected that the rate of faults would be greater on older rail, as most rail faults are caused by wear and fatigue. The rates of the different fault types for rail installed in each decade since the 1960's is recorded in Table 5.3 and shown visually in Figure 5.4. The results show that the older the rail the higher the rate of faults, with rail installed in the 1960's having an 145% increase in faults compared to rail installed in the 2000's. It also appears that the link between the track age and fault rate is mostly linear as seen in Figure 5.5, where the total fault rates for each decade have been plotted against the middle of the decades. The linear fit had an r-squared value of 0.9301 and a Root Mean Squared Error (RMSE) of 0.0017.

Looking at the different types of rail faults in Table 5.3, it appears that weld, squat and tache ovale defects are the ones most affected by the rail age, whereas wear-related faults (lipping, side and head wear) are less so, but still affected. In general the different fault types follow a similar pattern to the total faults with newer rail having lower rates of faults. RCF, bolt hole and wheelburn faults do not show a correlation between age and fault rate. This makes sense for wheelburn faults, as these are not caused by fatigue but by the high forces from wheels slipping (acceleration which is too high for the grip between the wheels and rail). It is unknown why RCF has no relationship with track age as it is caused by fatigue. It is possible that rail on bends, where RCF tends to occur, is renewed more often than straight line rail, and hence the proportion of rail on corners is lower in the datasets of older rail. Weld faults see a large step (50% increase) in the rate of occurrence between rail installed in the 70's and 80's. This gives evidence to suggest

that the expected life of a weld is 30-40 years.

Table 5.3: The effect of rail age (denoted by rail installation decade) on the rates of different rail faults on 113 lb Flatbottom Continuous Welded Rail [Faults/Poskey/Equivalent Million Gross Tonnage]

	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
1960-69 (60's)	8.81e-3	1.07e-3	1.01e-4	3.02e-3	3.91e-3	9.28e-4	6.64e-4
1970-79 (70's)	8.83e-3	9.89e-4	1.14e-4	2.79e-3	3.16e-3	1.03e-3	9.06e-4
1980-89 (80's)	6.27e-3	6.97e-4	1.10e-4	1.88e-3	1.98e-3	1.00e-3	5.72e-4
1990-99 (90's)	6.10e-3	6.18e-4	9.99e-5	1.46e-3	2.00e-3	1.31e-3	7.18e-4
2000-09 (00's)	2.77e-3	3.77e-4	4.52e-5	9.51e-4	1.32e-3	8.99e-4	4.88e-4
70's % Var 60's	0.2	-7.4	12.9	-7.5	-19.3	10.6	36.4
80's % Var 60's	-28.8	-34.8	9.1	-37.7	-49.4	8.2	-13.9
90's % Var 60's	-30.8	-42.2	-1.2	-51.5	-48.9	41.2	8.0
00's % Var 60's	-68.5	-64.8	-55.3	-68.5	-66.1	-3.2	-26.6

	Lipping	Side wear	Head wear	Corrugation	Unknown	Total
1960-69 (60's)	1.09e-3	2.05e-3	4.66e-4	3.61e-6	2.42e-4	2.23e-2
1970-79 (70's)	1.18e-3	2.07e-3	4.58e-4	0.00e+0	2.33e-4	2.18e-2
1980-89 (80's)	8.92e-4	1.53e-3	3.38e-4	1.77e-6	1.33e-4	1.54e-2
1990-99 (90's)	7.08e-4	1.19e-3	4.57e-4	4.16e-6	1.93e-4	1.49e-2
2000-09 (00's)	7.65e-4	1.08e-3	2.75e-4	1.85e-6	1.26e-4	9.10e-3
70's % Var 60's	8.4	1.0	-1.7	-100.0	-3.8	-2.7
80's % Var 60's	-18.0	-25.5	-27.4	-51.0	-45.1	-31.1
90's % Var 60's	-34.8	-41.8	-2.0	15.2	-20.0	-33.5
00's % Var 60's	-29.6	-47.5	-40.9	-48.9	-47.7	-59.3

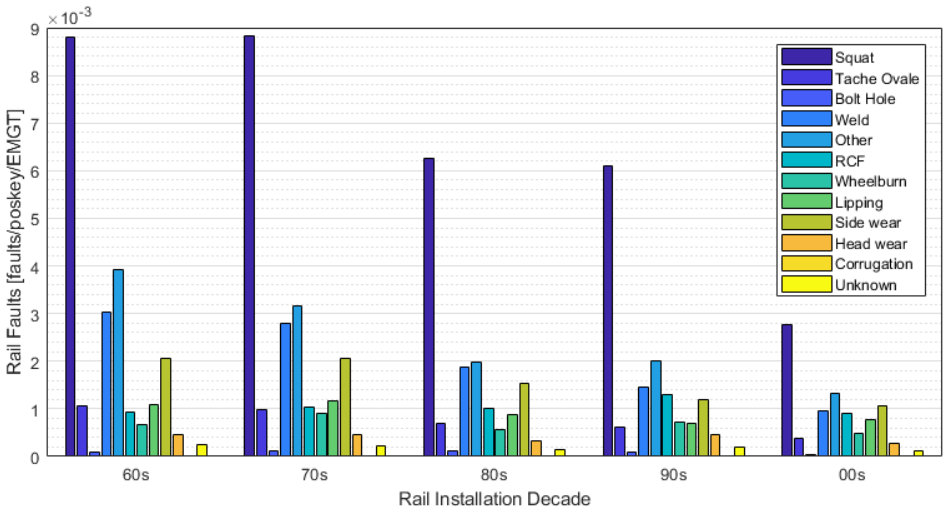


Figure 5.4: The effect of rail age (denoted by rail installation decade) on the rates of different rail faults on 113 lb Flatbottom Continuous Welded Rail rail

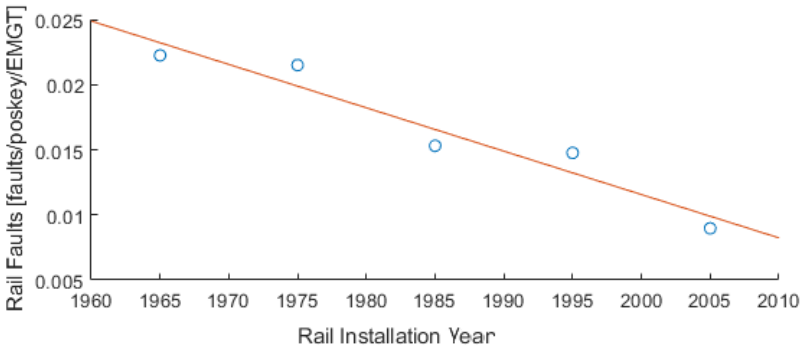


Figure 5.5: The affect of rail age (denoted by rail installation decade) on the rate of rail faults on 113 lb Flatbottom Continuous Welded Rail rail, linear fit

Curvature

Due to the increased sideways (centrifugal) forces, there is a shift of forces from the inner rail to the outer rail when a train traverses curved track. This results in the forces that occur on curved rails being higher than on straight track. Due to this it was expected that faults would be more common on curved sections of track. The track sections were split into four groups of curvature which are defined in Table 5.4. The table also shows the fault rates for the different curvatures for each fault type and the percentage variance from straight track. Figure 5.6 shows the data for all faults. These show that, in general, faults occur less often on track with curvatures between 0-0.001 m^{-1}, but there is a 40% increase for tightly curved track between 0.001-0.002 m^{-1}. These results differ from Sawley (2001) who predicted a reducing lifetime of rails as the curvature increased.

The impact of curvature on different rail faults varies as expected and is due to the individual failure mechanisms. Corrugations and RCF are related to fatigue caused by high forces, hence curvature has a greater effect on the rate of these faults occurring compared to wear-related faults. Tache ovales are seen to be strongly related to the curvature of the track which was unexpected as these tend to originate from manufacturing defects. The evidence points to the increased forces experienced on rails increasing the rate at which they grow internally. The reason that the rate of faults of slightly curved track is lower than straight is unknown but may be due to maintenance procedures (curved track rails replaced more often than straight) or how the track is used, with the speed restrictions common on curved sections reducing the forces. It is also thought that the use of a cant on curved sections, as is the case on the UK railway network, reduces the impact of the centrifugal forces.

Table 5.4: The effect of curvature on the rates of different rail faults [Faults/Poskey/Equivalent Million Gross Tonnage]

Curvature [m^{-1}]	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
0	6.447e-3	6.670e-4	1.900e-4	1.887e-3	2.959e-3	1.004e-3	1.088e-3
0-0.0005	4.870e-3	5.280e-4	7.091e-5	1.498e-3	1.637e-3	6.871e-4	3.882e-4
0.0005-0.001	5.465e-3	8.137e-4	1.092e-4	2.241e-3	2.219e-3	1.023e-3	5.618e-4
0.001-0.002	7.741e-3	1.587e-3	2.174e-4	3.368e-3	3.803e-3	2.199e-3	1.538e-3
0-0.0005 % Var 0	-24.5	-20.8	-62.7	-20.6	-44.7	-31.5	-64.3
0.0005-0.001 % Var 0	-15.2	22.0	-42.5	18.7	-25.0	2.0	-48.4
0.001-0.002 % Var 0	20.1	137.9	14.4	78.4	28.5	119.1	41.3

Curvature [m^{-1}]	Lipping	Side wear	Head wear	Corrugation	Unknown	Total
0	1.316e-3	2.147e-3	4.921e-4	1.585e-6	1.923e-4	1.839e-2
0-0.0005	4.814e-4	6.828e-4	1.671e-4	2.663e-6	8.623e-5	1.110e-2
0.0005-0.001	6.608e-4	9.366e-4	2.656e-4	1.365e-6	1.645e-4	1.446e-2
0.001-0.002	1.564e-3	2.435e-3	6.287e-4	9.349e-6	6.427e-4	2.573e-2
0-0.0005 % Var 0	-63.4	-68.2	-66.0	68.1	-55.2	-39.6
0.0005-0.001 % Var 0	-49.8	-56.4	-46.0	-13.8	-14.4	-21.4
0.001-0.002 % Var 0	18.8	13.4	27.7	489.9	234.3	39.9

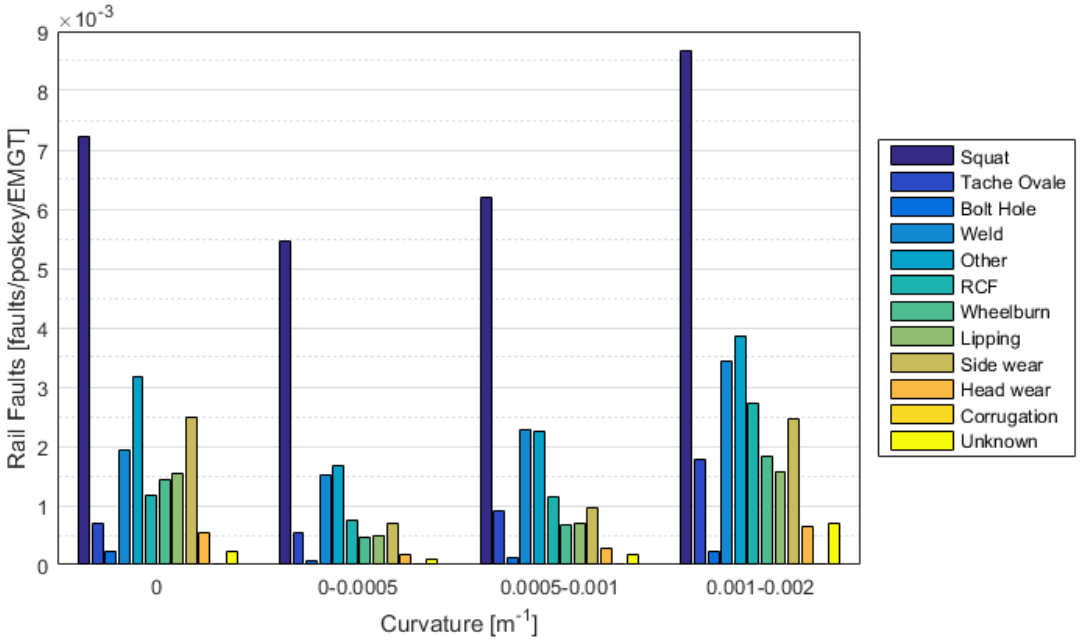


Figure 5.6: Rail fault rates split by track curvature

Rail Type

The UK railway infrastructure contains many types of rail. The main two designs are the superseded BH and currently used FB, which has a much larger foot profile. Within

the two design groups there are different sizes, described by lb/yard, varying from 95 lb to 121 lb. The only rail type not measured this way is the most recent, UIC 60, where the 60 stands for 60 kg/m which is equivalent to 121 lb/yd. The most common rail types installed are 113 lb FB, 110 lb FB, 95 lb BH, UIC 60 FB and 109 lb FB, with 113 lb FB being the most common. To analyse the effect the rail type has on the occurrence rate of faults it was required to remove other factors which influence the rate first. Due to rail types being installed over different time periods, to remove the impact of rail age it was decided to analyse rails installed over a similar time period. Additionally, the effect of the joint type was removed by only using CWR. It was decided to use CWR as it was the most common method of joining rails and is still in use today.

To assess the variation of older rail types it was decided to use rail installed between 1955-70. This period was chosen as there were many designs installed, allowing for a comparison between more rail types. The occurrence rates are recorded in Table 5.5, which is shown visually in Figure 5.7. The results show that larger rails have lower rates of faults, with 98 lb FB seeing a 10.6% reduction compared to 95 lb BH. 109 lb FB has a 28.9% reduction, 110 lb has 30.3% and 113 lb has 38.4%. Looking at the individual types of faults, the largest reductions occur in bolt hole, other, wheelburn and weld. Wear-related faults (lipping, side wear and head wear) also see large reductions. 98 lb FB rail appears to have an issue with tache ovale, with a rate that is over three times higher than the other rail types. Tache ovale faults primarily start as manufacturing defects, so it appears that the quality of 109 lb FB produced between 1955-70 was lower than that of other rail types. The results were as expected as larger rails reduce the stress within the rail caused by the forces of traversing trains. As most rail faults are caused by fatigue, lower stresses lead to less fatigue and hence reduced number of faults.

To compare more recent rail types it was decided to use rails installed between 1995-2010, with the results shown in Table 5.6 and Figure 5.8. 110 lb and 113 lb FB rails occur in both the older the newer rail datasets. For 1955-70 113 lb rail fault rate was 11.6% less than 110 lb in Table 5.5. For the years 1995-2010 it was 9.6% less. The similar results, even though the rails were installed in different decades, give confidence to saying that the amount of faults that occur on 113 lb FB rail is around 10% lower than 110 lb rail. As the sizes of these rails are similar it is envisaged that the costs would also lead to a strong case to use 113 lb over 110 lb. UIC 60 FB rail sees a large decrease in faults with a 48.6% reduction from 110 lb FB rail. This is primarily from a large reduction in squats, but large reductions are seen in all rail fault types but bolt holes, side wear and head wear. It is unknown why there has been an increase in side and head wear of over 100% especially as lipping (which is also wear-related) has seen a 43.9% reduction.

Table 5.5: The effect of rail type on the rates of different rail faults on Continuous Welded Rail installed 1955-1970 [Faults/Poskey/Equivalent Million Gross Tonnage]

	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
95 lb BH	9.98e-3	9.29e-4	5.80e-4	4.99e-3	7.20e-3	3.48e-4	1.28e-3
98 lb FB	6.27e-3	4.13e-3	0.00e+0	4.91e-3	8.53e-3	2.58e-4	1.23e-3
109 lb FB	9.35e-3	1.44e-3	4.52e-5	3.47e-3	3.72e-3	5.43e-4	9.77e-4
110 lb FB	9.97e-3	1.04e-3	6.37e-5	3.91e-3	3.60e-3	6.50e-4	4.96e-4
113 lb FB	8.25e-3	1.07e-3	3.46e-5	2.90e-3	3.14e-3	7.53e-4	6.12e-4
98 lb % Var 95 lb	-37.2	345.3	-100.0	-1.6	18.5	-25.8	-3.9
109 lb % Var 95 lb	-6.3	54.9	-92.2	-30.4	-48.3	55.8	-23.5
110 lb % Var 95 lb	-0.1	11.9	-89.0	-21.6	-50.0	86.6	-61.1
113 lb % Var 95 lb	-17.3	15.3	-94.0	-41.9	-56.4	116.3	-52.1

	Lipping	Side wear	Head wear	Corrugation	Unknown	Total
95 lb BH	2.90e-3	6.96e-4	3.48e-4	1.16e-4	3.48e-4	2.97e-2
98 lb FB	5.81e-4	1.29e-4	1.94e-4	0.00e+0	3.23e-4	2.66e-2
109 lb FB	8.86e-4	4.07e-4	1.45e-4	0.00e+0	1.54e-4	2.11e-2
110 lb FB	3.89e-4	3.69e-4	8.80e-5	3.74e-6	1.37e-4	2.07e-2
113 lb FB	5.70e-4	6.32e-4	2.11e-4	0.00e+0	1.35e-4	1.83e-2
98 lb % Var 95 lb	-80.0	-81.4	-44.3	-100.0	-7.2	-10.6
109 lb % Var 95 lb	-69.5	-41.6	-58.4	-100.0	-55.8	-28.9
110 lb % Var 95 lb	-86.6	-47.0	-74.7	-96.8	-60.8	-30.3
113 lb % Var 95 lb	-80.4	-9.2	-39.5	-100.0	-61.3	-38.4

Table 5.6: The effect of rail type on the rates of different rail faults on Continuous Welded Rail installed 1995-2010 [Faults/Poskey/Equivalent Million Gross Tonnage]

	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
110 lb FB	3.27e-3	3.06e-4	1.27e-5	1.25e-3	1.81e-3	1.06e-3	4.71e-4
113 lb FB	3.05e-3	4.04e-4	4.01e-5	9.70e-4	1.14e-3	8.20e-4	4.65e-4
UIC60 FB	1.37e-3	1.18e-4	1.43e-5	3.80e-4	6.19e-4	5.49e-4	2.26e-4
113 lb % Var 110 lb	-6.8	32.3	214.7	-22.3	-37.1	-22.5	-1.4
UIC60 % Var 110 lb	-58.1	-61.3	12.3	-69.6	-65.8	-48.1	-52.1

	Lipping	Side wear	Head wear	Corrugation	Unknown	Total
110 lb FB	5.73e-4	4.08e-4	1.02e-4	0.00e0	1.78e-4	9.44e-3
113 lb FB	5.23e-4	7.39e-4	2.49e-4	1.39e-6	1.27e-4	8.53e-3
UIC 60 FB	3.22e-4	1.02e-3	2.07e-4	7.95e-7	2.62e-5	4.86e-3
113 lb % Var 110 lb	-8.8	81.3	144.3		-28.8	-9.6
UIC60 % Var 110 lb	-43.9	150.7	107.7		-85.3	-48.6

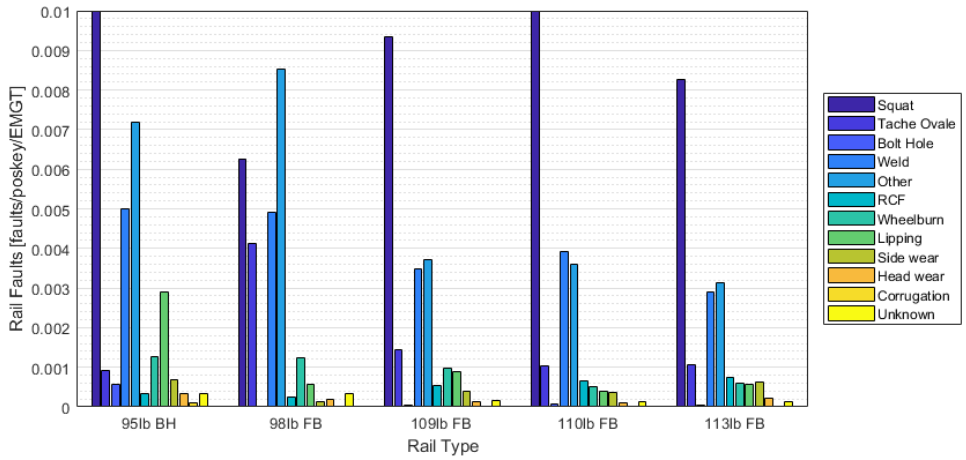


Figure 5.7: Rail fault rates split by rail type [Continuous Welded Rail installed 1955-1970]

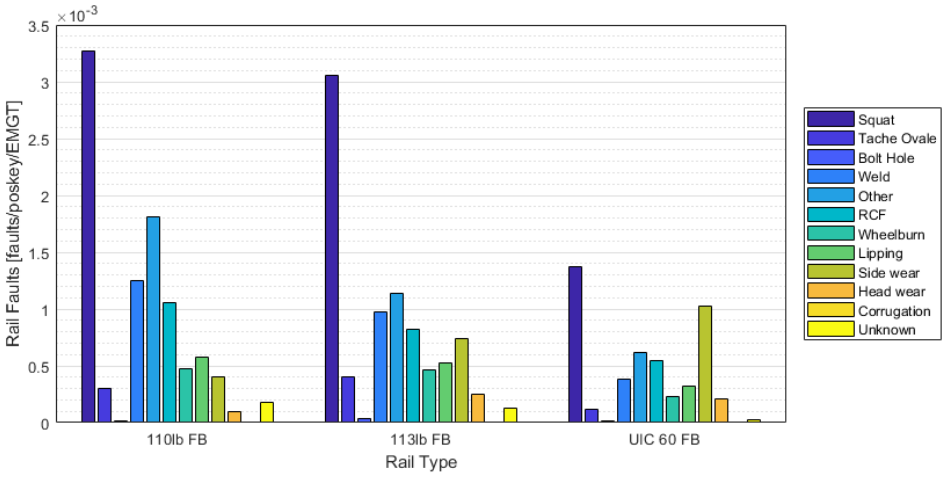


Figure 5.8: Rail fault rates split by rail type [Continuous Welded Rail installed 1995-2010]

Stations and Tunnels

It was thought that the existence of stations and tunnels could have an impact on rail faults due to the different loading found at stations and the protected environment within tunnels. As before, the rates of different faults on track containing stations and tunnels has been found and compared to the average of the rest of the network. These rates are shown in Table 5.7 and Figure 5.9. Table 5.7 also shows the percentage difference of the rates compared to the network average of track sections with neither tunnels and stations. This shows that rail faults are 44.8% more likely to occur at stations and 38.4% less within tunnels.

The reduction in tunnels is expected due to the protection from the environment such as temperature fluctuations and frost, with rails being affected by extreme temperatures due

to contraction and expansion. Only head wear, other and corrugations have increased rates in tunnels. The results from corrugations can be misleading due to a small sample size of only 24 cases being recorded to have occurred over the time period of the data. It is unknown why head wear is more common as other wear-related faults see large decreases in their occurrence rate (lipping and side wear). The slight increase in other faults could be due to difficulties performing inspections leading to faults being missed early on and them turning into other faults such as rail splits before they are noticed.

The increase in fault rates for all types of faults seen at stations was expected due to the dynamic forces produced from the trains braking and accelerating. Wheelburn faults see the largest increase as these are caused by the wheels slipping on the rail as the trains accelerate. Wear-related faults see the next greatest increase at stations, which would also be due to slipping between the wheels and rail causing friction and in turn wear of the rail.

Table 5.7: The effect of tunnels and stations on the rates of different rail faults [Faults/Poskey/Equivalent Million Gross Tonnage]

	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
Neither	5.9707e-3	6.970e-4	1.500e-4	1.909e-3	2.497e-3	9.933e-4	8.113e-4
Tunnels	2.229e-3	1.762e-4	5.287e-5	9.825e-3	2.837e-3	3.833e-4	5.155e-4
Stations	7.785e-3	8.987e-4	1.945e-4	2.183e-3	3.842e-3	1.273e-3	2.039e-3
Tunnels % Var	-62.7	-74.7	-64.7	-48.5	13.6	-61.4	-36.5
Stations % Var	30.4	28.9	29.7	14.4	53.9	28.1	151.3

	Lipping	Side wear	Head wear	Corrugation	Unknown	Total
Neither	1.010e-3	1.599e-3	3.403e-4	2.281e-6	1.834e-4	1.616e-2
Tunnels	2.071e-4	4.847e-4	1.740e-3	4.406e-6	3.481e-4	9.962e-3
Stations	1.859e-3	2.425e-3	6.100e-4	0.000e0	2.858e-4	2.340e-2
Tunnels % Var	-79.5	-69.7	411.4	93.1	89.8	-38.4
Stations % Var	84.1	51.6	79.2	-100.0	55.9	44.8

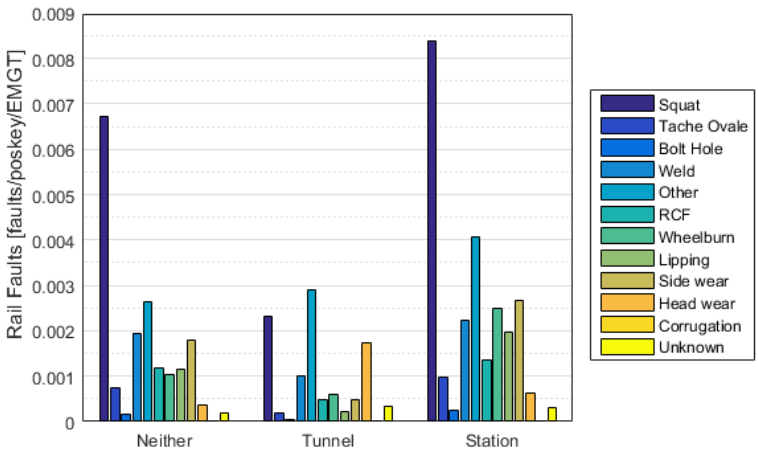


Figure 5.9: Rail fault rates split by the existence of tunnels and stations

Rail Joint Type

It was thought that the additional forces produced as a train travels over the joints in jointed rail would increase the rate of faults. Looking at the rates in Table 5.8, which are also shown in Figure 5.10, it can be seen that the rate of faults is 87% lower on CWR. Looking at individual fault types it could be seen that, as expected, bolt hole faults see the largest reduction of 97.5% (much fewer bolt holes in CWR) followed by wear-related faults. The reason for such a large reduction of wear faults is possibly due to the rail ages with jointed rail tending to be older than CWR as well as CWR being used on larger rail. To remove some impact of this, the analysis was undertaken using the same type of rail installed over a similar time period.

It was decided to compare using 95 lb BH rail as these had numerous faults on both jointed and CWR methods. 95 lb BH was primarily installed between 1950-1965, which was also the same period in which CWR started to become more common. Looking at the results in Table 5.9 and Figure 5.11, it can be seen that the variances between the joint types (64.6%) is much lower than in Table 5.8. This shows that the results in the first table were being impacted by the difference in rail ages and types between the joint types. Despite this the results demonstrate the large reduction in faults seen when moving from jointed rail to CWR. Looking at the individual faults in Table 5.9, again bolt holes is shown to have the largest reduction. Wear-related faults (lipping, side and head wear) still see a large reduction, showing that the additional movement and forces that occur as trains traverse the joints lead to high levels of wear. Weld is the only fault type that saw an increase from the move to CWR, which is expected due to the increase in welds added into the rail.

Table 5.8: The effect of rail joint type on the rates of different rail faults [Faults/Poskey/ Equivalent Million Gross Tonnage]

	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
Jointed	2.105e-2	2.413e-3	2.283e-3	5.546e-3	1.930e-3	6.086e-4	7.605e-4
CWR	5.274e-3	6.153e-4	5.395e-5	1.735e-3	1.800e-3	7.620e-4	5.420e-4
% Variance	-75.0	-74.5	-97.6	-68.7	-90.7	-87.5	-92.9
	Lipping	Side wear	Head wear	Corrugation	Unknown	Total	
Jointed	1.079e-2	1.703e-2	3.701e-3	1.098e-5	1.684e-3	9.749e-2	
CWR	5.836e-4	9.130e-4	2.304e-4	1.864e-6	1.234e-4	1.263e-2	
% Variance	-94.6	-94.6	-93.8	-83.0	-92.7	-87.0	

Table 5.9: The effect of rail joint type on the rates of different rail faults on 95 lb Bullhead rail installed between 1950-1965 [Faults/Poskey/Equivalent Million Gross Tonnage]

	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
Jointed	1.61e-2	1.54e-3	4.33e-3	4.91e-3	1.99e-2	1.69e-3	1.86e-2
CWR	1.08e-2	8.11e-4	5.80e-4	5.45e-3	8.00e-3	3.48e-4	1.74e-3
% Variance	-33.2	-47.3	-86.6	10.9	-59.9	-79.4	-90.7
	Lipping	Side wear	Head wear	Corrugation	Unknown	Total	
Jointed	7.79e-3	8.40e-3	1.92e-3	0.00e0	2.41e-3	8.77e-2	
CWR	2.09e-3	3.48e-4	1.16e-4	1.16e-4	4.64e-4	3.11e-2	
% Variance	-73.2	-95.9	-81.9		-80.8	-64.6	

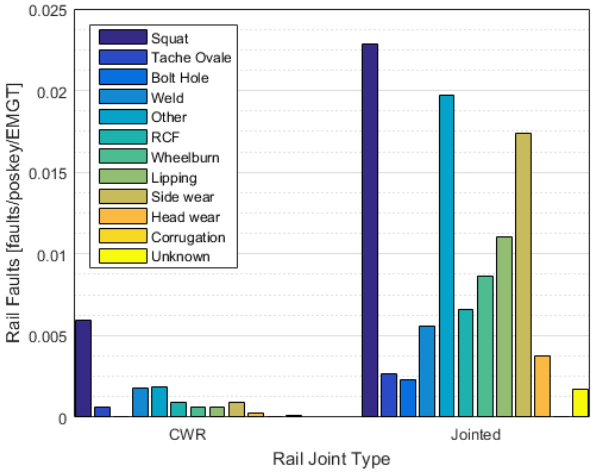


Figure 5.10: Rail fault rates split by rail joint type

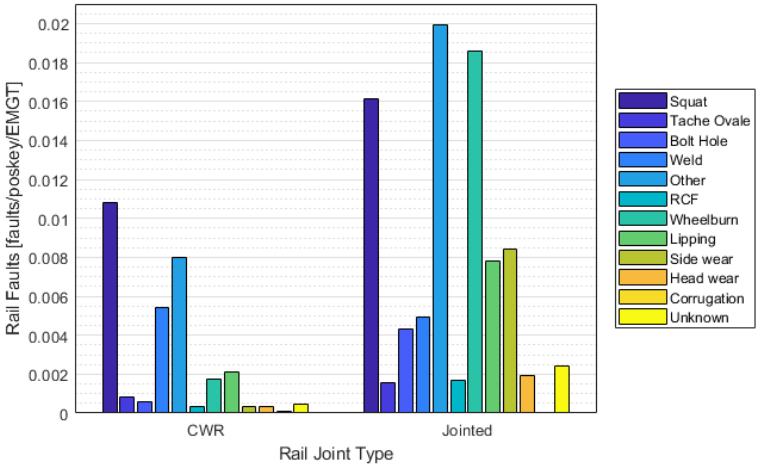


Figure 5.11: Rail fault rates split by rail joint type [95 lb Bullhead Rail, Installed 1950-65]

Track Geometry

A railway track is a complex system made up of many assets, ballast, rail and sleepers, which degrade differently. It was thought that the individual conditions of the assets may have an impact on the degradation rate of the other assets. The primary link was thought to be the track vertical geometry and the rate of rail faults, as more asperous track would endure higher dynamic forces. In order to analyse a possible link the linear geometry degradation fits from Chapter 3 were used to estimate the track geometry when each rail fault occurred. The geometry fits were also used with the usage data to categorise the average usage which occurred over track between certain geometry bands (Standard Deviation (SD) every 1 mm). With this it is possible to calculate the rate of faults that occur in different geometry bands.

The results are shown in Table 5.10 and in Figures 5.12 and 5.13, where the numbers above the bars are the number of faults. The results for all faults show an increasing rate of rail faults as the track geometry worsens, with the rate also increasing with worsening track vertical geometry. Looking closely at the individual fault types it can be seen that all types of faults see an increasing rate as the SD of the vertical geometry increases. Squats, tache ovales and welds have a more linear relationship than the other faults. Bolt holes show a large step increase of 4-5 mm, suggesting that the increased dynamic forces at this point starts to damage the bolt holes, with very little damaged being caused on track with better vertical geometry. Wear-related faults (lipping, side wear and head wear) seem very dependent on the track geometry with very low rates seen below 4-5 mm, which increase rapidly above this. This makes sense as when the track geometry worsens, the traversing trains tend to move on the rail more (side to side motion) with greater force accelerating the rate of wear.

To ensure that the results were not being skewed by rail type, joint or age it was decided to perform the same analysis but on smaller sets of data with consistent factors. The most common rail types were analysed to maintain large datasets. This included UIC60 CWR installed between 2000-2010 (Figure 5.14), 113 lb FB CWR installed between 1990-2010 (Figure 5.15) and 113 lb FB CWR installed between 1980-1990 (Figure 5.16). Looking at these figures it can be seen that they follow a similar trend to the results obtained when all the data was used in Figure 5.12, showing that the rail type, joint and age was not affecting the shape of the results. By reducing the dataset size it was not possible to have confidence in the results at an individual fault level.

Table 5.10: The effect of vertical track geometry on the rates of different rail faults [Faults/Poskey/Equivalent Million Gross Tonnage]

Geo SD [mm]	Squat	Tache Ovale	Bolt Hole	Weld	Other	RCF	Wheelburn
0-1	1.35e-3	1.34e-4	1.03e-5	4.05e-4	2.37e-4	1.01e-4	2.54e-4
1-2	3.79e-3	3.94e-4	1.17e-5	1.26e-3	9.37e-4	2.51e-4	3.22e-4
2-3	5.98e-3	8.00e-4	5.34e-5	2.01e-3	1.92e-3	5.51e-4	6.82e-4
3-4	7.38e-3	9.03e-4	2.05e-4	2.65e-3	3.33e-3	1.00e-3	1.11e-3
4-5	9.48e-3	1.28e-3	7.43e-4	3.32e-3	6.14e-3	1.54e-3	2.50e-3
5-6	1.23e-2	1.56e-3	7.58e-4	3.72e-3	9.07e-3	2.82e-3	3.86e-3
6-7	1.58e-2	2.11e-3	9.28e-4	3.97e-3	1.32e-2	4.39e-3	6.24e-3
7-8	2.19e-2	2.40e-3	9.02e-4	6.31e-3	1.86e-2	4.81e-3	7.52e-3

Geo SD [mm]	Lipping	Side wear	Head wear	Corrugation	Unknown	Total
0-1	1.86e-5	1.24e-5	1.03e-5	0.00e+0	8.26e-6	2.54e-3
1-2	5.15e-5	5.64e-5	3.50e-5	2.26e-6	4.10e-5	7.15e-3
2-3	2.59e-4	3.47e-4	1.23e-4	7.31e-7	1.18e-4	1.28e-2
3-4	5.76e-4	9.97e-4	3.84e-4	4.27e-6	1.77e-4	1.87e-2
4-5	1.28e-3	2.48e-3	7.76e-4	6.69e-6	4.02e-4	3.00e-2
5-6	2.41e-3	4.12e-3	1.40e-3	4.73e-5	6.15e-4	4.26e-2
6-7	6.92e-3	1.04e-2	3.80e-3	0.00e+0	1.77e-3	6.95e-2
7-8	3.61e-3	1.77e-2	3.31e-3	0.00e+0	2.40e-3	8.96e-2

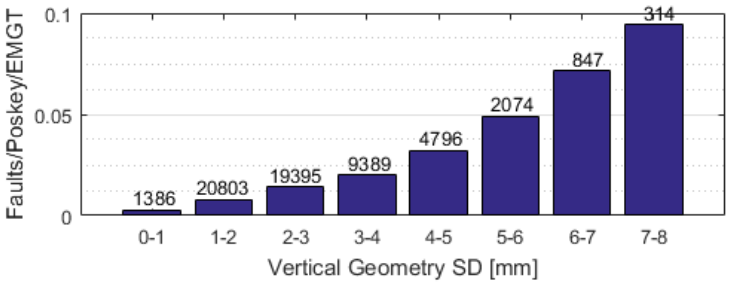
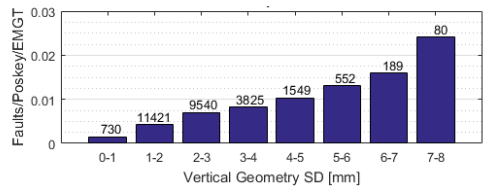
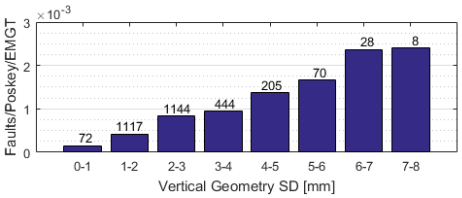


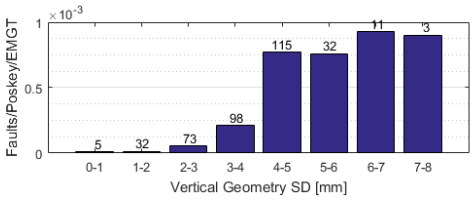
Figure 5.12: Effect of vertical track geometry on the occurrence rate of rail faults



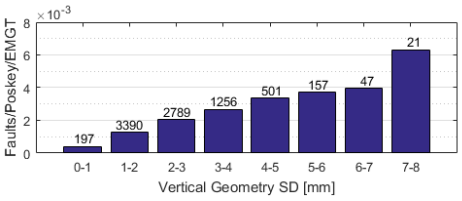
(a) Squat



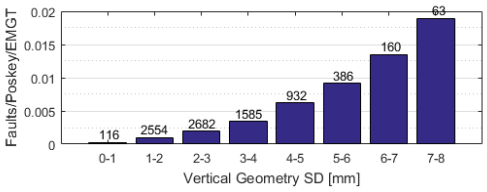
(b) Tache Ovale



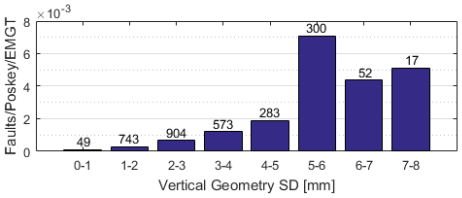
(c) Bolt Hole



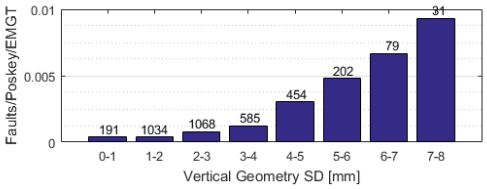
(d) Weld



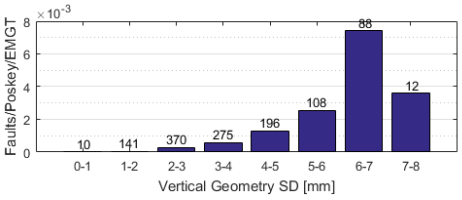
(e) Other



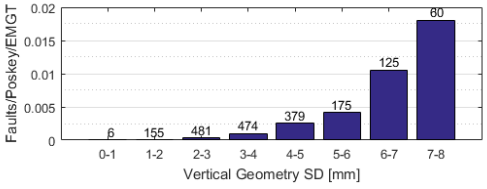
(f) Rolling Contact Fatique



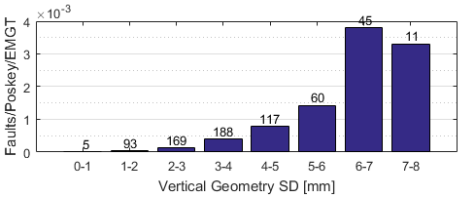
(g) Wheelburn



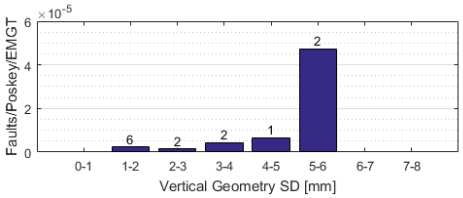
(h) Lipping



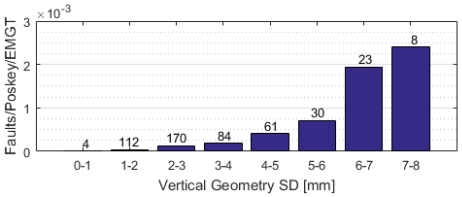
(i) Side Wear



(j) Head Wear



(k) Corrugation



(l) Unknown

Figure 5.13: Effect of vertical track geometry on the occurrence rate of different rail fault types

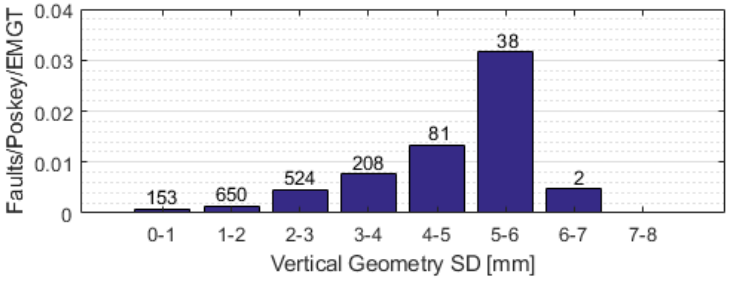


Figure 5.14: Effect of vertical track geometry on the occurrence rate of rail faults on UIC 60 Continuous Welded Rail rail installed between 2000-2010

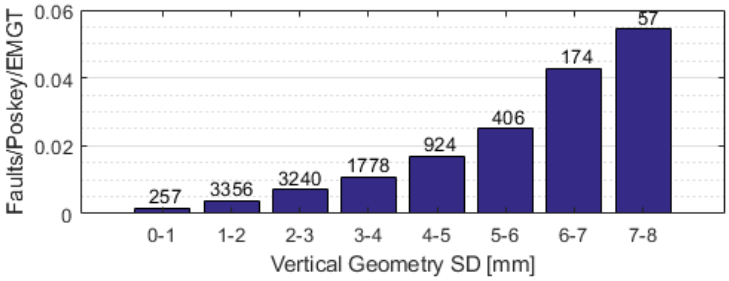


Figure 5.15: Effect of vertical track geometry on the occurrence rate of rail faults on 113 lb Flatbottom Continuous Welded Rail rail installed between 1990-2010

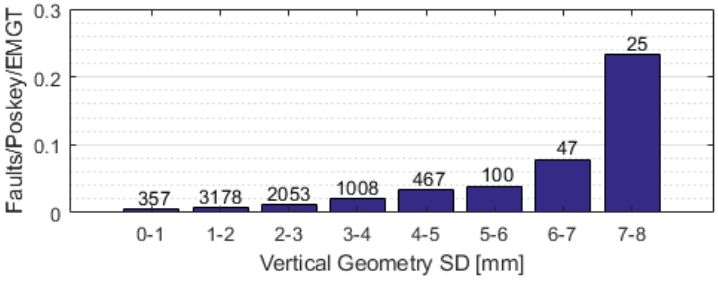


Figure 5.16: Effect of vertical track geometry on the occurrence rate of rail faults on 113 lb Flatbottom Continuous Welded Rail rail installed between 1980-1990

5.4 Model Relating Track Geometry to Rail Faults

The analysis has shown that the rate of faults is heavily dependent on the track vertical geometry SD. Due to this it is required to be able to calculate the rate of each fault type for all possible track qualities. The link is important to model because it would allow the impact on rail faults to be calculated for decisions which affect track geometry, such as changing the tamping thresholds. To create the model the results from Table 5.10 on page 199 have been related to the mid-points of each vertical geometry band.

The proposed model is stochastic, with a random aspect to account for the chance of a fault having occurred, and is designed to fit in a larger asset management model which is solved using Monte-Carlo, such as the one proposed in Chapter 6. If a probability of a rail fault was 0.01/poskey/EMGT, and the section of track being analysed was 10 poskeys long (2200 yards) over a period of 3 EMGT, the probability of a fault would be $0.01 * 10 * 3 = 0.3$ or this can be said as 0.3 faults would occur. If a uniform random number generator between 0-1 was used, any number above 0.3 would mean no fault occurred whereas if the random number was lower a fault occurred. We know that the fault rate is not as simple as a single number, with the link to track geometry needing to be included. Short time periods (small usage) and/or shorter sections of track should be used to ensure the probabilities stay below 1.

This is accomplished by fitting equations to the relationships seen previously in Section 5.3.2. It was decided to stack the data, hence a line is fitted to the data from squats, then a line is fitted to the sum of squats and tache ovale. By doing this we can simplify the decisions within the model. If the probability of a squat was 0.004 and the probability of a squat or tache ovale was 0.006, then the random number generated just needs to be checked in order of the stacking. For example; the random number sampled is 0.005 then we check if it is less than the rate for squats (0.004), which it is not, then we check if its less than squats or tache ovale (0.006), which it is, hence a tache ovale occurred as we already know it was not a squat. For the purpose of readability the stack groups will be known by the key in Table 5.11.

For each group a polynomial fit to an order of three was used to link the vertical track geometry to the rail fault rates. A polynomial was used as the analysis from Section 5.3.2 showed that the link was not linear, and the polynomial fit allows for many shapes. It was decided to hold an intercept through the origin as this improved the fits (visually they make more sense), stopping the fits going negative (fault rates values less than zero). Linear, polynomial (order two and three), exponential and power fits were tested but the polynomial fit with an order of three was best for all fault groups. Comparing the fits of polynomial order two and three showed an average r^2 improvement of 1% and a reduction in the RMSE of 53% whilst still visually fitting the trends in the data. The polynomial order 3 fits are shown in Figures 5.17 and 5.18. A summary of the polynomial order three fits are shown in Table 5.12 with the polynomial form denoted in Equation 5.3, where Geo is the vertical geometry SD in mm. This equation will be known as poly3, hence the equation for the rail fault rate of G-Squat will be poly3-G-Squat(Geo). The stacked results of the fits are compared to the actual in Figure 5.19, in which it can be seen how closely it resembles the actual data.

$$\text{Rail Fault Rate [/poskey/EMGT]} = A*(Geo)^3 + B*(Geo)^2 + C*(Geo) + D \quad (5.3)$$

Table 5.11: Stacked Rail Fault Groups Key

Group	Included Rail Fault Types
G-Squat	Squat
G-TacheOvale	Squat or Tache Ovale
G-BoltHole	Squat, Tache Ovale or Bolt Hole
G-Weld	Squat, Tache Ovale, Bolt Hole or Weld
G-Other	Squat, Tache Ovale, Bolt Hole, Weld or Other
G-RCF	Squat, Tache Ovale, Bolt Hole, Weld, Other or RCF
G-Wheelburn	Squat, Tache Ovale, Bolt Hole, Weld, Other, RCF or Wheelburn
G-Lipping	Squat, Tache Ovale, Bolt Hole, Weld, Other, RCF, Wheelburn or Lipping
G-SideWear	Squat, Tache Ovale, Bolt Hole, Weld, Other, RCF, Wheelburn, Lipping or Side Wear
G-HeadWear	Squat, Tache Ovale, Bolt Hole, Weld, Other, RCF, Wheelburn, Lipping, Side Wear or Head Wear
G-Corrugation	Squat, Tache Ovale, Bolt Hole, Weld, Other, RCF, Wheelburn, Lipping, Side Wear, Head Wear or Corrugation
G-Unknown(All)	Squat, Tache Ovale, Bolt Hole, Weld, Other, RCF, Wheelburn, Lipping, Side Wear, Head Wear, Corrugation or Unknown

Table 5.12: Rail fault rate against track vertical geometry polynomial fits

	G-Squat	G-TacheOvale	G-BoltHole	G-Weld	G-Other	G-RCF
A	7.64E-5	7.85E-5	6.90E-5	9.38E-5	1.18E-4	1.13E-4
B	-6.45E-4	-6.55E-4	-5.45E-4	-7.96E-4	-6.94E-4	-5.46E-4
C	3.44E-3	3.73E-3	3.56E-3	4.85E-3	5.22E-3	5.07E-3
D	0.00E+0	0.00E+0	0.00E+0	0.00E+0	0.00E+0	0.00E+0
r2	0.999	0.999	0.999	0.998	0.999	1.000
RMSE	2.38E-4	2.37E-4	2.43E-4	5.53E-4	6.29E-4	3.81E-4
	G-Wheelburn	G-Lipping	G-Sidewear	G-Headwear	G-Corrugation	G-Unknown(All)
A	1.12E-4	7.34E-5	1.66E-4	1.61E-4	1.60E-4	1.70E-4
B	-3.87E-4	1.31E-4	-3.48E-4	-1.98E-4	-1.92E-4	-2.31E-4
C	4.93E-3	3.87E-3	4.61E-3	4.28E-3	4.27E-3	4.33E-3
D	0.00E+0	0.00E+0	0.00E+0	0.00E+0	0.00E+0	0.00E+0
r2	1.000	0.997	0.997	0.996	0.996	0.995
RMSE	4.23E-4	1.59E-3	1.78E-3	2.34E-3	2.33E-3	2.48E-3

5.4.1 Step by Step Approach

Below describes a step by step approach to assessing

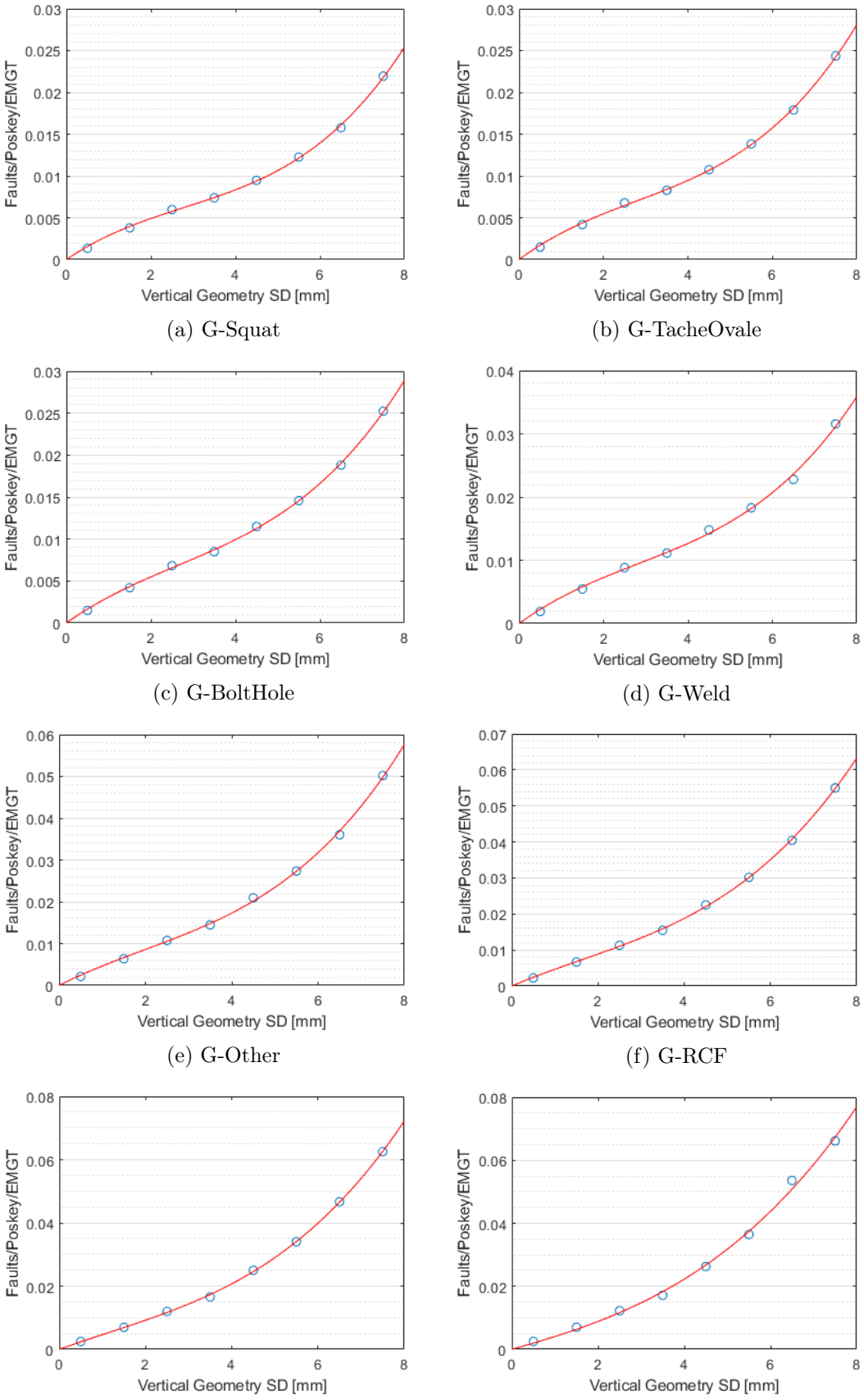
1. Decide the length of track in poskeys (L) and amount of traffic in the period being modelled in EMGT (T).
2. Estimate the expected average vertical geometry SD over the period (Geo)
3. Sample a random number between 0-1 (uniform distribution) (R)
4. Calculate poly3 for each fault type using Geo, and multiply by L and T
5. If R $<$ poly3-G-Unknown(All)(Geo)*L*T then a fault occurred, continue to check the fault type, if R $>$ poly3-G-Unknown(All)(Geo)*L*T no fault occurred
6. If R $<$ poly3-G-Squat(Geo)*L*T then a squat occurred
7. If R $<$ poly3-G-TacheOvale(Geo)*L*T then a tache ovale occurred
8. Continue checking each fault type in order of the stacking until the fault type is found

A graphical example of how the model works is shown in Figure 5.20. In the example the track vertical geometry SD was 4.2, and the amount of traffic was 1 EMGT and the amount of poskeys was 1. If a random number of 0.013 was sampled then the model predicts weld fault (where the red lines cross), whereas if 0.035 was sampled the model reports no fault occurring.

5.4.2 Limitations

The model proposed has some limitations which are listed below:

1. Cannot predict more than one fault appearing in the assessed period, hence a recommendation of short track lengths and low usage is recommended to reduce the impact. This reduces the impact as the probability of more than one fault occurring would be negligible.
2. Does not take into account rail and joint type or age. These could easily be added with more data, but for the purpose of this model it was decided that splitting the data into smaller datasets would reduce the strength of the geometry link, which was the main aspect that was being modelled.



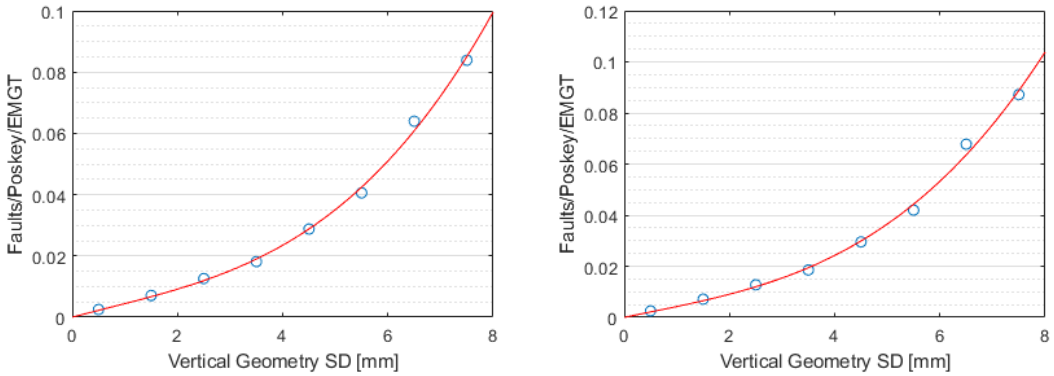
(a) G-Squat (b) G-TacheOvale

(c) G-BoltHole (d) G-Weld

(e) G-Other (f) G-RCF

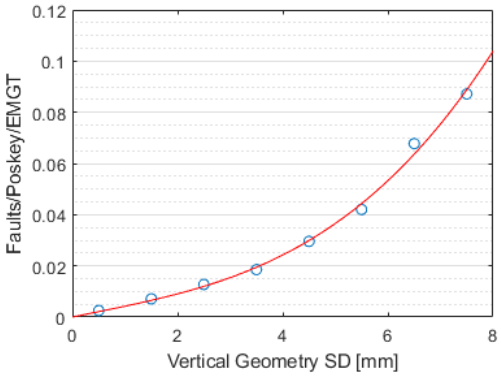
(g) G-Wheelburn (h) G-Lipping

Figure 5.17: Vertical track geometry impact on rail fault rates, polynomial order 3 fits (Part 1)

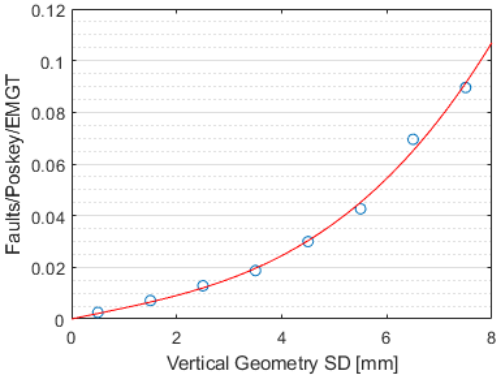


(a) G-Sidewear

(b) G-Headwear

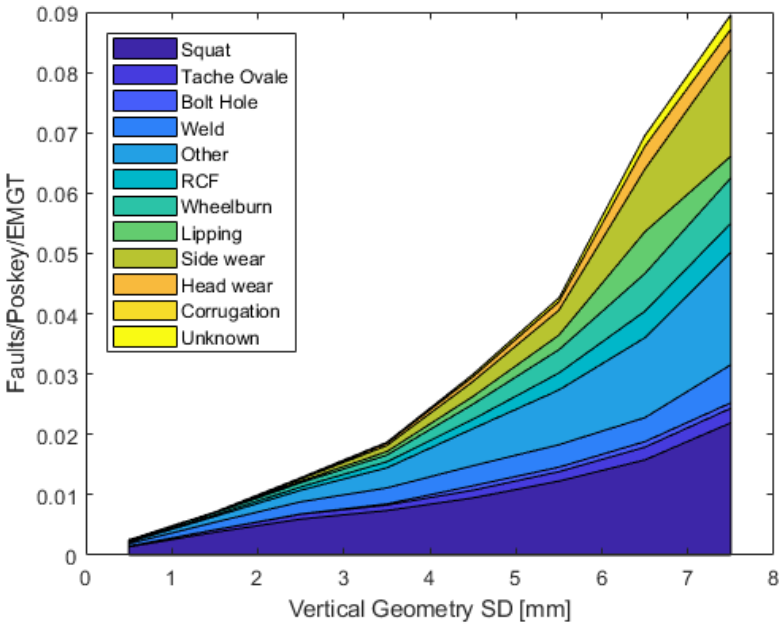


(c) G-Corrugation

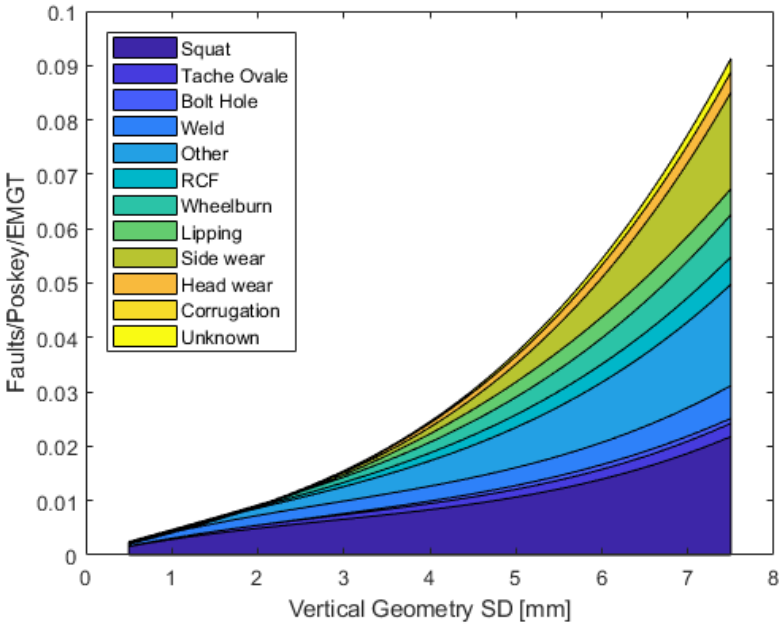


(d) G-Unknown(All)

Figure 5.18: Vertical track geometry impact on rail fault rates, polynomial order 3 fits (Part 2)



(a) Actual Data



(b) Model

Figure 5.19: Stack rail fault rates against vertical track geometry. Model compared to actual data.

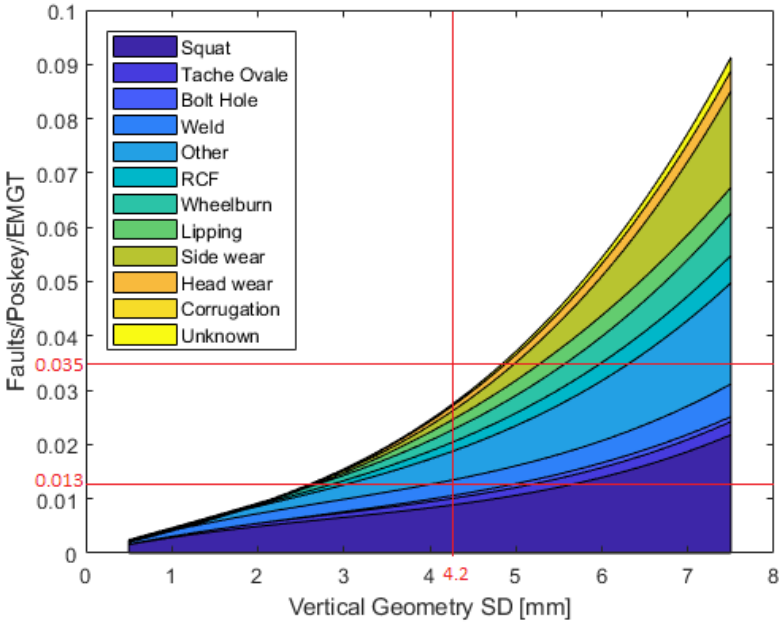


Figure 5.20: Rail Fault Model Example

5.5 Summary

In this chapter rail faults and the factors which influence them have been investigated. The analysis has shown that the distinct fault types are impacted differently by the factors explored. It has also demonstrated that the occurrence rate of rail faults is impacted by many factors, including age, rail and joint type, curvature, tunnels, stations and track vertical geometry.

The most common types of faults on the network are squats (35%), other (16%) and weld (11%). The types of faults that are most likely to results in an ESR are bolt hole (43%), followed by tache ovale (9%), other (6%) and RCF (4%). Faults can also turn into rail breaks with the highest percentage being from bolt hole (6%) then other (2%), welds and tache ovales (1%).

There are two main types of inspections, visual and ultrasonic, where ultrasonic is used to detect internal defects such as tache ovale and squats and visual is used to detect external defects like head and side wear. There are also three main types of maintenance: weld, grind or replace the rail. Wear-related defects such as lipping tend to be ground, whereas squats are commonly welded and bolt hole faults require replacements.

The average rate of rail faults per poskey (220yd track section) per EMGT of traffic is 0.016, meaning a fault occurs every 62 EMGT which traverses a poskey.

The rail age was shown to have a linear impact on the rate of faults, with older rail experiencing higher rates. Fatigue related faults (squats and tache ovales) were most

affected but wear-related faults were also highly dependent on the rail age. The impact of age is high with fault rates on rail installed in 2000-10 being 60% lower than rates in similar rail installed in 1960-70.

Track curvature was seen to reduce the rate up to 0.001 m^{-1} compared to straight track, but the rate of faults increased significantly above this. The reason that the rate was lower on slightly curved track is thought to be due to the inclusion of cant in track designs as well as lower speeds on curved sections. The results of the analysis showed that fatigue-related faults were more related to track curvature than wear.

The results of the analysis exploring the impact of rail being installed in tunnels or at stations, showed a reduction in faults in tunnels, but an increase at stations. The reduction in tunnels was expected as the rail is protected from the elements, and the increase at stations is most likely due to the dynamic forces produced from trains accelerating and braking.

Rail type was shown to impact all faults with small increases in the size of the rail having large impacts on the rate of faults. UIC60 (the newest type of rail type) is seen to have a big step improvement over the commonly-used 113 lb FB rail, nearly halving the rate of faults.

The type of rail joint is the factor that had the largest impact on the rate of faults, with CWR showing a large reduction in faults compared to the legacy jointed rail. As expected, bolt hole faults show the largest reduction which is very important due to the percentage of these faults that lead to ESRs and rail breaks.

It was thought that the track vertical geometry increasing the rate of faults due to the unevenness of track with poor geometry would increase the forces experienced by the rail. The analysis proved this to be the case, with the rate increasing as the vertical geometry worsened. The results showed that this relationship was not linear, with the rate at which the fault rate was increasing also increasing with worsening geometry. Squats, tache ovales and welds had a more linear relationship whereas wheelburn, lipping, side wear, and head wear had a step change at geometry SD of 4-5 mm with very low fault rates being seen on lower levels of vertical track geometry SD.

The proposed stochastic model to link the track geometry to the rate of faults closely matches the data, while taking into account track length and traffic.

With an understanding of the factors which influence the track geometry degradation, maintenance effectiveness and rail faults it is possible to design a railway track asset management model, which will be explored in the following chapter.

Chapter 6

Railway Track Asset Management Model

6.1 Introduction

Good quality railway infrastructure is an important economic driver for any country. Due to this it is imperative that it is kept in a good condition but often this is carried out with very limited financial resources. To maximise the benefit seen from the expenditure it is important to make the correct asset management decisions leading to efficient maintenance regimes and the correct asset choices. To enable efficient asset management decisions to be made, modelling is used to estimate the impact of decisions and compare different options. Modelling can also be used to predict the future costs of maintaining an asset, enabling future planning.

In this chapter a Petri Net model is introduced which has been designed as an asset management tool for railway track. The knowledge obtained from the literature review in Chapter 2 and the analysis performed in Chapters 3, 4 and 5 is used to help to design a model. The model is populated using the data obtained from Network Rail (NR), which is introduced in Chapter 3.

6.2 Modelling Technique Choice

It was decided to utilise a Petri Net (PN) modelling approach which is introduced in Section 2.4.2. Due to the high levels of variability seen in the data in the previous chapters, it was felt that a stochastic model would best suit a railway track asset management model. Out of the possible modelling techniques, Coloured Petri Nets (CPNs) were chosen due to their highly flexible nature, working as a model, or a framework which smaller models, such as the track geometry maintenance model discussed in Section 4.4, can be inserted. Other proven uses are discussed by Jensen (1997), including, communication

protocols, distributed systems, imbedded systems, automated production systems and work flow analysis.

The model has been built in a free-ware computer software called CPN-Tools, which was based on work by Jensen and Kristensen (2009).

6.3 Introduction to Coloured Petri-Nets

CPNs are an extension of PNs which add a wide array of extra functionality. Similar to PNs such as the one introduced by Andrews (2012), which is discussed in Section 2.4.2, CPNs still consist of places which tokens move between via along arcs and via transitions.

Tokens

CPNs add coloured sets which are linked to tokens and can contain information, which moves around the net with the token. Each coloured set is defined and can only contain one type of data, such as boolean, integer, real number (decimals) and string (text). Each token can have many coloured sets attached to it, these are known as multisets. The multisets attached to tokens can change as the tokens move through the net with the information in the mutliset changing or even coloured sets being added or removed. This is important for a railway track model as historical and classification data, such as maintenance histories and track types, are important factors.

Places

Places are denoted by circles and contain tokens. Each place can only contain one type of multiset tokens, which is known as its colour set inscription. This is denoted to the bottom right of a place. For the purpose of this thesis all inscriptions will be denoted by the data types, i.e. if a place allowed multisets consisting of two real numbers (R), an integer (I), three boolean (B) and a string (S) such as (1.2, 3.4, 5, TRUE, TRUE, FALSE, "Test"), the inscription would be RRIBBBS.

Places also contain the initial markings, which are shown at the top left of the place. These are the initial locations of the tokens and the information they contain in their coloursets. An example of initial marking would be 1‘(1,TRUE)++2‘(3,FALSE), which creates three tokens of a multiset of one integer and one boolean, IB. The multiset is denoted by the part in brackets, with the quantities of tokens with these multisets marked in front of the brackets. In the markings ++ means "and".

Transitions

Transitions allow tokens to move between places and allow time to pass (for processes that requires time to complete) or can be used to make decisions. The top right of each transition shows the time inscription which describes the time it takes for the transition to fire (move a token), with transitions denoted by rectangles. It is denoted with @+, i.e. @+60 would mean that it takes 60 time steps to fire. By utilising Monte-Carlo simulation (random number generation), it is possible to use distributions to describe the time period for a transition to fire. This means that a time such as @+N(100, 10) would denote the time to fire being related to a normal distribution with a mean of 100 (time steps) and a Standard Deviation (SD) of 10 (time steps), hence each time the transition is activated (has a token to fire) the time to fire varies.

To enable transitions to make decisions, guards are used. These are rules which the inputting tokens are required to pass for a transition to fire. This is useful for ensuring transitions only fire certain types of tokens or tokens containing certain information. Due to limitations with CPN-Tools guards (in particular the lack of support for random values or time related functions), they have not been used in the proposed model, with code segments used instead.

Transitions also have code segments. These calculate when a transition is activated (all input arcs have required tokens). The calculation can use any of the information contained within the coloured sets of the inputting tokens. This allows token information to be changed, as well as new colour sets and decisions to be made. Code segments are written in a coding language called Standard Meta Language (SML), but for the purpose of the diagrams in this chapter normal mathematical expressions have been used instead to maintain legibility. Code segments allow normal mathematical models to be inserted, into a CPN. Code segments plus conditional output arcs can give the same result as a guard, where the code segment results are used to return the token back to the original place if a condition is failed.

Transitions also have priorities which are recorded in the bottom left. These tell the model what order to activate transitions if multiple transitions have the required inputs at the same time. In this chapter five priorities are included, and are denoted as either lowest, low, medium, high or highest, with highest always firing first.

Arcs

There are two main types of arcs, normal and inhibitor. Normal arcs connect places and transitions and create the net that the tokens move around. These are denoted as lines with arrows showing the direction of movement. Normal arcs in CPNs have a large amount of flexibility. They bind the token colour sets to variables which are passed to the transition. These variables can then be used in the transition as part of a guard or

code segment. The variables are then used on the output arcs to create new tokens with differing colour sets to the inputs. When a transition has multiple inputs, if the same variables appear on different input arcs then the variable values have to be equal. This allows the model to choose which tokens to pull through the transition based on other inputs, instead of picking a random token from each input place.

Output transitions have added functionality, with the ability to run functions via conditional arc inscriptions or even add a time for the token to pass through the arc, known as an arc delay. Conditional arc inscriptions allow decisions to be made on what the output tokens should contain, i.e. if an output arc inscription was "if y = 2 then 2‘(1) else if y = 1 then 1‘(1) else empty" then the output would be based on the value of y (either from the input arc or from a transition code segment). If y = 2 then two tokens with an integer colour set with value 1 are sent to the output place, if y = 1 then one token with an integer colour set with value 1 is sent to the output place, if y = any other value no token is outputted. Conditional arc expressions along with transition code segments add a high level of functionality to CPNs, allowing any complexity of guards to be created. Output arc delays are extremely useful allowing a transition to take different amounts of time for each separate output. Arc delays can be conditional and are also related to the input variables.

Inhibitor arcs are a feature of simple PNs but still have uses in CPNs. An inhibitor links a place to a transition but tokens cannot move along it, instead while the connecting place has a token the connecting transition cannot be activated.

Hierarchical Features

CPNs can be hierarchical. This means that a transition can be used to represent another PN known as a subnet. This allows large complex nets to be split into smaller sections, so they are easier to draw and understand. Additionally, a net can be simplified by having repeating parts of a net design removed and the one subnet used instead. The top of the hierarchy is known as the superpage, and those below which contain subnets are called subpages. A page can be a superpage and a subpage if it has both levels above and below (is a subnet and also contains transitions to further subnets). Hierarchical transitions are denoted by a double outline and are known as substitution transitions. Instead of the transition firing as normal, the token moves into the subnet. The substitution transition has to have the same inputs and outputs places as the subnet, with all connecting places on both nets. Places which are in subnets which are related to a place in the higher level nets are denoted by a double outline and are connected such that a change to one (on any page) happens to all of them, these places are known as fusion places. The fusion place on a superpage is known as a socket, whereas when it occurs on the subpage it is known as a port, (CPN).

6.4 Coloured Petri-Net Design

6.4.1 Overview

The hierarchical CPN discussed in this section is designed to act as an asset management tool for railway track, replicating the degradation of the track as well as maintenance decisions. Figure 6.1 shows the top-level of the CPN. The model consists of four main parts: "Track Geometry Model", "Rail Fault Model", "Planned Upgrades (Speed and Assets)" and "Asset Upgrades at Renewal". Within the nets are:

1. Variables, which the colour sets are attached to when the tokens move along arcs. These are written as combinations of letters based on what information they hold and appear in brackets on arcs as well as in code segments. Throughout this chapter all variable will be written in round brackets within the text. A full list of the variables within the model can be seen in Table 6.2.
2. Initial markings, which show the initial tokens in the model. These are denoted in red to the top left of each place which tokens start in. Within the explanation of the CPN model initial markings are denoted in square brackets. A full list of initial markings with examples can be seen in Table 6.1.
3. Functions, which are denoted by F! and occur within code segments and conditional arc inscriptions. A full list of the functions within the model are included in Table 6.4.
4. Global variables, are values stored in the model which do not change within the simulations. These are outlined in Table 6.3.

The model starts in place P1, where the information about the railway line is entered as tokens, with a token per track section. This input is known as [track] and consists of: unique section number (s), vertical geometry SD at the last inspection (q), number of previous tamping and stoneblowing operations (mt, ms), time since the last inspection (t), track speed (ts) and track asset types (tt) for each section to be modelled. An example input could be: track section 1, which at the last inspection, which was 34 days ago, had a vertical geometry of 1500 um, has a track type of 2 (large concrete sleepers on a track speed between 80-110 MPH), a track speed of 9 (110 MPH), and has undergone 3 previous tamping actions and no stoneblowing. This track section would have an initial marking of 1‘(1, 1500.0, 3, 0, -34, 2, 9), following the order of the variables on the output arc of P1 and P1’s colour set inscription.

With a token for every track section to be modelled in P1, transition T1 is activated. T1 takes each token from P1, and calculates a random vertical geometry degradation rate (d) based on the track type and maintenance history by sampling a Weibull distribution. The code segment attached to T1 first checks if stoneblowing had previously occurred on the section (last maintenance was stoneblowing) ((ms)>0). If stoneblowing had occurred the model samples a degradation rate from the Weibull distribution with parameters from the tables in Section 8.2.25 using the corresponding track assets and stoneblowing

history (columns labelled "After Stoneblowing") (function F!SD). If no stoneblowing had occurred previously on the track section but there had been previous tamping operations, the degradation results from after tamping operations are used (F!TD) and if no previous tamping or stoneblowing operations were recorded (mt = 0 and ms = 0) then the results after a renewal are used (F!RD). Again the tables in Section 8.2.25 are referenced but the columns "After Tamping" are used instead. A degradation rate is sampled from the Weibull distribution using the inverse Cumulative Distribution Function (CDF) with a uniform random number between zero and one. The Weibull parameters in the tables in Section 8.2.25 are included within the model. Using the information of track type (tt) stored in the track section tokens, the model function use the corresponding table in Section 8.2.25 for the Weibull parameters. When T1 fires the tokens into P2 it brings through all the input arc information (variables) as well as the sampled degradation rate (d). An additional boolean colour set is added with a value of FALSE, this is to mark that no maintenance has been scheduled for the track section yet and is required for the "Track Geometry Model" subpage.

Once all the track sections have had a degradation rate sampled and moved to P2, the substitution transition track geometry model, ST1, is activated. Throughout the model copies of all the track sections will always be in P2, with the values in the colour sets being changed by the models in the subnets. P2 is a socket of a fusion place that has ports on all subpages, linking them all together.

Table 6.1: CPN model initial markings

Name	Variables Included	Purpose
upgradeS	s,Uts,Ut	List of sections with times to track speed changes
upgradeT	s,Utt,Ut	List of sections with times to track type changes
upgradeR	s,Utt,Ut	List of sections with times to track type changes also has to contain everything from upgradeT
inspectT	s,TRUE	List of all track section numbers with TRUEs attached (inspection undertaken)
inspectF	s,FALSE	List of all track section numbers with FALSEs attached (inspection required)
renewals	s,tt	List of all track section numbers and their track type at the start of the model
track	s,q,mt,ms,t,ts,tt	Full information on each track section to be modelled
sectionEnds	s	Contains two tokens, one for the maximum value of s and one of the lowest

Table 6.2: CPN model variables

Variable	Type	Purpose
s, $s2$	Integer	Poskey ID Number. The track IDs of the line modelled have to be integers and sequential i.e. 1, 2, 3, 4 etc. This is due to the model presuming section ID 9 is between sections IDs 8 and 10.
Utt	Integer	Track type to upgraded to
Uts	Integer	Track speed to upgrade to
Ut	Integer	Time to upgrade
q, $q2$	Real	Initial Track Vertical Geometry SD [μm, micrometre]
d, $d2$	Real	Current Vertical Geometry Degradation Rate [μm/Equivalent Million Gross Tonnage (EMGT)]
mt, $mt2$	Integer	Number of Previous Tamping Maintenance Actions
ms, $ms2$	Integer	Number of Previous Stoneblowing Maintenance Actions
t	Integer	Time of last geometry maintenance
m	Boolean	Marks if vertical geometry maintenance has been scheduled
u	Boolean	True if urgent maintenance
M	Integer	Maintenance type; 3 = renewal, 2 = stoneblow and 1 = tamping
tt, $tt2$	Integer	Track type
ts, $ts2$	Integer	Track speed
f	Integer	Fault type
ft	Integer	Fault time
u	Real	Usage since the last rail fault check
Q	Real	Current vertical geometry
fmt	Integer	Rail maintenance type; 1 = grind, 2 = weld and 3 = rerail

Table 6.3: CPN model global variables

Variable	Purpose
CT	Current Model Time
TI_mu	Average time between track geometry inspections (normal distribution mu)
TI_sigma	Variation in time between track geometry inspections (normal distribution sigma)
TM_mu	Average time to perform track geometry maintenance (normal distribution mu)
TM_sigma	Variation in time to perform track geometry maintenance (normal distribution sigma)
$Urgent\_thresh$	Table of vertical geometry SD threshold values for urgent maintenance which relate to the track speed.
$Normal\_thresh$	Table of vertical geometry SD threshold values for normal maintenance which relate to the track speed.
Opp_thresh	Table of vertical geometry SD threshold values for opportunistic maintenance which relate to the track speed.
RIV_mu	Average time between visual rail inspections (normal distribution mu)
RIV_sigma	Variation in time between visual rail inspections (normal distribution sigma)
RIU_mu	Average time between ultrasonic rail inspections (normal distribution mu)
RIU_sigma	Variation in time between ultrasonic rail inspections (normal distribution sigma)
RMG_mu	Average time to perform rail maintenance (Grind) (normal distribution mu)
RMG_sigma	Variation in time to perform rail maintenance (Grind) (normal distribution sigma)
RMW_mu	Average time to perform rail maintenance (Weld) (normal distribution mu)
RMW_sigma	Variation in time to perform rail maintenance (Weld) (normal distribution sigma)
RMR_mu	Average time to perform rail maintenance (Rerail) (normal distribution mu)
RMR_sigma	Variation in time to perform rail maintenance (Rerail) (normal distribution sigma)
$Shift\_Hrs$	Number of hours per maintenance shift of work

Table 6.4: CPN model functions

Function	Inputs	Outputs	Description
F!TQ	CT, t, q, d, ts	q	Calculates vertical track geometry SD after tamping [μm]. Uses the stochastic model outlined in Section 4.4. As part of this the function outputs the variables to a text file. This data, with the additional data from F!TD, F!SQ, F!SD and F!RD, can then be used to trace the degradation history of each track section and hence find an average condition for the track.
F!TD	tt, mt	d	Calculates vertical track geometry SD degradation rate after tamping [μm/EMGT]. Uses the stochastic model proposed in Section 3.5. External output to text file like F!TQ.
F!SQ	CT, t, q, d	q	Calculates vertical track geometry SD after stoneblowing [μm]. Uses the stochastic model outlined in Section 4.4. External output like F!TQ.
F!SD	tt, ms	d	Calculates vertical track geometry SD degradation rate after stoneblowing [μm/EMGT]. Uses the stochastic model proposed in Section 3.5. External output like F!TQ.
F!RD	tt	d	Calculates vertical track geometry SD degradation rate after renewal [μm/EMGT]. Uses the stochastic model outlined in Section 3.5.
F!RC	ts, t, CT	*Bool*	Checks to see if the poskey passes the conditions required for a renewal instead of maintenance. Example could be after two previous stoneblowing operations and/or when the time between maintenance is less than 6 months.
F!SC	tt	*Bool*	Checks to see if the poskey passes the conditions required for a stoneblow instead of tamping. Example could be after 7 previous tamping operations.
F!U	time	usage	Uses a cumulative usage equation specified as an input to calculate the usage at any model time.
F!ML	M	Maintenance length in Poskeys	Calculates the total length of track maintained in the maintenance operation using the model introduced in Section 4.5.3. $Shift\_Hrs$ also required.
F!RF	CQ, u	r	Calculates if a rail fault occurred and what type it was using the results from Section 5.4.
CQ			Current Vertical Geometry SD [μm] calculated by: $q + d(\text{F!U}(CT) - \text{F!U}(t))$
F!RFM	f	fmt	Calculates the type of maintenance performed, using a simple uniform random number between 0-1 with the stacked probabilities in Table 8.237.

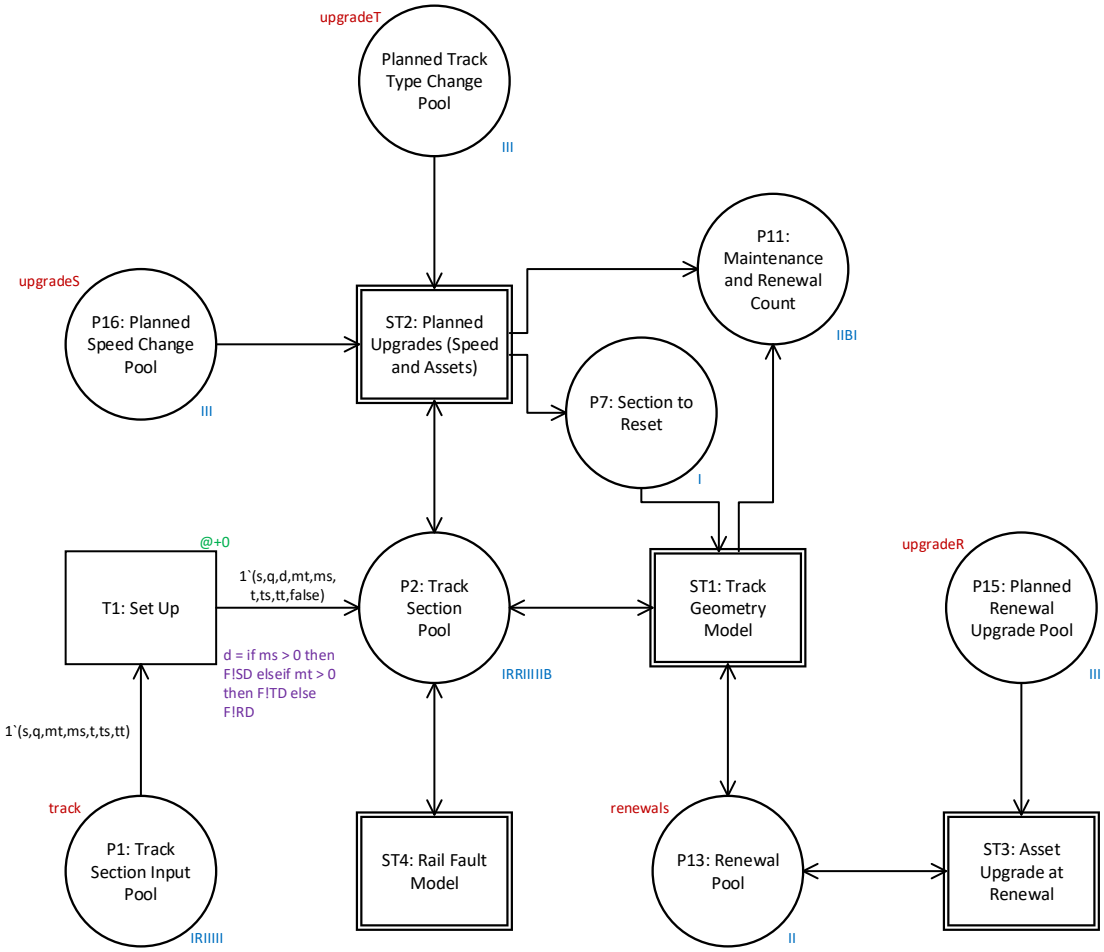


Figure 6.1: Rail Track Asset Management Model - superpage

6.4.2 Track Geometry Inspections

Moving down the hierarchy into the "Track Geometry Model" subpage, shown in Figure 6.3, a port of track section pool, P2, where all the track section information is kept can be seen. With the tokens in P2, they move into the substitution transition Track Geometry Inspection, ST5. The subnet for ST5 is shown in Figure 6.2. Again there is a port of P2 on the subpage. The purpose of the subnet is to model an inspection train being used on the track, and order maintenance if required.

P3 contains a token for every track section with the section number (s) and a boolean of FALSE, denoted by the multiset InspectF. The FALSE tells the model that an inspection is required. These tokens activate T2, which takes all the tokens at once and moves them to P4, adding a variable length of time to the first inspection which follows a uniform (discrete) distribution, sample values between zero and the average length of time between inspections (TI_mu). By using InspectF on the arcs the transition can only activate when all the tokens, with the correct colour sets, denoted in InspectF are in P3. They all activate T2 which fires them all at the same time. This means that they

will all reach P4 at the same time; this is desirable as an inspection train would inspect a line all at once and not individual sections at a time.

With the tokens in P4, T3 being activated. Due to T3 having two input arcs with the variable s, the model ensures these values are the same, i.e. T3 takes track section 3 from P2 and P4. The code segment attached to T3 checks the current vertical geometry of the track. As the degradation rates that are being used are usage-based (as that is the driver of the degradation, not time), a conversion in required. As part of the model an equation for the expected cumulative usage is required (F!U); this can be any formula, for example a quadratic equation. The code segment first calculates the current cumulative usage since the model started by using the current model time (CT) and the cumulative usage at the time of the last maintenance action occurred (t). The difference between the two is the usage that has occurred since the last maintenance action. This is then multiplied by the degradation rate (d) and added to the initial quality (q) to obtain the current vertical geometry SD for each track section.

With the current quality known, it is checked against maintenance thresholds. First, the model checks if the current quality is higher than the threshold for emergency maintenance, where the threshold is dependent on the track speed (ts) (due to safety). If emergency maintenance is not required it checks if normal maintenance should be scheduled, also checking if normal maintenance has already been planned (m). If normal maintenance has already been planned T3 does not add to the maintenance plan again, unless its become urgent. When T3 fires (inspection complete) the token for the track section in P4 is replaced, with a TRUE variable to indicate it has already been inspected. During an inspection (T3), if the track section quality is below the normal threshold for maintenance ($M = 0$) then the transition replaces the track section token from P2 (with no changes). If maintenance is required ($M > 0$), the track section is returned to P2, but with a TRUE variable to inform the model during future inspections that maintenance has already been planned (if inspection intervals are less than the time to undertake maintenance it is possible the model will schedule maintenance more than once).

If normal maintenance is required ($M = 1$) then a token is fired into P6, which is a port of a place on the subpage for the "Track Geometry Model". The section number (s), maintenance histories (mt, ms) and time since last maintenance action (m) are passed into P6 as these are required later to decide on the type of maintenance. An additional colour set of FALSE is also passed into P6 so the model knows that it is normal maintenance and not urgent. If urgent maintenance is required ($M = 3$) then a TRUE is passed instead. If no maintenance is required no token is passed to P6. The arc between T3 and P6 also has a time aspect. When normal maintenance is planned ($M = 1$) then the time the token takes to reach P6 is sampled from a normal distribution with a mean time to undertake maintenance (TM_mu) and a SD (TM_sigma). These parameters are one of the many which can altered by the user depending on real data or expected future values. A distribution was used as there would be variability in how long it takes to undertake maintenance due to issues with resource allocation across the

railway network (only a certain amount of tampers and stoneblowers). If the maintenance is urgent the model uses a time period of one day to undertake the maintenance.

When a track section requires urgent maintenance ($M = 3$) but normal maintenance has already been planned (but not undertaken yet) (m = TRUE), a token is also fired from T3 into P7 (which is a port of a fusion place on the subpage "Track Geometry Model") as part of the inspection. This is used in the "Track Geometry Model" subpage to stop the planned normal maintenance.

Once all the track sections have been inspected P4 will contain a list of tokens with section numbers (s) and TRUE values. T4 will only activate when all the tokens have TRUE values (inspection complete), as InspectT along the arc between them is a multiset containing all the track sections with TRUE values attached. T4 is the time between inspections, with the time sampled from a normal distribution with a mean of TI_mu and a SD of TI_sigma. When T4 fires after the sampled time, all the tokens from P4 are returned but with FALSE values attached (which will activate T3 again). An additional token is fired to P5, which contains the current model time (CT), hence P5 contains a record of the number of inspections that have occurred and when they occurred.

The "Track Geometry Inspection" subnet keeps going during the whole modelling time period, sending track sections that need maintenance to the fusion place P6 which moves it into the "Track Geometry Model" subnet.

6.4.3 Track Geometry Model

The subnet of the "Track Geometry Model" is shown in Figure 6.3 and is designed to perform maintenance.

Once the track sections in P2 have been inspected in the substitution transition ST5 ("Track Geometry Inspection") any track sections needing to be maintained are fired into P6 (which is the port to P6 in the "Track Geometry Inspection" subnet). From the last inspection the sections to be maintained will reach P6 at different times, with sections requiring urgent maintenance reaching P6 faster. When a track sections token reaches P6 the transition T5 will be activated if the same track section (s) has already been sent to P7 and the scheduled maintenance is not urgent (FALSE in colour set). For example, if an inspection occurred which identified a track section as requiring normal maintenance a token for that track section will be sent to P6. If another inspection occurs before the token reaches P6 (due to the delay in the arc between T3 and P6 in the subnet "Track Geometry Inspection"), and a future inspection identifies that urgent maintenance is required (due to further degradation occurring), a second token will also be sent to P6 with a TRUE colour set to signify it is urgent. The urgent token will reach P6 first and cause the section to be maintained. When the normal scheduled maintenance finally reaches P6, maintenance would already have occurred and hence does not need to be maintained again. In this situation though the inspection that decided urgent

maintenance was required would also have sent a token for the track section to P7. As the token in P7 and the normal maintenance required token in P6 will have the same (s) value, T5 is activated and the tokens move into T5, which is a sink transition (no outputs), hence the tokens are removed from the model. As the transition priority of T5 is high it will always activate instead of T6 if its input arc conditions are met.

Track section tokens, which reach P6 and do not activate T5, activate T6, which chooses the type of maintenance to perform. It first checks if the renewal conditions have been met (F!RC), if they have not been met or if the maintenance required is urgent (cannot perform urgent renewals) then the model checks if the conditions for stoneblowing have been met (F!SC). If these are not met tamping is required. If tamping or stoneblowing is required ($M < 3$), a token is fired to P8 with colour sets containing the track section number (s), the type of maintenance (M) and whether the maintenance was urgent or not (u). If a renewal is required ($M = 3$), then a token is fired to P12 with just the section number (s).

When P12 is marked T7 is activated. T7 also has input arcs from the track section pool, P2, and the renewal pool, P13. T7 takes the token for the track section being maintained from P2 (with all the information about the track section in the multiset), and the corresponding track section token from P13. P13 holds a token for every track section being modelled, with the track type ($tt2$) which should be used for the renewal. When T7 fires it replaces the token it took from P13. It also fires a new track section token to P2 with updated information in the multiset. The updated information consists of an initial vertical geometry SD of 800 μm (presumed quality after a renewal), resetting the maintenance histories (mt, ms) to zero, set the maintenance boolean back to FALSE (maintenance not scheduled), updates the track type (tt) to the type marked in the renewal pool, P13 ($tt2$). The time of the last maintenance action (t) is updated with the current time (CT) and a new geometry degradation rate is sampled (F!RD). T7 also fires a token to P11 where a record of maintenance and renewals is kept. The record consists of the section number, the maintenance type (3 as a renewal), whether it was urgent (FALSE as renewals cannot be urgent), and the current time (CT) (time renewal occurred). The fusion place Renewal Pool, P13, is used to allow for the track type to be changed in the model at the first renewal after a certain time period. This is performed in the subnet "Asset Upgrades at Renewal" which is described in Section 6.4.4.

If tamping or stoneblowing is required when T6 fires, P8 is marked. This activates T8, which takes the associated track sections token from P2. When T8 fires a token is sent back to P2 with updated multisets. The time of last maintenance (t) is updated with the current time (CT) and the boolean for maintenance scheduled (m) is changed too FALSE. If the type of maintenance (M) from T6 is one, tamping occurs. As such a new initial geometry is found from F!TQ using the tamping model from Section 4.4, and a new degradation rate from F-TD. The number of previous tamping operations (mt) is also increased by one. If stoneblowing occurs ($M = 2$) then a new initial geometry is found using F!TS, a new degradation rate from F!SD and the number of previous

stoneblowing operations (ms) is increased by one. T8 also fires a token to P9, which consists of the track section number (t) and the type of maintenance (M). This is then used in the "Opportunistic Maintenance" subnet. P10 also receives tokens when T8 fires, these are simple tokens with no colour sets. The number of tokens is dependent on the length of track access for maintenance work, with the number of additional track sections maintained within the access time (input in the model) being calculated using F!ML which follows Section 4.5. T8 also updates the maintenance count, P11.

6.4.4 Asset Upgrades at Renewal

The subnet "Asset Upgrades at Renewal" is designed to allow changes to the track assets to be made when a renewal occurs. The subnet can be seen in Figure 6.4. In the net the fusion place Renewal Pool appears again and starts with a token for every track section with their track types at the start of the model. P15 contains tokens for upgrades, denoted by upgradeR, taking the form of the section number (s) the time to upgrade (Ut) and the track type to upgrade the section to (Utt). At the start of the model running each token activates T10 which fires them to P14. Due to the delay on the arc between T10 and P14, the token does not reach P14 till the time to upgrade (Ut). T9 is then activated, which takes the token for the track section to be upgraded from P13 and updates the track type. The next time the track section is renewed, it will be renewed with the new track asset type specified in Utt.

6.4.5 Opportunistic Maintenance

When performing maintenance on a railway track it is possible to maintain more than one track section, hence the track sections around the one requiring maintenance are also maintained. This can have negative impacts, as seen in Section 4.3.2, where performing maintenance earlier increasing the chance of doing more harm than good. Due to this it is important that the correct decisions are made regarding opportunistic maintenance. Within the proposed model the "Opportunistic Maintenance" sub net, shown in Figure 6.5, tries to mimic the decisions of a track engineer picking the cluster of track sections around the one being maintained that are in the poorest condition. This includes avoiding opportunistic maintenance being performed too early.

The subnet connects to the "Track Geometry Model" via the fusion places P2, P7, P9 and P10. When maintenance is performed on a track section a copy of the section number and maintenance type are fired into P9 and the number of additional sections that can be opportunistically maintained based on the working window are fired into P10 (one token for each section). If no additional sections can be maintained, P10 is empty of tokens and hence T28 can be activated. When T28 fires it removes the token from P9 and no opportunistic maintenance occurs.

If P10 has tokens T20 is activated. When it fires, it finds the adjoining track sections.

If section number, s, is 7, $smin$ is 3 and $smax$ is 123 then two tokens will be fired into P26, one was $s = 8$ and one where $s = 6$. If $smin = 7 = s$ then there is no track section in one direction (end of line/track being modelled), hence only one token is fired to P26 ($s = 8$). An additional boolean colour set is added to the tokens which records if the section was up or down the line (up). T21 then fires, this calculates the current vertical geometry of the track, using the track section information stored in the multiset in P2. T21 then fires the token to P27 if the current quality is above the opportunistic thresholds (which are related to the track speeds). If a second token is in P26 then this also fired into P27 if its quality is above the threshold.

If there are two tokens in P27, T22 is activated. This checks that the track sections require the same type of maintenance. This stops tamping occurring if stoneblowing already has. If both are applicable to the type of opportunistic maintenance (M), then the T22 picks the one in the worst condition and fires it to P28, while firing the better quality track section back to P27. If only one token made it to P27 (quality was below the threshold for one of the tokens or end of line), T23 is activated instead of T22. This again checks that the type of maintenance (M) is applicable to the track section, if it was it is fired into P28. The token in P28 activates T25. This maintains the track section similar to T8 in the subnet "Track Geometry Model", updating the multiset attached to the token with the same section number in P2. If that section already had maintenance scheduled a token is fired into P7 to stop the scheduled maintenance. T25 takes a token from P10. P10 holds a token for each track section can be maintained within the maintenance access period, for example three tack sections. T25 removing a token represents using up one of the possible sections which could be maintained in the maintenance action. A token is also fired to P29, so a new starting place for checking adjacent is known. When T25 fires, if it removes the last token from P10, the sink transitions called reset (T26 and T27) removes the remaining tokens from the sub net and opportunistic maintenance finishes.

The token in P29 activates T24 if there are still tokens remaining in P10. T24 identifies for the next track section to be checked for maintenance, using the token information. This is always the section next to the last one maintained by T25. For example, if the value of the colour set, up, is TRUE then it is known that the last section had a lower section number (s) than the original which means the next section to be checked would be one lower again (as long as it is not the end of the track being analysed). If the original section number to be maintained was 27, then after the first opportunistic maintenance check, section 28 was maintained, T24 would input section number 29 into P26, to mark this should be checked next. T24 fires a token into P26 and the cycle continues until there are no tokens left in P10 or none of the adjacent track sections are applicable for maintenance (wrong sort of maintenance or quality level is below the threshold). When this happens T28 activates and removes all the tokens from P10 leaving no tokens on the sub net.

6.4.6 Planned Upgrades (Speed and Assets)

To allow for the track speed to be changed or to undertake renewals on specific dates (not waiting till a renewal naturally occurs as Section 6.4.4) the subnet "Planned Upgrades (Speed and Assets)" was designed and shown in Figure 6.6. As the model is designed to predict long term (>20 years) of railway track degradation and maintenance it is likely that aspects such as speed would change over time. As the track speed is a required model input, the model needs to know if it will change and when. This allows to assess the impact of changing the future track speed. Planned upgrades is also important as it may not be desirable to wait until a renewal is required to change the track assets. For example if a track speed increase is desirable in 10 years, which will require improved track assets, the model will need to know to upgrade the track assets at this time.

Speed Upgrade

P16 has an initial marking of upgradeS which consists of a selection of tokens for each speed change that is desired throughout the model. Each token equates to one speed change and must include the section number (s), the speed to change to (Uts) as well as the time to the speed is changed (Ut). The tokens in P16 activate T11. T11 fires the tokens to P17, with the tokens being delayed by Ut. When the tokens reach P17 they activate T12 which takes the track section token to be changed from P2 and returns it with a new track speed (Uts). Due to the track speed changing, the rate of degradation would also change. This is accomplished by T12 first calculating the current vertical geometry SD (CQ) and also the type of the last maintenance (M). The token returned from T12 to P2 also updates the initial quality (q) with the current quality (CQ), updates the last time of maintenance (t) to the current time (CT). It also recalculates the degradation rate (d), sampling from the Weibull distributions corresponding to the track type (tt) and the previous type of maintenance (M), where $M = 2$ for tamping, 1 for stoneblowing and 3 for renewal.

Track Type Upgrade

Similar to the speed upgrade, P18 starts with tokens of which sections to upgrade (s), what to upgrade them to (Utt) and after how long (Ut), specified by upgradeT. The tokens in P18 activate T14 which fires them to P19 after delay Ut. T13 is then activated, which takes the corresponding track section token from P2. When T13 fires, it sends a token back to P2, where it has been renewed with an updated track type of Utt. As part of the forced renewal, the maintenance histories are returned to zero, a new degradation rate is found, the initial geometry is set to 800 μm and the time of last maintenance is updated to the current time. Additionally, a token is sent to P11 to count the renewal. A token is also sent to P7 if maintenance is already planned on the track section to stop the maintenance occurring.

6.4.7 Rail Fault Model

In addition to predicting track geometry the model has also been designed to take into account rail faults, in particular the link between rail faults and vertical geometry seen in Section 5.3.2. This is accomplished in the subnet "Rail Fault Model", which is shown in Figure 6.7. A port of P2 appears again in this subnet, linking the geometry and rail fault models together. P20 has an initial marking (inspectT), which adds a token for every track section with an associated TRUE variable. P20 activates T16, which takes all the tokens in P20; then after 7 time steps, these are returned but with FALSE values attached instead. This activates T20 which checks each section type similar to the "Track Geometry Inspections" subnet described in Section 6.4.2. As part of the check the track sections current quality (geometry) is found, which is used in the rail fault model introduced in Section 5.4 to determine if a fault occurred and what fault type it was (F!RF). The usage since the last fault check is also calculated and used in the rail fault model to alter the rates (as the rates are for 1 EMGT). As each section is one poskey there is no need to make an adjustment for track section length. If a fault occurred ($f > 0$) a token is fired into P21. The token includes the track section (s), the current time (CT) and the fault type (f) and a boolean value to specify the inspection type that can find the fault, based on the analysis in Section 5.2.1 and the f values in Table 8.237. For example, if a squat occurs ($f = 1$) a token is fired to P21 with an inspection type, it, of FALSE which means only ultrasonic inspections can find the fault. This is based on the analysis in 5.2.1 which showed 84.4% of squats being found using ultrasonic inspection. If the fault type can be found by either inspection method (ultrasonic or visual) i.e. weld faults, then the token is duplicated, one with a TRUE and one with a FALSE. P21 holds rail faults that have occurred but have not been found yet.

P22 holds tokens for rail inspections. The initial marking adds two boolean tokens, one TRUE and one FALSE. The TRUE token is for visual inspections and the FALSE is for ultrasonic. Initially T18 is activated by P22. When T18 fires there is a delay on the output arc back to P22. This delay is the time between inspections and are normally distributed, with separate inspection regimes possible for visual and ultrasonic inspections. When the tokens reach P22, if a fault has occurred and a token is in P21 which can be seen by the inspection type in P22 (same inspection type, it), T17 will activate instead of T18 due to the higher priority of T17. This removes the fault from P21 and moves it to P24 (fault identified). If the fault type was one that could be found by either inspection type a token will be left in P21 of the inspection type that did not occur. If this is the case T20 will activate and remove the extra token from P21 (the fault had already been found). T19 is then activated. T19 finds the type of maintenance used to fix the fault using the fault type, f, and the probabilities in Table 8.237. If a grind is required it will return 1, weld = 2 and rerail = 3. T19 then fires the token to P25. The token is delayed by the arc delay which calculates the time to perform the maintenance based on the type of maintenance scheduled in T19. P25 collects a count

of all the rail faults as well as their fault type (f), inspection type (it), track section it occurred on (s), time fault occurred (ft), time fault was fixed (CT) and maintenance type (fmt)

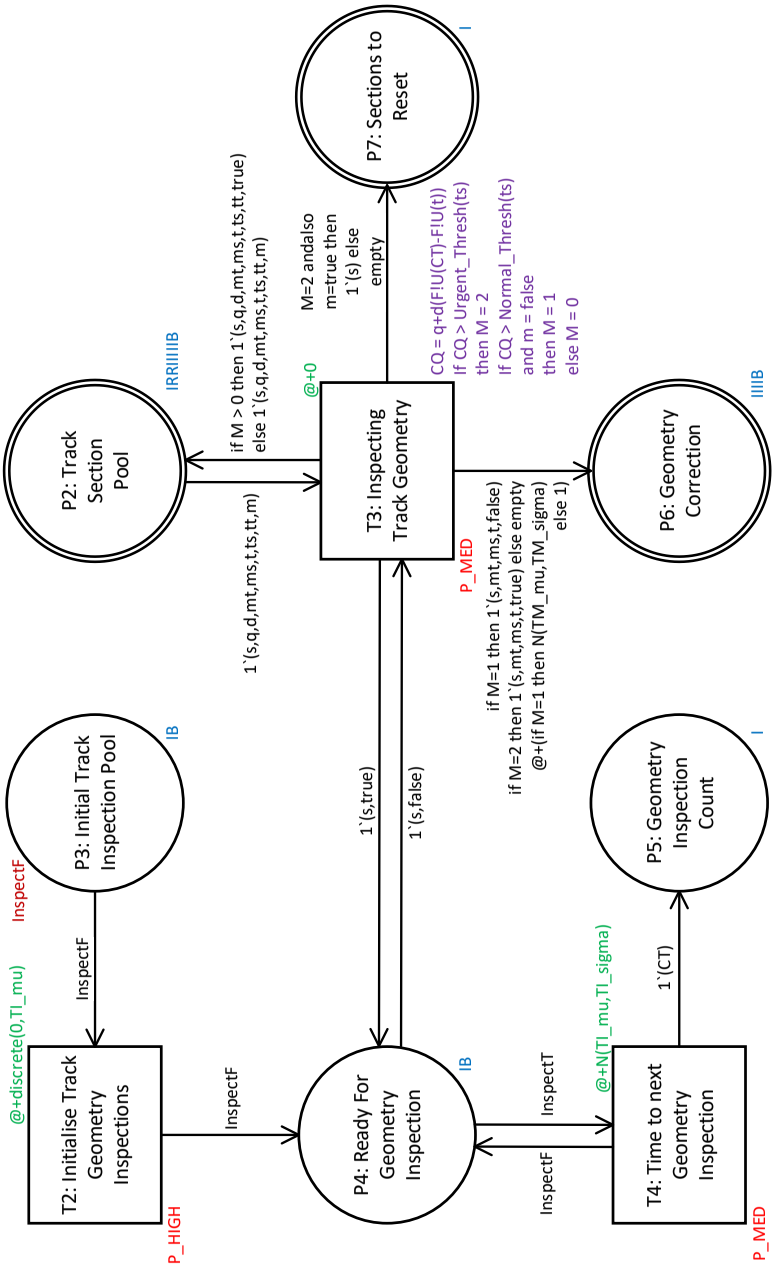


Figure 6.2: Track Geometry Inspections - subpage

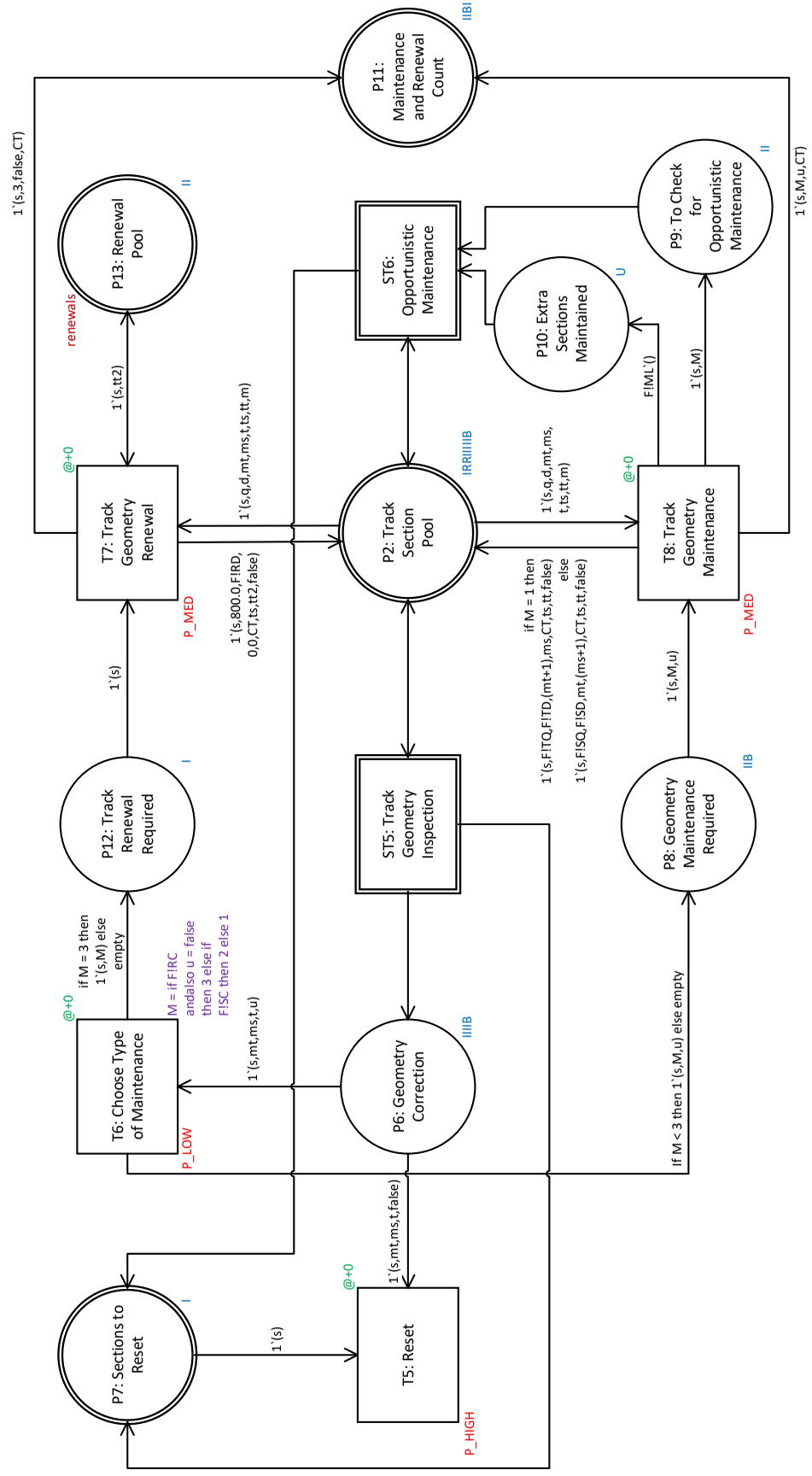


Figure 6.3: Track Geometry Model - subpage

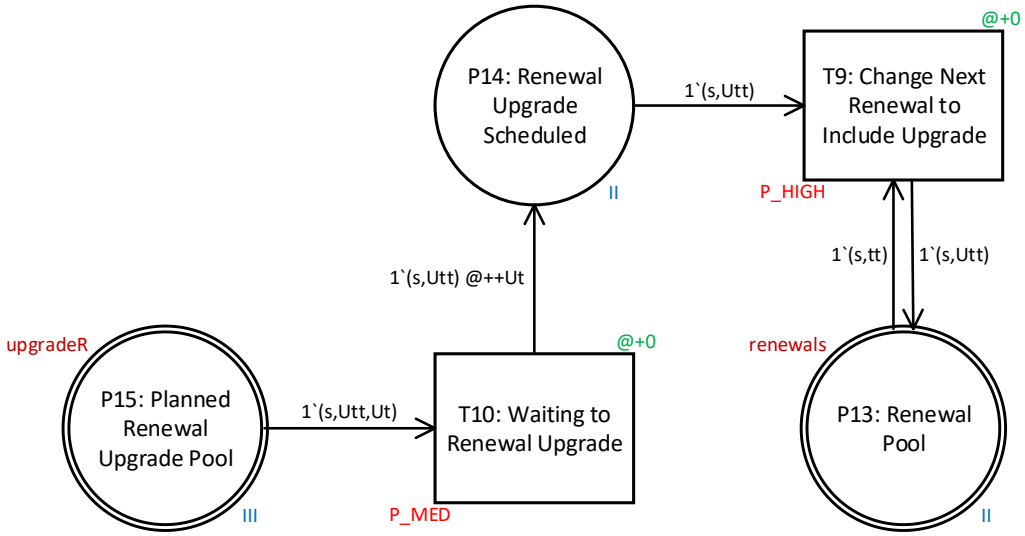


Figure 6.4: Asset Upgrades at Renewal - subpage

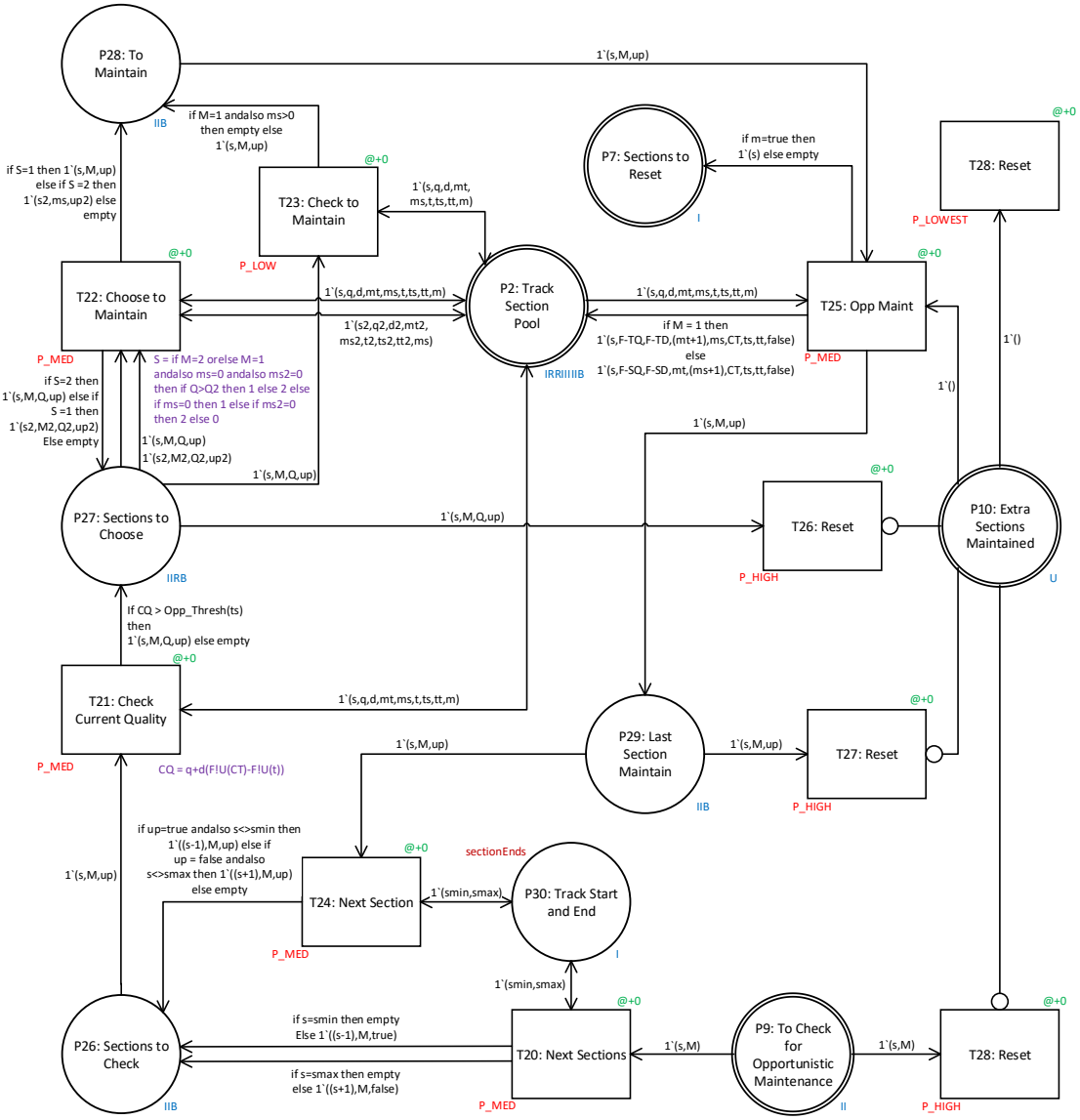


Figure 6.5: Opportunistic Maintenance - subpage

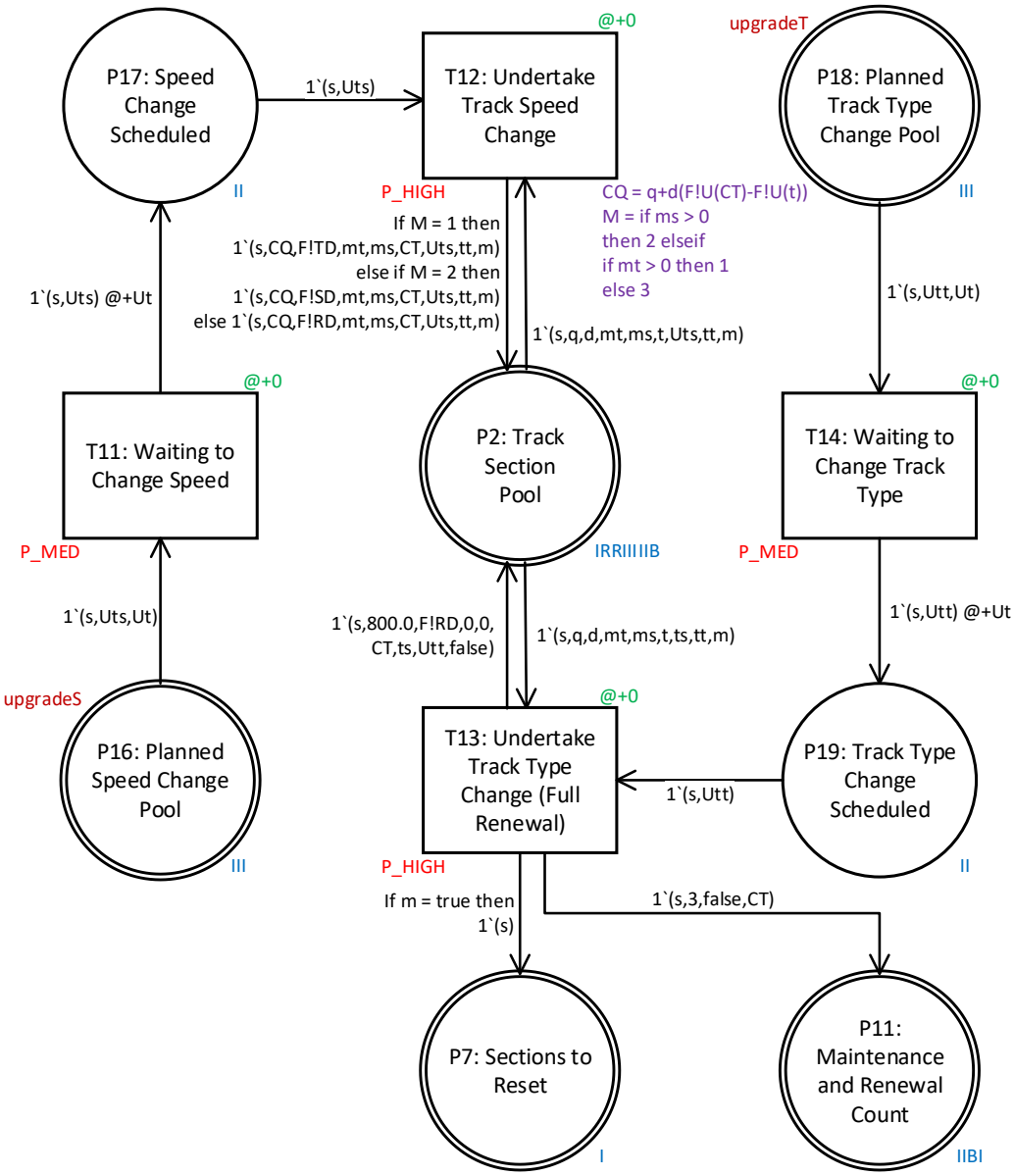


Figure 6.6: Planned Upgrades (Speed and Assets) - subpage

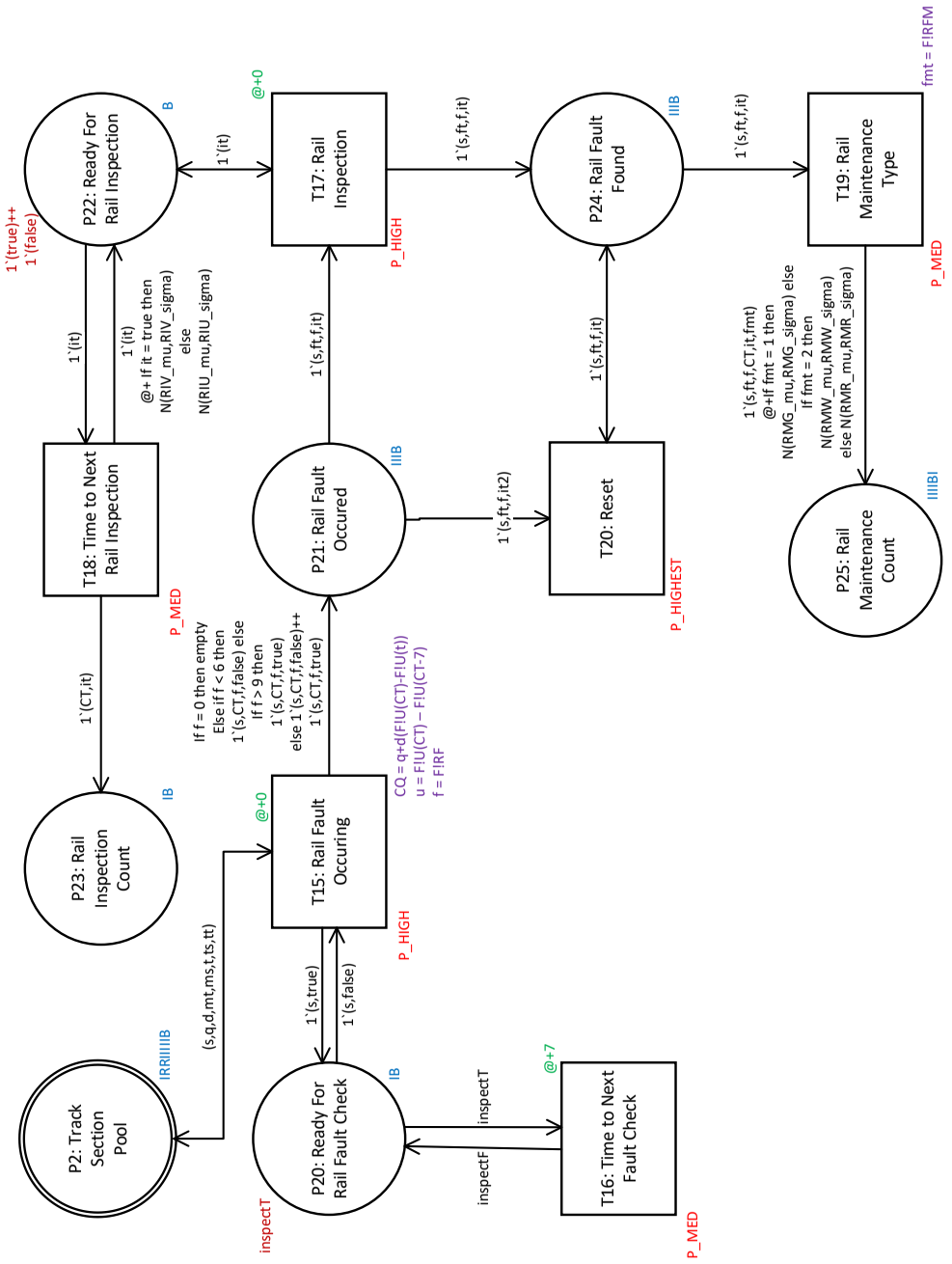


Figure 6.7: Rail Fault Model - subpage

6.4.8 Possible Outputs

The model can output a large amount of data stored in the counting places (P5, P11, P25). These include:

1. Quantity of track geometry and rail (visual and ultrasonic) inspections and when they occurred, over a period of time.
2. Quantity of tamping, stoneblowing and track renewals, when they occurred and on what track sections, over a period of time.
3. Number of rail faults that occurred, their types, what inspection method found them and what maintenance method was used
4. Time from a rail fault occurring to it being found and maintained, involves both inspection regimes and time to perform maintenance.
5. Amount of maintenance actions that are urgently required (speed restrictions applied) as well as the amount of time speed restrictions are in place.
6. Track quality conditions using the outputs from F!TQ (track geometry after tamping), F!TD (track geometry degradation rates after tamping), F!SQ (track geometry after stoneblowing), F!SD (track geometry degradation rates after stoneblowing) and F!RD (track geometry after renewal), can be obtained using additional software such as Matlab.

These outputs in addition to the large amount of possible inputs the model user can change (maintenance thresholds, access time periods,time to perform different maintenance, upgrades etc.), allows many aspects to be investigated including the impact of:

1. Usage change. Amount of trains traversing the track can change in the modelled period as the degradation is related to usage not time. This allows the impact of traffic changes on maintenance costs to be modelled and understood.
2. Change in the working window, and how this impacts the number of track sections that can be maintained in a single maintenance visit. Shorter working windows will reduce the quantity of opportunistic maintenance performed leading to more maintenance visits.
3. Changes to inspection intervals and the speed that maintenance is undertaken. If too long is spent undertaking maintenance or long times between inspections, more urgent maintenance and hence speed restrictions will occur.
4. Changes to maintenance boundaries. Higher geometry boundaries will lead to less geometry maintenance but a higher number of rail faults and lower average track quality (and hence lower passenger comfort and safety).
5. Changing the track speed or asset types. This allows the impact of upgrades, and whether they are good financial value, to be modelled. If an upgrade costs more but is shown by the model to have minimal impact of maintenance saving, the investment may not be worth it. Track speed is also important to understand, as increasing the track speed, with reduced maintenance thresholds and increased degradation, can lead to an increased amount of maintenance actions and hence

more money required.

6. Changes to the rules of what maintenance type to choose. When should stoneblowing start to be used, what conditions require a renewal over maintenance.

6.5 Example of Multi-Section Track Modelling

6.5.1 Introduction

To test the model it was decided to run an example line of track, using a time step of one day and a time period of 60 years.

6.5.2 Line Selection and Variables

A rail track of 50 miles in length (400 poskeys) was tested within the model over a 60 year period. The chosen line section included a selection of different speeds but was primarily between 80-125MPH with sleepers choices (Track Type, *tt*, in the model) matched to which track speeds they were most likely to occur on. For example, large concrete sleepers for fast track and small steel and timbers for the slower track sections. This made sure the track section to be tested was realistic, but also ensured that the related distributions used were the ones based on the most data (higher confidence). The included track sections were given starting conditions and maintenance histories which varied per section instead of start from new. The global variables used are recorded in Table 6.5. For the example sensible values were chosen by assessing past data. As no data was available for, the times to perform maintenance once a threshold is past, rail inspection periods and the time to action rail maintenance, these values were picked logically. Where urgent maintenance is performed much faster and more complex maintenance such as Rerail, takes longer to plan and action than grinding. The model rules are recorded in Table 6.6, where the usage has been given as a cumulative function. This allows the model to calculate the expected cumulative tonnage (EMGT) at any point. The equation chosen increases the traffic slightly each year, with the first year having 10 EMGT of traffic and year 60 will have 16.3 EMGT.

Table 6.5: Model Variables

Variable	Value [days]
TI_mu	100
TI_sigma	10
TM_mu	60
TM_sigma	10
RIV_mu	50
RIV_sigma	5
RIU_mu	200
RIU_sigma	40
RMG_mu	14
RMG_sigma	2
RMW_mu	28
RMW_sigma	4
RMR_mu	60
RMR_sigma	5
$Shift\_Hrs$	7

Table 6.6: Model Rules

Rule	Boundaries
Stoneblowing	Occurs after seven previous tamps
Renewal	Occurs after two stoneblowing operations
Usage	$10 * (T/365)^{1.1}$ where T is in days

6.5.3 Results

Impact of Maintenance Thresholds

To assess the impact of changing the maintenance thresholds the model was run using the thresholds in Table 6.7. Model A contains lower thresholds, hence it would be expected that the average condition of the track will be better compared to B at the expense of further maintenance actions. The thresholds for Model A were based on NR's quality bands, with opportunistic maintenance occurring when a track condition degraded to very good, normal maintenance when it became good and urgent maintenance if the quality became poor. Model B thresholds were set slightly higher to see the impact of increasing the maintenance thresholds.

The model, as with all stochastic models, has completed running when convergence is reached on the outputs. Convergence occurs when the running average becomes steady (stops changing drastically). For the model proposed this occurs in about 1000 simulations as seen in Figure 6.8a, each simulation took 12 minuets to calculate, with the required 1000 simulations taking over a week to run. The high calculation time is due to limitations within the computer program used (CPN-Tools). The average number of tamping operations per year is given in Figure 6.8b. The figure shows that there is a small steady increase in the number of tamping operations required a year for both Model A and B. This is due to the usage not being constant throughout the 60 years, instead following the usage equation in Table 6.6, which gives a steady increase of traffic on the line.

Comparing the results given in Table 6.8, it can be seen that Model A with the lower maintenance boundaries has a higher amount of tamping operations, stoneblowing operations and renewals over the 60 year period. Remembering that renewals count the number of 220 yd sections renewed as the model only undertakes one section renewal at a time (no opportunistic renewals). Tamping and stoneblowing can occur over multiple sections at a time via opportunistic maintenance and modelling the working window, but each operation is only counted once (not the number of sections) as maintenance costs tend to be per whole shift. By associating costs to maintenance procedures it is possible to understand which maintenance thresholds lead to lower running costs of the railway track. Some fabricated costs have been outlined in Table 6.9. Multiplying the costs by the maintenance quantities for Model A and B in Table 6.8, gives Table 6.10. In this table it can be seen that Model A, which had lower maintenance thresholds has higher track geometry maintenance costs. This is due to more maintenance being required than Model B to keep a higher quality level. Alternatively, the cost of rail maintenance costs in lower in Model A (due to less rail maintenance required). This is due to the link proved in Section 5.3.2, where worse track geometry leads to more rail faults. With the average track geometry quality higher in Model A, due to the lower maintenance thresholds, fewer rail faults occur. Overall, using the fabricated maintenance costs in Table 6.9, Model B was shown to have lower costs. This is due to the saving in track geometry maintenance compared to Model A, being larger than the increase in rail maintenance costs. If rail maintenance costs were much higher, Model A could become cheaper. The track condition is not the same between Models A and B and also needs to be taken into account when making asset management decisions. As the higher costs of Model B maybe worse the improvement in average quality, and the advantages this brings like improved passenger comfort safety.

The average geometry condition of each model can be seen in Figure 6.9. Track quality is related to the required level of service, for example, a vertical geometry SD of 5 mm on a slow track is safe but would not be on a fast track. Due to this the SD values have been changed into quality bands. This is accomplished using Table 6.11 which aligns to NR's current quality band thresholds. Over the model simulations in each year for each

track section the time spent in each band is recorded and the data collated across all simulations and track sections to obtain an average condition band over the line for each year. The average percent of each band is multiplied by a value between 0-1 to achieve one number between 0-1 for the geometry quality, *GeoQ*. This is given by:

$$GeoQ = \%Excelent * 1.0 + \%V.Good * 0.75 + \%Good * 0.5 + \%Poor * 0.25 + \%Red * 0 \tag{6.1}$$

If a track section spent its entire time in excellent condition its condition score would by 1. Whereas if it spent half the time in good and half in poor the condition score would be 0.375.

The condition results found at the bottom of Table 6.8 show that the average condition of Model A is slightly above very good, whereas Model B is in between good and very good. Looking at Figure 6.9 it can be seen that the condition of the track in both Model A and B is decreasing with time, which is most likely to do with the growing level of traffic being modelled. When the track condition falls into Super Red (the lowest quality band) a speed restriction is imposed on the line to maintain safety. Model A spent 0.01% of time (2.2 days) in Super Red whereas Model A was 0.03% of the time (6.6 days). Any financial penalty of having to impose speed restrictions should be included. With the condition, time with speed restrictions and costs predicted a decision can be made on the preferable scenario.

Table 6.7: Track Geometry Maintenance Thresholds

Speed Band (MPH)	Model A Threshold [um]			Model B Threshold [um]		
	Oppor-tunistic	Normal	Urgent	Oppor-tunistic	Normal	Urgent
5-20	5200	7400	8300	6000	8000	8300
25-30	4300	6100	7000	5200	7400	7000
35-40	4100	5800	6700	4300	6100	6700
45-50	3800	5400	6300	4100	5800	6300
55-60	3500	5000	5900	3800	5400	5900
65-70	3000	4300	5400	3500	5000	5400
75-80	2700	3800	4800	3000	4300	4800
85-95	2200	3200	4000	2700	3800	4000
100-110	1900	2700	3400	2200	3200	3400
115-125	1700	2400	3000	1900	2700	3000

Table 6.8: Maintenance Threshold Results Comparison

Activity	Number in Model A	Number in Model B
Tamping Operations	2603	1943
Stoneblowing Operations	302	212
Sections Renewed (220yd)	401	290
Rail Faults (Rerail)	2027	2317
Rail Faults (Weld)	2259	2582
Rail Faults (Grind and Other)	1036	1184
Average Condition ($GeoQ$)	0.79	0.68

Table 6.9: Example Maintenance Activity Costs

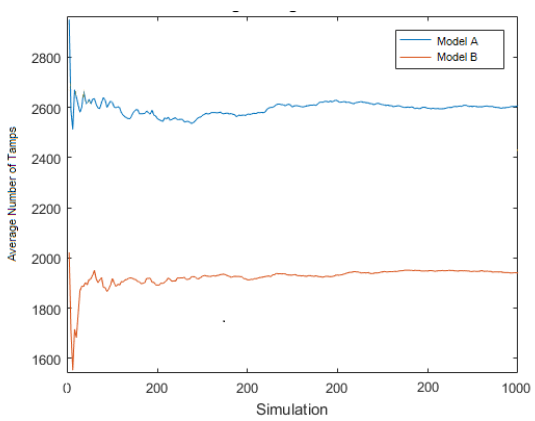
Activity	Cost [£]
Tamping Operation	5,000
Stoneblowing Operation	8,000
Renewal	40,000
Rerail	2,000
Weld Rail	500
Grind Rail	400

Table 6.10: Maintenance Threshold Results, Maintenance Costs Comparison

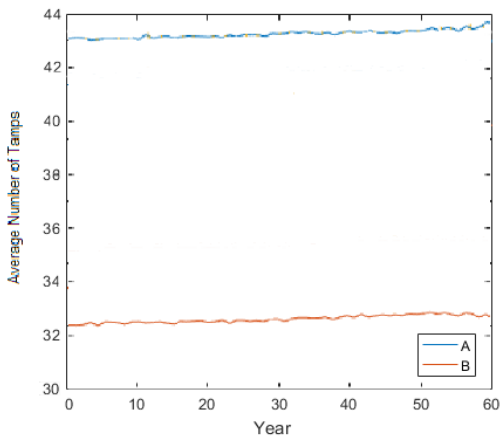
Activity	Cost in Model A [£]	Cost in Model B [£]
Tamping Operations	13,015,000	9,715,000
Stoneblowing Operations	2,416,000	1,696,000
Sections Renewed (220yd)	12,030,000	8,700,000
Track Geometry Maintenance Subtotal	27,461,000	20,111,000
Rail Faults (Rerail)	4,054,000	4,634,000
Rail Faults (Weld)	1,129,500	1,291,000
Rail Faults (Grind and Other)	414,400	473,600
Rail Fault Maitnenance Subtotal	5,597,900	6,398,600
Total	33,058,900	26,509,600

Table 6.11: Track Geometry Condition Thresholds

Speed Band (MPH)	Threshold [um]				
	Excellent	Very Good	Good	Poor	Super-Red (Speed Restriction)
5-20	5200	7400	8300	9900	>9900
25-30	4300	6100	7000	7700	>7700
35-40	4100	5800	6700	7200	>7200
45-50	3800	5400	6300	6700	>6700
55-60	3500	5000	5900	6300	>6300
65-70	3000	4300	5400	6000	>6000
75-80	2700	3800	4800	5700	>5700
85-95	2200	3200	4000	5300	>5300
100-110	1900	2700	3400	5000	>5000
115-125	1700	2400	3000	4700	>4700

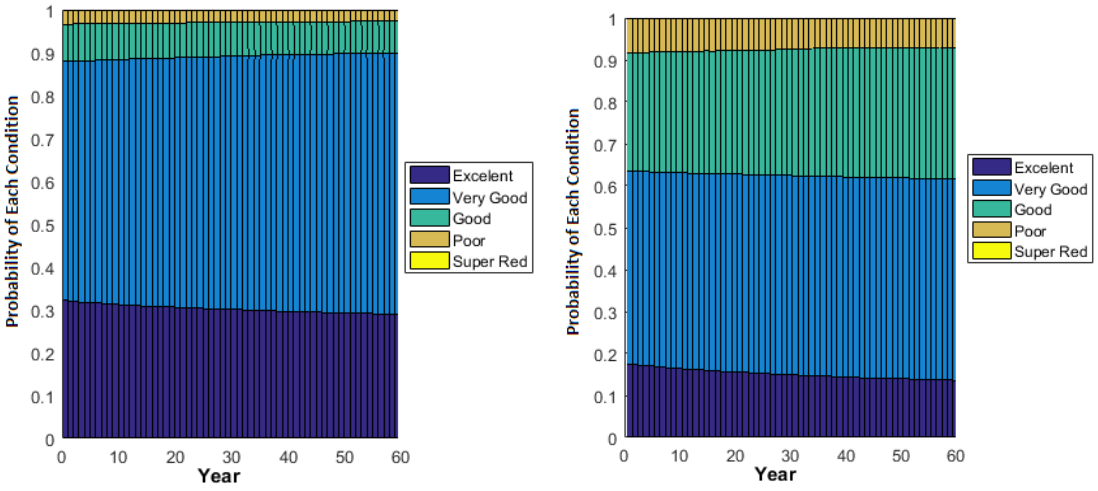


(a) Model Convergence



(b) Yearly Averages

Figure 6.8: Number of Tamping Operations (Model A and B)



(a) Model A (b) Model B

Figure 6.9: Yearly Condition Results

Impact of Increased Time to Perform Maintenance

The developed model has many inputs which can be altered to assess the impact on the track. When vertical geometry maintenance, tamping, stoneblowing and renewals, are identified as required, the maintenance does not happen instantly. There maybe other higher priority maintenance, or a lack of resources, which delays it. In the time it takes to perform the maintenance, the track continues to degrade. Which may lead to urgent maintenance being required (possibly at a higher cost) or even line speed restrictions. The previous Model A was used as the baseline. This decision reduced the number of times the model needed to be run, which was required due to the long calculation time. Model A is compared to Model C. Where Model C has all the same variables and rules as Model A (Tables 6.7, 6.6 and 6.5), but the variable TM_mu has been increased from 60 days to 120 days.

Results for Model C were calculated and compared to Model A, and can be seen in Table 6.12. For this comparison tamping and stoneblowing maintenance operations have been split between normal and urgent. Urgent maintenance occurs when an inspection finds a track section within the urgent threshold of Table 6.7. These are track sections which are in poor condition and close to requiring speed restrictions for safety, so maintenance is prioritised. The results show that, increasing the time to perform geometry maintenance led too fewer operations and renewals occurring. Despite there being less tamping and stoneblowing operations in Model C, there are more urgent operations required. Due to the longer time from scheduling maintenance before it happens in Model C, more degradation occurs before the section is maintained. If a section degrades quickly, and an inspection happens before maintenance is complete, urgent maintenance is scheduled instead. Urgent maintenance is likely to be more costly, due to short time frames, as obtaining the required resources is more difficult and hence normally costly. Also, urgent

maintenance is often completed instead of another normal maintenance action, delaying other maintenance which is required. In Table 6.13 it was presumed urgent maintenance would be 50% more expensive. Note in this table Model C is still cheaper overall for track geometry maintenance due to the cost saving from the reduction in maintenance actions being greater than the cost of additional urgent maintenance actions.

The additional time to maintenance has reduced the overall amount of geometry maintenance. There are two probable reasons. Due to the extra time to perform maintenance, the average time between maintenance operations has increased, leading to less possible in the 60 years modelled. Another reason will be due to opportunistic maintenance. The extra time before maintenance increases the chance of track sections around the targetted one degrading to a point where opportunistic maintenance is undertaken. This leads to more track sections being maintained at a time. Reducing the over number of maintenance actions, but not necessarily reducing the average number of tamps per track section.

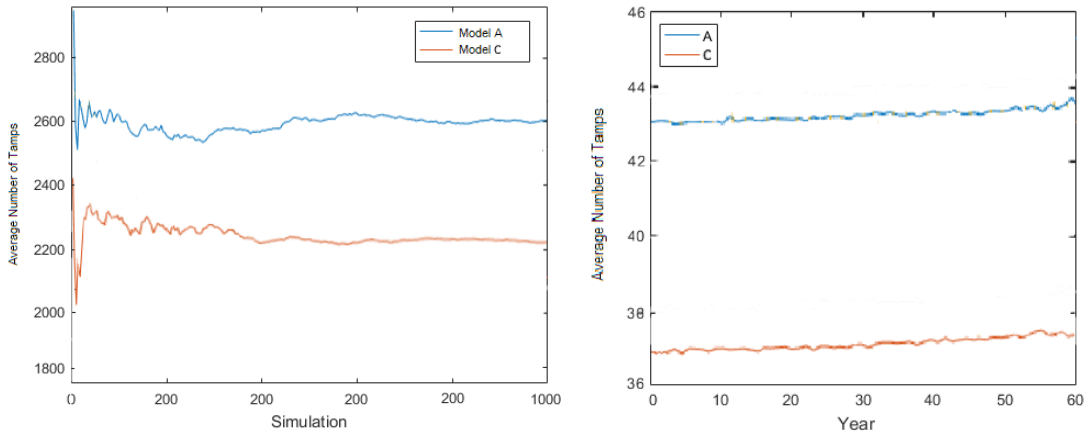
The conditions in Table 6.12 and in Figure 6.11 show that the average condition of the track for Model C was lower than Model A. This is due to the extra time in good and poor track conditions while waiting for maintenance. Due to the increasing traffic on the track, which is governed by the commutative usage formula in Table 6.6, the average number of maintenance actions per year slowly increases as seen in Figure 6.10b. The lower average condition of the track in Model C led to an increase in the number of rail faults and hence maintenance actions, as seen in Table 6.12.

Table 6.12: Time to Perform Maintenance Results Comparison

Activity	Number in Model A	Number in Model C
Tamping Operations	2542	2112
Urgent Tamping Operations	61	109
Stoneblowing Operations	290	243
Urgent Stoneblowing Operations	12	13
Sections Renewed (220yd)	401	327
Rail Faults (Rerail)	2027	2170
Rail Faults (Weld)	2259	2351
Rail Faults (Grind and Other)	1036	1075
Average Condition ($GeoQ$)	0.79	0.73

Table 6.13: Time to Perform Maintenance Results, Maintenance Costs Comparison

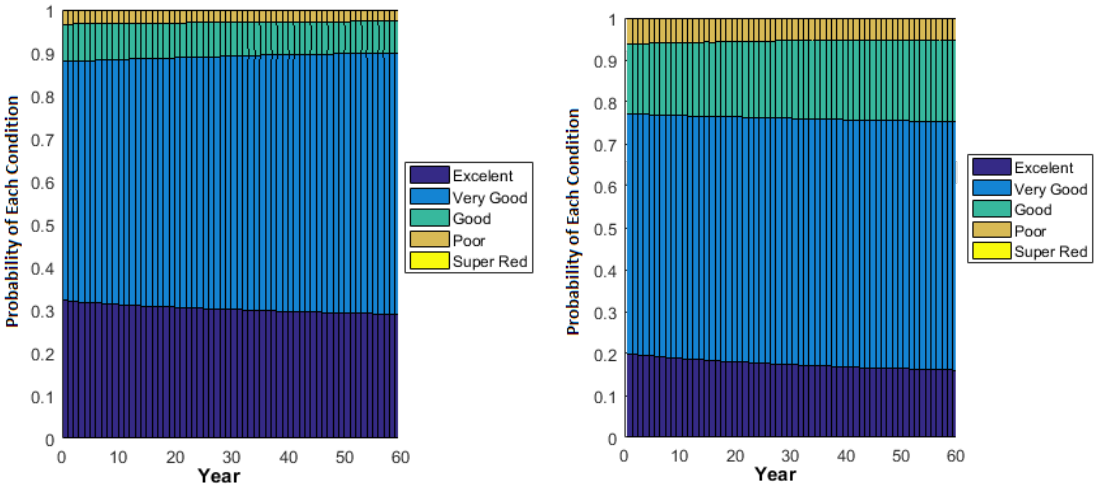
Activity	Cost in Model A [£]	Cost in Model C [£]
Tamping Operations	12,710,000	10,560,000
Urgent Tamping Operations (+50%)	457,500	817,500
Stoneblowing Operations	2,320,000	1,944,000
Urgent Stoneblowing Operations (+50%)	144,000	156,000
Sections Renewed (220yd)	12,030,000	9,810,000
Track Geometry Maintenance Subtotal	27,661,500	23,287,500
Rail Faults (Rerail)	4,054,000	4,340,000
Rail Faults (Weld)	1,129,500	1,175,500
Rail Faults (Grind and Other)	414,400	430,000
Rail Fault Maitnenance Subtotal	5,597,900	5,945,500
Total	33,259,400	29,233,000



(a) Model Convergence

(b) Yearly Averages

Figure 6.10: Number of Tamping Operations



(a) Model A (b) Model C

Figure 6.11: Yearly Condition Results

Impact of a Decreased Maintenance Window Length

The cost of maintenance is more related to the number of shifts of work than it is the number of track sections maintained. Due to this the length of a maintenance shift, and hence how much work can be completed in a shift, are very important to the cost of maintaining a railway track. This is also an aspect previous models in literature have not considered. In the previous Models A-C, the length in hours of a maintenance shift was 7 hours. Using the maintenance output model introduced in Section 4.5.3, on Page 178, the median number of tamps and stoneblows in an 7 hour shift are 7.2 and 7.0 poskeys, respectively. To assess the impact of the maintenance window length, Model D was calculated. Model D was given all the same parameters as Model A except for the shift length which was decreased to 4 hours. Hence, in Model D, the median number of tamps and stoneblows per shift are 4.1 and 4.0, respectively.

Reducing the working window was shown to increase the number of tamping operations, but decrease the number of stoneblows and renewals. These results are seen in Table 6.14. Due to the reduced working window in Model D, each tamping operation maintains fewer track sections (poskeys). To perform the same amount of tamping operations on each track section, Model D requires more work shifts of tamping. The reduction in stoneblowing and renewals occurs due to opportunistic maintenance. As less track sections are being maintained at a time, fewer track sections are maintained within the opportunistic thresholds in Table 6.7. This reduces the average quality of the track slightly (0.76 vs 0.79) as on average track is being left longer until it is maintained, despite the same maintenance thresholds. The model rules were set to switch to stoneblowing after seven previous tamps, then renew after two stoneblows. As, on average, tamping is being performed later in Model D, this means it takes longer (time and usage) before the maintenance method is changed to stoneblowing. This Resulted in less stoneblows

being required within the 60 year period modelled for model D. As renewals occur after two stoneblows, Model D also sees a reduction in renewals.

Due to fewer renewals and less opportunistic maintenance the average track quality in Model D is lower than A, as seen in Figure 6.13. This leads to the slight increase in rail maintenance seen in Table 6.14. The costs in Table 6.15 show that, for the example monetary values in Table 6.9, despite fewer renewals and stoneblowing operations in Model D, the overall costs are still higher. Model D also had lower average track geometry. So reducing the working window not only increases maintenance costs and occurrences, but also leads to a reduction in the average track quality.

Table 6.14: Maintenance Window Length Results Comparison

Activity	Number in Model A	Number in Model D
Tamping Operations	2603	3214
Stoneblowing Operations	302	260
Sections Renewed (220yd)	401	351
Rail Faults (Rerail)	2027	2080
Rail Faults (Weld)	2259	2411
Rail Faults (Grind and Other)	1036	1099
Average Condition ($GeoQ$)	0.79	0.76

Table 6.15: Maintenance Window Length Results, Maintenance Costs Comparison

Activity	Cost in Model A [£]	Cost in Model D [£]
Tamping Operations	13,015,000	16,070,000
Stoneblowing Operations	2,416,000	2,080,000
Sections Renewed (220yd)	12,030,000	10,530,000
Track Geometry Maintenance Subtotal	27,461,000	28,680,000
Rail Faults (Rerail)	4,054,000	4,160,000
Rail Faults (Weld)	1,129,500	1,205,500
Rail Faults (Grind and Other)	414,400	439,600
Rail Fault Maitnenance Subtotal	5,597,900	5,805,100
Total	33,058,900	34,485,100

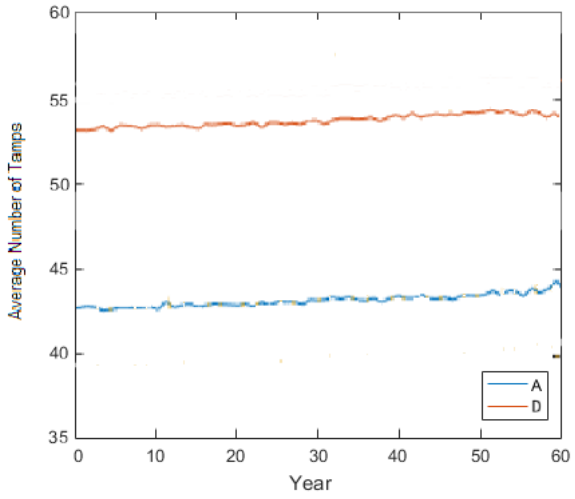
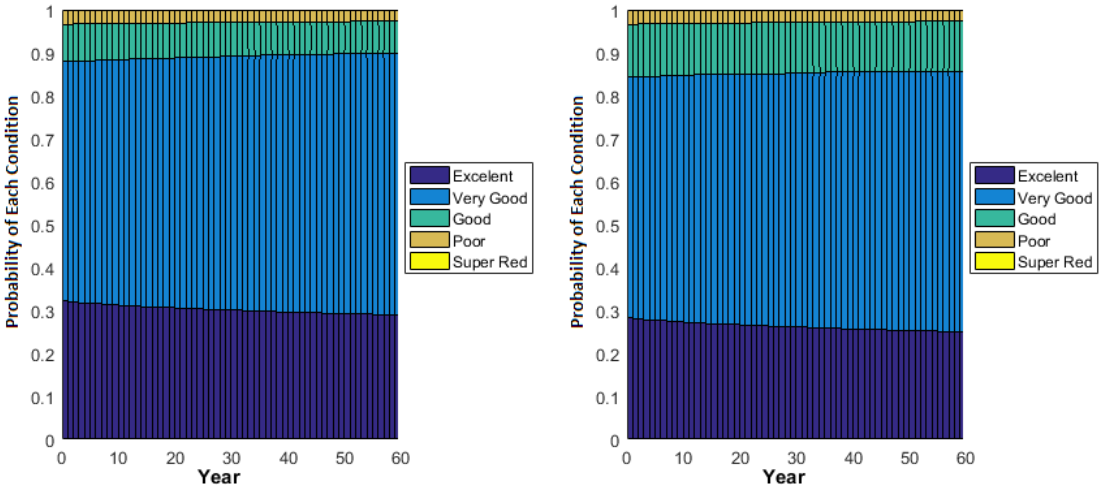


Figure 6.12: Yearly Averages of Tamping, Model A and D



(a) Model A

(b) Model D

Figure 6.13: Yearly Condition Results

6.5.4 Summary

A railway track is a complex system that can be difficult to model. Despite this, the CPN model proposed in this chapter has many features which include:

1. Geometry degradation and rail fault occurrences with the link between them.
2. Usage based, hence changes in the track usage are included in the model.
3. Rail model includes two types of inspections and three types of maintenance to take into the variation in inspection and maintenance type depending on the rail fault type.
4. Track model includes renewals, tamping and stoneblowing, as well as opportunistic maintenance with clustering techniques which resembles track engineer decisions.
5. Can be used on any length of single line track over any time period.

6. Changes to the track speed at any point in the time period being modelled, as well as forced renewal upgrades and upgrades when a renewal is required.
7. Allows change in working window lengths to be modelled.
8. Works as a framework that additional modules can be added to improve the model's capability.
9. Outputs large amounts of useful data such as the number of different maintenance actions and inspections which can be used to calculate the cost of the track modelled. Track quality outputs can be calculated from the outputs separately (Matlab has been used in the Thesis).
10. Due to the model inputs allowing for the user to set the initial track qualities, the model can be applied to any track section with known data. It does not require the track sections to start in a new state. This allows the model to be used on any existing railway.

The modular approach used in the CPN model's design allows additional aspects to be added to the model relatively simply. This is due to utilising a CPN as a decision-making model and framework for other models to run within, which is an extension on how previous research has utilised a CPN technique, (Yianni, 2017; Kilsby, 2017; Audley, 2014). The track geometry degradation model is a linear model with distributed parameters, geometry maintenance is a linear model with distributed residues and the rail fault model is a simple probability based model. All the individual models are stochastic and solved using Monte-Carlo but deterministic models would also work. For example, if a sleeper degradation model was developed, it would be inserted as its own subpage within the model, linking to P2, the primary track section pool (where all track information is stored within the model), similarly to the rail fault model.

6.5.5 Limitations

The developed model has some limitations:

1. Lack of data. Due to lack of data not all significant factors are included within the model, such as the impact of rail type on rail faults.
2. Running Time. The model time to calculate (full-convergence) can take many days depending on the time period and track length modelled. This is due to the complex nature of the model and limitations in the software its built in (CPN-Tools).
3. The clustering method used for opportunistic maintenance assumption of sequential track sections means it cannot account for branches or gaps in the line. Only a single line can be modelled at a time.
4. Rail fault model can only predict the occurrence of one fault per track section at a time (7 day period), when multiple faults could occur in real life. The 7 day period can be reduced to lower the impact of this at the expense of the models running time.

6.5.6 Results

The developed model has demonstrated its ability to assess many aspects within a railway tracks asset management process. The impact of different inputs of the model have been assessed by keeping all other aspects constant. This includes:

- Higher (less strict) geometry thresholds;
  - Lead to less geometry maintenance operations but an increase in rail faults and hence repairs.
  - A reduction in the average geometry quality which may impact safety/passenger comfort.
  - A larger percentage of speed restrictions was also seen (Model A 0.01%; Model B 0.03%) which have associated costs as well as reducing the reputation of the network operator.
- Increasing Traffic;
  - Reduces the track quality as the traffic increases.
  - As traffic levels increase the number of maintenance operations, like tamping, required per year, increases. This is due to the increased rate of geometry degradation and rail fault occurrences caused by increased traffic.
- Increase in time from identifying the need for geometry maintenance and it being completed;
  - Reduces the overall number of track geometry maintenance actions, but increases the number of urgent interventions.
  - Less geometry maintenance, reduces the average track quality, and increases the number of rail faults that need to be maintained.
- Decrease in the period of a track geometry maintenance shift (tamping and stoneblowing);
  - As fewer track sections can be maintained in a shift, less opportunistic maintenance occurs. Leading to more shifts of tamping required.
  - Due to the period which tamping is used to correct the track geometry increases, the number of stoneblowing and renewals within the 60 years decreases.
  - Despite more tamping operations, the average track quality still reduces.

Within this Thesis costs of maintenance activities have been fabricated. To fully understand and analyse the outputs from the model, accurate costs will have to be incorporated. This will enable the user to understand aspects such as if the cost saving from reducing geometry maintenance is greater than the cost of repairing the additional rail faults.

Chapter 7

Conclusions

This thesis outlines the development of a railway track model, capable of accurately predicting future maintenance requirements, and hence costs. The model allows asset management strategies to be predicted and compared to help make informed decisions. The production of the model was undertaken after extensive analysis of not only the geometry degradation, but also maintenance effectiveness, maintenance output rates and rail faults. This is the first study to consider and analyse the maintenance output rates.

Due to the amount of available actual real life data and the degree of analysis undertaken (such as number of factors considered), the analysis in this thesis is one of the most detailed and overarching undertaken in the field of railway track. The multiple aspects analysed were individually modelled within a Coloured Petri Nets (CPNs), which was used as a decision framework and to capture outputs. This is a very novel use of CPN, and makes the final model design extremely flexible (with separate modules (smaller models), joined together). Due to this the model design will scale better than models developed in previous research as more track aspects are included.

The proposed model covers more aspects of a railway track than the ones found it literature. This includes, track geometry degradation, maintenance effectiveness (tamping, stoneblowing and renewals), rail faults (multiple types), rail maintenance (multiple types related to the faults found), geometry and rail inspections, opportunistic maintenance, maintenance shift lengths and changes in traffic levels. Previous models have tended to concentrate on one aspect. The link between the track geometry and the number of rail faults has never been quantified or modelled before, neither has the impact of changing maintenance windows lengths.

The final model uses a novel approach to CPNs, is based on a large amount of research and analysis (as shown in Chapters 2-5), and has demonstrated its ability to represent a railway over a long time period. Additionally, it has shown the impact of changing different asset management policies on the number of maintenance actions and track

condition can be compared and analysed.

7.1 Track Geometry Degradation

Using data supplied by Network Rail (NR), a detailed analysis of degradation rates was performed. It was found that track geometry tended to follow a linear relationship with usage. Many techniques were employed to maximise the amount of usable data, including predicting historical maintenance actions and renewal dates, allowing for the full cycle of a railway track to be analysed. This included up to 10 tamps and 3 stoneblows, whereas previous research by Audley (2014), using a similar sized dataset also from NR, only included up to 2 tamps and did not analyse stoneblowing.

The analysis showed that maintenance history, track speed, sleeper and rail joint types were the primary factors influencing vertical geometry degradation rates. These were expected based on previous literature. The analysis was able to prove that certain options within these factors performed similarly, so could be grouped to maximise the size of datasets to feed into a stochastic model. For example, there were three distinct groups of sleepers (small concrete and large steel, large concrete and timber and small steel), which performed significantly different from each other. As steel sleepers are still relatively new, the results give confidence in their use as large steel sleepers were shown to perform similarly to small concrete ones.

Many aspects were explored, many of which had not been analysed in literature (or minimal previous work), such as the embankments, soil and rock cuttings, tunnels and stations, curvature and cant, electrification, percentage of dirty traffic and the geology of layers under the ballast. The results showed that the bedrock geology, electrification type, dirty traffic percentage and curvature were factors which impacted the rate of degradation. This is important as many of these have not been identified in previous work. Previous research had singled out rail type as being one of the leading factors. On initial analysis the results agreed, but when a multi-level analysis was undertaken (to remove as many factors as possible before comparing), it was shown that the rail type has minimal impact. As the rail type is linked to the sleeper type, it was proved that the improvement from larger rails actually comes from the larger sleepers used with them.

Using the results from the analysis, data was split into similar track sections, based on the primary factors and the groups found within these. Weibull distributions were then used to represent the variability of linear degradation rates of vertical track geometry Standard Deviation (SD) against usage.

7.2 Maintenance Effectiveness

Railway geometry maintenance does not return the track to an as good as new state. The degree of improvement returned from the primary two maintenance actions, tamping and stoneblowing, were analysed. Tamping is much more common than stoneblowing and has been used for a much longer time period. Due to this very few sources in literature discuss the effectiveness of stoneblowing.

The results showed a linear relationship between the initial and resultant quality. Both tamping and stoneblowing were proved to become less effective with continual use on the same track section, this has been previously proved for tamping but not stoneblowing. Additionally, the data showed that maintenance can have a negative effect, with many tamps, and even more for stoneblows, actually reducing the track quality. This has not been mentioned in most previous literature where it is incorrectly assumed all maintenance improves the track quality. Previous research identified track speed as an important factor, but the analysis results showed no impact on stoneblowing and minimal on tamping, except for track speeds >=115MPH, which was significantly different from the rest. Due to the decrease in effectiveness found in stoneblowing after many uses, it is recommended to renew the track after the second stoneblowing operation. This is important knowledge when deciding maintenance regimes.

A model was developed that combined a linear relationship with a stochastic element to account for the variability in the data. This type of model has been used previously for track degradation but not maintenance effectiveness. The linear model related the resultant geometry after maintenance to the geometry before. The analysis showed that the variability in improvement increases with the initial geometry (heteroscedasticity). This has previously not been considered in literature and was used to normalise the variability before a Weibull distribution was fitted. It is an important aspect to consider, as waiting longer to perform maintenance (higher vertical geometry SD), increases the average improvement but also decreases the confidence on estimating the resultant geometry.

The maintenance output rate was analysed as no previous literature existed. The length of maintenance activities (in time) and the output rate (yds/hr) were considered. The results showed that the output rate for stoneblowing is slightly lower than tamping. This increases the cost of stoneblowing (£/yd) compared to tamping due to most costs being per shift (plant and labour)). With an understanding of the output rates of maintenance the impact of the length (in time) of the maintenance shifts can be modelled. This allows decisions such as running trains later in the day, which will reduce the time for maintenance work, to be modelled and understood in advance. The model consists of two Weibull distributions to represent the spread of outputs rates found for tamping and stoneblowing, these are multiplied by the maintenance length (in hours) to obtain an estimate of the length of track that could be maintained in one maintenance shift.

7.3 Rail Faults

Rail faults are common and vary in nature, with 12 separate rail fault types identified on the UK rail network. The analysis undertaken used the occurrence of rail faults associated to the usage not time. This is uncommon in research, where previous research has used time instead, but very important due to the amount of rail faults caused by fatigue. Many aspects were shown to impact the rate of rail faults. By analysing the fault types separately, it was possible to see that the same factors do not impact all types of faults in the same way. This is important as faults types have different inspection and maintenance regimes as well as some being more dangerous than others. This adds more depth to the understanding of rail fault occurrences than previous analysis has shown. For example fatigue related faults were shown to be most depended on the rail age, as these are discovered using ultrasonic inspections, it might be desirable to increase the number of ultrasonic inspections on track with older rail (>20 years old).

Previous research mentioned that rail faults were related to the forces on the rails. As the track geometry governs the smoothness of the running surface, and hence impacts the forces, it was thought that the two must be linked. Previous research has never demonstrated a link between the conditions of the assets within the railway track system. As the railway track is operated as one system, it is important to understand how decisions based on one asset in the system, impacts the other assets. For example, a decision to save money by reducing the number of geometry maintenance actions, may increase the number of faults and hence costs of maintaining the rail. By relating fault rates (by usage), with the track geometry recordings it was possible to quantify the link between them. This proved, for all rail fault types, that the occurrence rate increases with worsening track geometry.

A rail fault model has been proposed that includes the proven link between the occurrence rate of rail faults and track geometry. As this link has never been shown in previous literature, it has never been considered in a rail fault model. The link was quantified using a cubic function. To obtain the probability of each fault type, the average track geometry for a chosen time period is entered, with the results multiplied by the usage in the same period. Once a probability is found, occurrences are sampled using Monte-Carlo method. The model allows the impact on the amount of rail faults from different track geometry SD to be calculated and understood.

7.4 Railway Track Asset Management Model

A railway track is a complex system, with many assets and degradation mechanisms. There are also a lot of possible asset management decisions which need to be understood. Due to this a complete railway track model is complex. Utilising the knowledge obtained from the extensive analysis undertaken, a hierarchical CPN model was developed. It

incorporated the individual models proposed in this thesis and the required decision making aspects to make an asset management model (such as what type and when to maintain). The model was designed as separate modules joined together by a master Net, where each module has a different aspect to model, such as rail faults or opportunistic maintenance. This design allows the model to easily scale to include more aspects, such as including a sleeper model. By combining the separate models together within the CPN, information is shared and interactions can be modelled. This is required as aspects such as the rail model, need the geometry information from the track geometry degradation and maintenance model to calculate (due to the new link found in this thesis between geometry and rail faults). The track geometry degradation model is a linear model with distributed parameters, geometry maintenance is a linear model with distributed residues and the rail fault model is a simple probability based model. This novel approach to using CPNs as a decision framework for other models, allowing them to work together and interact, has been proved to be very successful with many advantages, such as scalability and readability.

The model introduced includes many aspects which have not previously been considered in literature including: modelling the working window length for maintenance, maintenance of geometry up to a high amount of previous tamps, stoneblows being based on actual data, rail faults link to track geometry and different types of rail faults and inspections (some rail faults can only be found with certain inspection methods). Other aspects such as opportunistic maintenance with a simple clustering methodology are also included.

The model allows for many inputs to be changed, allowing the user to understand the possible impact of different asset management policies. This includes aspects such as inspection intervals, time to complete different maintenance actions, rules of when to use stoneblowing and renewals, changes in traffic, assets or line speeds, maintenance thresholds and maintenance shift lengths. Many aspects can be outputted for analysis, most importantly; the number of different inspections and maintenance actions (overall and in each year), number of these which were urgently required, expected track quality (overall and in each year) and number of speed restrictions. Due to the stochastic nature of the model, each of these outputs results give a range of possibilities. All these aspects together create a very comprehensive railway track model.

7.5 Future Work

Despite this thesis developing a detailed asset management model for railway track, due to the complex nature of degradation and maintenance of track there are still many aspects of improvement. These include:

1. Improvements to the separate models. Many models, such as rail faults, have been introduced, each of which can be improved individually, improving the overall

model. Some possibilities include:

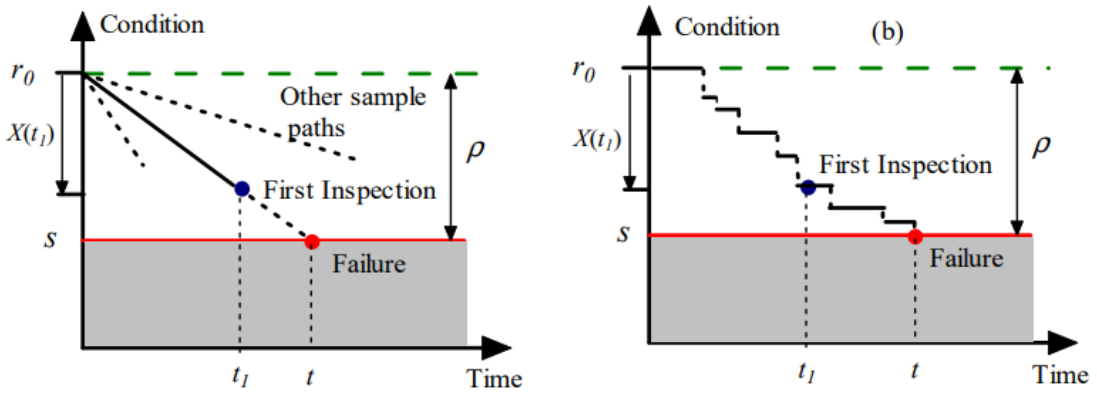
- Geometry Degradation Model. One possible approach assessed was the Gamma process. The Gamma Process is well suited to model gradual degradation, which is monotonic (condition always decreasing), such as wear and erosion, (Abdel-Hameed, 1975). The Gamma process relates the degradation rates to the gamma distribution. The distribution varies in time, as time goes by the time to failure (passing maintenance thresholds) decreases, so does the variation. The gamma process is related to Markov and such is memoryless, calculating degradation in small, independent, random steps. This is shown in Figure 7.1, where the linear fit presumes a constant degradation rate with time, unlike the gamma process. The linear fit presumes that after two inspections (or one with a known starting quality), the point of failure is known. The Gamma process results have been shown to be more realistic, especially at the tails of possible results, (Yuan, 2005). Due to the random movements in each step, issues such as the sampling of very low degradation rates leading to track sections in the model not requiring maintenance for an unrealistic period of time maintenance, are removed. The Gamma process could be utilised in the CPN similarly to the method currently used in this thesis (Weibull distribution of degradation rates), as it can be used to calculate the track quality at any point of time (or usage), which can then be checked against the maintenance thresholds by the CPN model. The Gamma process would need to be linked to the track geometry maintenance, to account for maintenance not returning track to as good as new.
- Rail Fault Model. The current model links the impact of track geometry on the fault occurrence rate using regression. Due to this it is not possible to simply calculate the times to a fault, but have to solve by Monte-Carlo for each time step separately to assess if a fault occurred. This makes the fault rate memoryless with a constant hazard rate. Reliability/survival models are very common and have been well researched. One of the most common types are Cox models, in particular Proportional Hazards Models. These use regression to link survive times (times to failure) to predictors. The model type allows for both categorical and quantitative inputs (such as rail joint type or track geometry), (Cox, 1972). Due to the link to track geometry, which is a function of time, the model will have to be extended as discussed by Martinussen and Scheike (2006). This type of model would bring many advantages over the current memoryless model. The Cox model would allow times to the next fault to be calculated, which is much quicker than the current process of checking probabilities in time steps, until a fault occurs. The Cox model could also take into account more of the factors which influence the rate of faults, as these would become inputs to the regression model. The assumptions of the Cox model will have to be checked before its application, such as constant proportionality between the different variables hazard curves, Calavas (1994).

If the rail fault model was changed to a Cox model approach, some minor alterations to the CPN model would be required.

2. Maintenance machine availability. The model developed presumes that there will always be tampers and stoneblowers available. This will not always be true with situations where stoneblowers are used purely because a tamping machine may not be available. The inclusion of this would also allow decisions of the maintenance fleet size to be modelled and informed decisions made. Adding a machine availability model into the present Petri Net (PN) would be possible and has previously been done Audley (2014).
3. Develop a clustering methodology for the Opportunistic Maintenance which can take into account track sections with multiple lines and branches. This will allow larger sections of railway lines to be modelled.
4. Integrate more maintenance techniques. Only the main three types of track geometry maintenance; tamping, stoneblowing and renewals have been included within the model. Additional methods such as partial renewals and manual maintenance could be included.
5. Sleeper and Switches and Crossing (S&C) models. This thesis has only explored the degradation and maintenance of plain line track and rails. An additional module which includes a model for railway sleepers could be added to the Petri-Net model in this thesis. A S&C model would be important due to the higher levels of derogation and risk of failure at these track locations.
6. Obtain accurate costs of work. These would have to be related to shift costs (of varying shift lengths) for performing tamping and stoneblowing. The inspections and rail faults would have to be a cost each. Due to the varying nature of costs of maintenance these costs could also be distributed. Multiplied by the quantities of inspections and maintenance outputted by the model to obtain predicted costs.
7. More data. This will improve the analysis undertaken, the distribution fits used in the model and the amount the data can be split by the significant factors. This could be from more years of NR collecting data or obtaining data from other sources such as countries.
8. Speed up the model. The model is slow due to its complexity as well as the limitations of the program it was built in (CPN-Tools), such as lack of multi-core Central processing unit (CPU) support. The PN model could be sped up by developing the model in a coding language such as C++ as shown by Yianni (2017) and Kilsby (2017). Yianni (2017) also shows how utilising a computer Graphics Processing Unit (GPU) can sped up PNs even further. Additional changes to the calculation method could also speed up the time for the model results to converge such as using Latin Hypercube instead of Monte-Carlo. Speeding up the model would allow further results to be run and compared.
9. Optimisation. Despite the developed model being able to output results based on the users input it currently cannot tell the user the optimum inputs and hence the optimum asset management strategy. Optimisation works well on PNs models as

shown by Yianni (2017), Kilsby (2017) and Audley (2014) who all utilised generic algorithms. Genetic algorithms basically run the model a certain number of times varying the inputs. It then takes the inputs with the best results and combines them (like natural selection, best survive and have offspring) . Results are calculated for the new inputs and the process repeats until the results stop improving. For an asset management model if everything has accurate costs minimising the model output cost would be the most effective asset management decisions.

As the model does not need asset conditions to start in a new state, it can be applied to any railway track, as long as the basic input information is known (last inspection and quality, track assets, line speed and maintenance history). To enable the model to be used commercially, such as by NR, points 5 and 7 will need to be rectified. The model needs to be rebuilt utilising faster technology to greatly reduce the running time, as the current runtime of over a week to obtain results is not feasible in a commercial setting. Also, to fully compare the results out of the model, the outputs such as number of tamping actions, need to be accurately costed.



(a) Linear Degradation

(b) Gamma Process Degradation

Figure 7.1: Example of Track Recordings of Vertical Alignment, (Yuan, 2005)

Bibliography

CPN Tools Online Knowlege Base. URL `http://cpntools.org/`.

Abdel-Hameed, M. A gamma wear process. *IEEE Transactions on Reliability*, R-24 (2): pp. 152–153, 1975.

Akhtar, M., Davis, D., Maal, L., Gordon, J., and Jeong, D. Effects of track parameters on rail joint bar stresses and crack growth. Technical report, AREMA, 2010.

Alva-Hurtado, J. E. and Selig, E. T. Permanent strain behavior of railroad ballast. In Balkema, A., editor, *Proceedings of The 10th International Conference on Soil Mechanics and Foundation Engineering*, pages 543–546, Stockholm, Sweden, June 1981.

Andrade, R. A. and Teixeira, F. P. Uncertainty in rail-track geometry degradation: Lisbon-oporto line case study. *Journal Of Transportation Engineering*, 137(3):193–200, March 2011.

Andrews, J. A modelling approach to railway track asset management. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, pages 1–18, July 2012.

AREMA. Arema manual for railway engineering - roadway and ballast. Technical report, American Railway Engineering and Maintenance-of-way Association, 2012.

ARTC. Steel sleepers - usage and installation standards. Technical report, Austrailian Rail Track Corporation Ltd, February 2009.

Audley, M. and Andrews, J. D. The effects of tamping on railway track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(4):376–391, July 2013.

Audley, M. *Rail Track Geometry Degradation and Maintenance Decision Making*. PhD thesis, University of Nottingham, 2014.

Aursudkij, B. *A Laboratory Study of Railway Ballast Behaviour under Traffic Loading and Tamping Maintenance*. PhD thesis, University of Nottingham, Nottingham, 2007.

Balfour Beatty Rail. *Track Sections Datasheet*, 2010.

Bright, A. Office of rail regulation v network rail infrastructure limited. Technical report, Judiciary of England and Wales, 2011.

British Standard Institution. Bs 5760-23:1997: Reliability of systems, equipment and components - part 23: Guide to life cycle costing. Technical report, BSi, 2004.

BS EN 13450. Aggregates for railway ballast. Technical report, British Standard Institution, 2002.

Caetano, L. F. and Teixeira, P. F. Predictive maintenance model for ballast tamping. *Journal of Transportation Engineering*, 142(4), April 2016.

Calavas, F. B. C. D. D. Advantages and inconveniences of the cox model compared with the logistic model: application to a studyof risk factors of nursing cow infertility. *Veterinary Research,BioMed Central*, 25 (2-3):pp.134–139, 1994.

Calla, C. *Two Layered Ballast System for Improved Performance of Railway Track*. PhD thesis, Coventry University, Coventry, December 2003.

Chattopadhyay, G., Larsson-Kråik, P. O., and Kumar, U. Study of ndt rail inspection on malmbanan. Internal research report, Banverket, Luleå Railway Research Center, Sweden, 2005.

Chrismer, S. M. and Selig, E. T. Mechanics-based model to predict ballast-related maintenance timing and costs. In *Report No. R-863*, pages 208–211, AAR Technical Centre, Chicago, Illinois, USA, 1994. Association of American Railroads.

Clarke, S. and Prescott, D. R. Identifying the factors influencing track degradation on the uk rail network. In *The Stephenson Conference 2017: Research for Railways*, 25-27, London, UK, April 2017.

Cox, D. R. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 34, No. 2:pp. 187–220 (34 pages), 1972.

Dahlberg, T. Railway track settlements - a literature review. Report for the eu project supertrack, Division of Solid Mechanics, IKP, Linköping University, SE-581 83 Linköping, Sweden, July 2007.

Dell'orco, M., Ottomanelli, M., Caggiani, L., and Sassanelli, D. New decision support system for optimization of rail track maintenance planning based on adaptive neuro-fuzzy inference system. *Transportation Research Record: Journal of the Transportation Research Board*, 2043:49–54, 2008.

Dell'orco, M., Ottomanelli, M., Pace, P., and Pascoschi, G. Intelligent decision support tools for optimal planning of rail track maintenance. In *Proceedings of 13th Mini-EURO Conference on "Handling Uncertainty in the Analysis of Traffic and Transportation Systems"*, pages 218–223, Dept. of Highways and TransportationŰ Polytechnic University of Bari, Italy, June 2002.

Denby, N. Rail defect analysis issue 2. Internal NR Document, October 2013.

Dernoncourt, F. Introduction to fuzzy logic. Technical report, Massachusetts Institute of Technology, January 2013.

Dingwall, P. *Track Standards Manual - Section 8: Track Geometry*. Railtrack, Railtrack House, Euston Square, London, NW1 2EE, issue 2 edition, December 1998.

Dolven, O. F., Lindqvist, B. H., and Hokstad, P. R. Statistical modelling and analysis of failure and inspection data for a railway line. In Spitzer, C., Schmocker, U., and Dang, V., editors, *Probabilistic Safety Assessment and Management*, pages 591–596. Springer London, 2004.

Ebersohn, W. and Selig, E. T. Use of track geometry measurements for maintenance planning. Technical report, Transportation Research Record, 1994.

Ebrahimi, A. *Deformational Behavior of Fouled Railway Ballast*. PhD thesis, University of Wisconsin-Madison, 2011a.

Ebrahimi, A. *Deformational Behavior of Fouled Railway Ballast*. PhD thesis, University of Wisconsin-Madison, 2011b.

Ebrahimi, A. and Keene, A. K. Maintenance planning of railway ballast. Technical report, AREMA, 2011.

ERRI. *Decision Support System for Track Maintenance and Renewal - ECOTRACK - Decision Rules and Data Requirements*. European Rail Research Institute, Utrecht, erri d 187/dt 299 edition, April 1994.

Gershenson, C. Artificial neural networks for beginners. Technical report, Cornell University, 2003.

Gordon, D. Track strategic planning application: Technical basics - report to network rail. Technical report, Serco Assurance and Network Rail, December 2006.

Grabowski, B. Ş$p < 0.05\check{T}$ might not mean what you think: American statistical association clarifies p values. *J Natl Cancer Inst.*, 2016.

Grassie, S. L. 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*, 226:235–242, October 2011.

Gross, T. and Sayama, H. *Adaptive Networks - Theory, Models and Applications*. Springer, 2009.

Gular, H., Jovanovic, S., and Evren, G. Modelling railway track geometry deterioration. *ICE Proceedings - Transport*, 164:65–75, May 2011.

Halcrow. Track tier 2 whole life cycle cost model. Technical report, Network Rail, June 2012.

Hokstad, P., Langseth, H., Lindqvist, B. H., and Vatn, J. Failure modeling and maintenance optimization for a railway line. Technical report, SINTEF Technology and Society, Department of Safety and Reliability, Trondheim, Norway, 2005.

HSE Potters Bar Investigation Board. Train derailment at potters bar 10 may 2002. Technical report, Health and Safety Executive, 2003.

Institute of Asset Management. Pas 55-1:2008 - asset management - part 1: Specification for the optimized management of physical assets. Technical report, The British Standard Institution, 2008a.

Institute of Asset Management. Pas 55-2:2008 - asset management - part 2: Guidelines for the application of pas 55-1. Technical report, The British Standard Institution, 2008b.

Interflon. Rail jointing - fishplate lubrication, 2011.

Iwnicki, S., Grassie, S., and Kik, W. Track settlement prediction using computer simulation tools. *Vehicle System Dynamics*, 33(13):37–46, 2000.

Jang, J., Sun, C., and Mizutani, E. Fuzzy inference systems. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, pages 73–91, 1997.

Jang, J.-S. R. Anfis : Adaptive-network-based fuzzy inference system. *IEEE Transactions On Systems, Man, And Cybernetics*, 23(3):665–685, May/June 1993.

Jensen, K. A brief introduction to coloured petri nets. In *Brinksma E. (eds) Tools and Algorithms for the Construction and Analysis of Systems. TACAS 1997. Lecture Notes in Computer Science, vol 1217*, 1997.

Jensen, K. and Kristensen, L. M. *Coloured Petri Nets.* Springer-Verlag Berlin, 2009.

Jovanovic, S. and Pearce, M. Ecotrack: An overview of the systemŠs functionality and implementation to date. Technical report, European Rail Research Institute / Delft University of Technology, The Netherlands, 2000.

Kalogirou, S. A. *Artificial Intelligence in Energy And Renewable Energy Systems.* Nova Science Publishers, 2006.

Kilsby, P. *Modelling railway overhead line equipment asset management.* PhD thesis, University of Nottingham, 2017.

Kirchhof, G. *Encyclopedia of Soil Science*, volume 2. CRC Press, 2006.

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

Kuttelwascher, C. Track ballast in austria. Technical Report 2, Plasser & Theurer, 2012.

Lei, X. and Feng, Q. Analysis of stability of continuosly welded rail track with finite elements. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit May 1, 2004 218: 225-233*, pages 225–226, 2004.

Lewis, R. Track geometry recording and usage. Notes for a lecture to Network Rail, April 2011.

Li, D. and Selig, E. T. Cumulative plastic deformation for fine-grained subgrade soils. *Journal of Geotechnical and Geoenvironmental Engineering, ASCE*, 122:1006–1014, December 1996.

Li, S. Railway sleeper modeling with deterministic and non-deterministic support conditions. Master degree project, Royal Institute of Technology, Stockholm, 2012.

Lichtberger, B. *Track Compendium.* Eurailpress Tetzlaff-Hestra GmbH and Co., Hamberg, 2005.

M Miwa, T. I. and Oyama, T. Modeling the transition process of railway track irregularity and its application to the optimal decision-making for multiple tie tamper operations. In *Railway Engineering 2000 - 3rd International Conference & Exhibition*, London, July 2000.

Martinussen, T. and Scheike, T. H. *Dynamic Regression Models for Survival Data.* Springer, 2006.

Ministry of Railways. *Manual of Instructions on Long Welded Rails.* Government of India, 2 edition, 2005.

Mundrey, J. S. *Railway Track Engineering.* Tata McGraw-Hill Education, 2010.

Network Rail. Categorisation of track. Technical Report 1, Network Rail, March 2011.

Network Rail. Track geometry and ballast condition on nr track - parametric analysis for tier 1 and 2 modelling. Technical Report 1, Network Rail, December 2012.

Network Rail. Annual expenditure 2018/19. Online, 2019.

Norris, J. R. *Markov Chains.* Number no. 2008 in Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998. ISBN 9780521633963.

Ottomanelli, M., Dell'Orco, M., and Sassanelli, D. A decision support system based on neuro-fuzzy system for railroad maintenance planning. In *Proceedings of the Seventh International Conference on Enterprise Information Systems*, pages 43–49, Miami, USA, May 2005. ICEIS.

Oxera. What is the contribution of rail to the uk economy?, July 2014.

Peacock, C. F. M. E. N. H. B. *Statistical Distributions.* Wiley, forth edition, 2011.

Prescott, D. and Andrews, J. Investigating railway track asset management using a markov analysis. *Rail and Rapid Transport*, 229(4):402–416, May 2015.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. *Numerical Recipes - The Art of Scientific Computing.* Cambridge University Press, third edition, 2007.

Profillidis, V. A. *Railway Management and Engineering.* Ashgate, third edition, 2006.

Quiroga, L. M. and Schnieder, E. Monte carlo simulation of railway track geometry deterioration and restoration. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 226:274–282, 2012.

Reid, L. *Rail Quality and Maintenance for Modern Railway Operation.* Kluwer Academic Publishers, 1993.

Rojas, R. *Neural Networks - A Systematic Introduction.* Springer, 1996.

Sato, Y. Japanese studies on deterioration of ballasted track. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, 24:197–208, 1995.

Sawley, K. J. Wheel/rail profile maintenance. Technical Report 461, International Union of Railways (UIC), Transportation Technology Center, Pueblo, Colorado, 2001.

Schneesweiss, W. G. *Petri Net Picture Book.* LiLoLe-Verlag GmbH, 2004.

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

Shafahi, Y. and Hakhamameshi, R. Application of maintenance management model based on markov chain and probabilistic dynamic programming for the iranian railways. *Trans A: Civil Engineering*, 16(1):87–97, February 2009.

Shafahi, Y., Masoudi, P., and Hakhamaneshi, R. Track degredation predict models, using markov chain, artificial neural and neuro-fuzzy network. Technical report, Sharif University of Technology, Theran, Iran, 2008.

Shenton, M. J. Ballast deformation and track deterioration. In *Track technology, Proceedings of a conference held at University of Nottingham*, pages 253–265. Thomas Telford Ltd, July 1985.

Shi, X. *Predict of Permanent Deformation in Railway Track.* PhD thesis, University of Nottingham, 2009.

Simões, G. M. P. Rams analysis of railway track infrastructure. Master's thesis, Insitituto Superior Technico - Universidade Technica de Lisboa, September 2008.

Smith, R. A. Railway fatigue failures: An overview of a long standing problem. *Mat.-wiss. u. Werkstofftech*, 36(11):697–705, 2005.

Steenbergen, M. J. M. M. and Esveld, C. Relationship between the geometry of rail welds and the dynamic wheel-rail response: Numerical simulations for measured welds. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 2006 220: 409*, pages 409–410, 2006.

Stephen M Famurewa, M. R., Tao Xin and Kumar, U. Optimisation of maintenance track possession time: A tamping case study. *Rail and Rapid Transit*, 229(1):12–22, 2015.

Stephens, M. A. Use of the kolmogorov-smirnov, cramer-von mises and related statistics without extensive tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 32(1):pp 115–122, 1970.

Sung, W., hsiang Shih, M., Lin, C.-I., and Go, C. G. The critical loading for lateral buckling of continuous welded rail. *Journal of Zhejiang University SCI*, 2005 6A(8): 878–879, 2005.

Swarnakar, P. K. Comparision, testing and statistical analysis of prestressed concrete sleeper manufactured for indian railway. Major project report, Delhi Technological University, Delhi, 2012.

UIC. Guidelines for the application of asset management in railway infrastructure organisations. Technical report, International Union of Railways, 2010.

Velt, P. *RTR Special - Maintenance and Renewal*, chapter Track Quality - Luxury or Necessity, pages 8–12. DVV Media Group GmbH, Eurailpress, July 2007.

Williams, J. Track asset policy. Technical report, November 2012.

Williams, J. *Network Rail Asset Report Manual - Definitions for the Report of Track Sustainability.* Network Rail, October 2013.

Wilson, A. *Rail Defects Handbook.* RailCorp, New South Wales, Australia, June 2012.

Wright, S. E. Damage caused to ballast by mechanical maintenance techniques (tm-td-015). Technical report, British Rail Research, 1983.

Yianni, P. *A modelling approach to railway bridge asset management.* PhD thesis, University of Nottingham, 2017.

Yuan, M. D. P. X. Gamma process model for reliability analysis and replacement of aging structural components. 2005.

Zakeri, J.-A. and Rezvani, F. H. Failures of railway concrete sleepers during service life. *International Journal of Construction Engineering and Management*, 1:3–4, 2012.

Zarembski, A. Stone blowing: An alternate approach to track surfacing. *Railway Track & Structures*, 1:177–178, 2005.

Zong, N. and Dhanasekar, M. Experimental studies on the performance of rail joints with modified wheel/railhead contact. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, pages 1–3, 2013.

Zwolski, J. Railways - continuous welded rails and grade crossings. Presentation, 2012.

Chapter 8

Appendix

8.1 Literature Review

Table 8.1: Comparison of deterministic track degradation models

Model	Input	Outputs	Additional Factors	Relationship	Degradation Rate	Page
Alva-Hurtado and Selig's strain model	Cycles	Strain		Semi-logarithmic	Decreasing	31
Selig and Waters' strain and settlement models	Cycles or tonnage	Strain or settlement		Power	Decreasing	32
Shenton's settlement model	Cycles	Settlement	Axle load, rail section, sleeper type and spacing, track and subgrade stiffness, previous maintenance, ballast condition	Power + Linear	Fast up to 1,000,000 cycles then slows	33
Sato's settlement models	Cycles or tonnage	Settlement	Vertical acceleration required to initiate slip, sleeper pressure, peek acceleration of the ballast particles (ballast condition and saturation), and ballast packing quality.	Exponential + Linear	Fast then slows with the point of change depending upon the initial ballast packing	33
TU Munich settlement models	Cycles	Settlement	Pressure in the ballast	Other	Fast up to 10,000 cycles then slows	34
Office of Research and Experiments' track geometry models	Tonnage	Settlement or vertical, alignment or twist geometry SD	Subgrade quality, ballast condition	Semi-logarithmic	Fast up to 2,000,000 standard axles then slows	34
Velt's track quality model	Time	Track quality		Exponential	Increasing	35

Continues on the next page

Model	Input	Outputs	Additional Factors	Relationship	Degradation Rate	Page
Office of Research and Experiments' track quality model	Tonnage	Combined Track Record (CTR) index	Axle load	Other	Decreasing	36
Sato's track damage model	Tonnage	Track irregularities	Speed, rail type (Continuous Welded Rail (CWR) vs jointed), subgrade quality, quasi static pressure in the ballast, acceleration of the rail and other rail properties.	Other	N/A	36
Wisc-Rail	Cycles or tonnage	Strain or vertical geometry SD and Ballast Fouling Index (BFI)	Subgrade classification, ballast saturation, traffic axle loads, ballast fouling, subgrade static strength, ballast thickness, rate of fouling.	Semi-logarithmic up to 0.3 Million Gross Tonnage (MGT) then linear	Fast but decreasing up to 10,000 cycles (0.3 MGT) then constant	36
Multivariate statistical analysis approach	Tonnage	Degradation rate of twist, alignment, cant or vertical geometry	Traffic load, velocity, curvature (cant), gradient, sleeper, rail, rail length, falling rocks, flood. No rail for vertical, twist and cant.	Linear	Constant	40
ECOTRACK geometry degradation rate model	Tonnage	Geometry degradation rate and maintenance effectiveness	Infrastructure with conditions	Linear	Constant	41
Network Rail's geometry degradation model	Time	Vertical geometry SD	Infrastructure, conditions, traffic (axles, volumes), speed, rail properties, static and dynamic forces, Rolling Contact Fatique (RCF), sleeper spacing and stiffness and BFI.	Exponential-Power	Decreasing	43

Table 8.2: Comparison of adaptive network and fuzzy interface systems track degradation models

Model	Inputs	Outputs	Model Type	Page
Type 2 fuzzy reasoning model	SD of alignment, vertical level, cross-level, gauge, rail wear and corrugations	Time to maintenance	Fuzzy Interface System (FIS)	46
Ottomanelli et al.'s maintenance occurrence model	SD until alignment and vertical level thresholds, time since last tamping action	Time to tamping maintenance	Adaptive Neural-based Fuzzy Inference System (ANFIS)	46
Dell'orco et al.'s Maintenance Occurrence Model	SD of alignment, vertical level and cross-level, previous number of tamping actions and time since last tamp	Time to tamping maintenance	ANFIS	47
Shafahi et al.'s Artificial Neural Network and Adaptive Neural-based Fuzzy Inference System Models	CTR index state for last five years, traffic volume, maximum speed, topography, gradient and curvature	CTR index state for the next year	Artificial Neural Network (ANN) and ANFIS	48

Table 8.3: Comparison of stochastic track degradation models

Model	Main Input	Outputs	Additional Factors	Relationship	Degradation Rate	Page
Quiroga and Schnieder's track quality model	Time	Track quality		Exponential	Increasing	49
Andrade and Teixeira's track geometry degradation model	Tonnage	Vertical geometry SD	Bridges, Plain Line (PL), stations and switches	Linear	Constant	49
Caetano and Teixeira's track geometry degradation model	Tonnage	Vertical geometry SD	Amount of Previous Data	Linear	Constant	50
Shafahi and Hakhamameshi's discrete-time Markov chain degradation model	Time	CTR index	Topography (plain, hilly etc.), traffic (light, heavy)	N/A	Increasing	51
Prescott and Andrews's continuous-time Markov chain model	Time	Amount of maintenance, renewals and speed and line restrictions. Time spent in each condition state.	Previous maintenance actions	Unknown		53
Andrews's Petri Net asset management model	Time	Amount of maintenance, renewals and line restrictions.	Previous maintenance actions	Linear	Constant	58
Network Rail's geometry faults model	Time	Occurrences for geometry faults	Vertical geometry SD	NA	NA	60

8.2 Track Geometry Degradation

Maintenance	R	1T	2T	3T	4T	5T	6T	7T	8T	9T	10T	1S	2S	3S	4S	5S	6S	7S
Data A (Usage)	6022	920	240	58	22	7	1	0	0	0	0	239	41	10	4	2	0	0
Data A (Time)	6080	927	242	59	23	7	1	0	0	0	0	241	41	10	4	2	0	0
Data B (Usage)	0	3942	3284	2584	1889	1423	1344	1210	1397	1422	1442	764	158	55	23	11	3	1
Data B (Time)	0	3990	3296	2595	1901	1431	1357	1224	1415	1430	1447	772	158	55	23	11	3	1
Total (Usage)	6022	4862	3524	2642	1911	1430	1345	1210	1397	1422	1442	1003	199	65	27	13	3	1
Total (Time)	6080	4917	3538	2654	1924	1438	1358	1224	1415	1430	1447	1013	199	65	27	13	3	1

8S	TS	T2S	T3S	T4S	T5S	T6S	2TS	2T2S	2T3S	2T4S	3TS	3T2S	3T3S	3T4S	3T5S	4TS	4T2S	4T3S	4T4S
0	33	5	1	0	0	0	7	0	1	0	1	0	0	0	0	1	0	0	0
0	33	5	1	0	0	0	7	0	1	0	1	0	0	0	0	1	0	0	0
1	671	153	52	12	3	1	694	141	35	7	691	110	26	13	2	574	126	41	9
1	673	153	52	12	3	1	696	141	35	7	692	110	26	13	2	576	126	41	9
1	704	158	53	12	3	1	701	141	36	7	692	110	26	13	2	575	126	41	9
1	706	158	53	12	3	1	703	141	36	7	693	110	26	13	2	577	126	41	9

4T5S	4T6S	5TS	5T2S	5T3S	5T4S	5T5S	6TS	6T2S	6T3S	6T4S	7TS	7T2S	7T3S	8TS	8T2S	9TS	Total	Maintenance
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7615	Data A (Usage)
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7686	Data A (Time)
2	1	591	132	36	10	2	598	80	23	8	718	141	27	691	137	779	28290	Data B (Usage)
2	1	591	132	36	10	2	602	80	23	8	721	141	27	692	138	782	28466	Data B (Usage)
2	1	591	132	36	10	2	598	80	23	8	718	141	27	691	137	779	35905	Total (Usage)
2	1	591	132	36	10	2	602	80	23	8	721	141	27	692	138	782	36152	Total (Time)

Table 8.4: Amount of Poskeys With Degradation Data

8.2.1 Stations and Tunnels

Data Amounts

Stations and Tunnels		Neither	Stations	Tunnels
Data A Usage	Datasets Used	7	2	2
	Data points Used	6185	282	318
Data B Usage	Datasets Used	34	15	6
	Data points Used	22413	1031	341

Table 8.5: Data Amounts for Stations and Tunnels

2-Tailed Hypothesis Tests

Stations and Tunnels	Neither	Stations	Tunnels
Neither	1.00	0.39	0.05
Station	0.39	1.00	0.03
Tunnel	0.05	0.03	1.00

Table 8.6: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Stations and Tunnels (Data A, Usage)

Stations and Tunnels	Neither	Stations	Tunnels
Neither	0.00	0.00	0.75
Station	0.00	0.00	0.75
Tunnel	0.75	0.75	0.00

Table 8.7: Mean 2 Tailed K-S and Mann Whitney U Test for Stations and Tunnels (Data A, Usage)

Stations and Tunnels	Neither	Stations	Tunnels
Neither	1.00	0.30	0.35
Station	0.30	1.00	0.26
Tunnel	0.35	0.26	1.00

Table 8.8: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Stations and Tunnels (Data B, Usage)

Stations and Tunnels	Neither	Stations	Tunnels
Neither	0.00	0.40	0.33
Station	0.40	0.00	0.33
Tunnel	0.33	0.33	0.00

Table 8.9: Mean 2 Tailed K-S and Mann Whitney U Test for Stations and Tunnels (Data B, Usage)

1-Tailed Hypothesis Tests

Stations and Tunnels	Neither	Stations	Tunnels
Neither	NA	0.20	0.49
Station	0.76	NA	0.50
Tunnel	0.52	0.50	NA

Table 8.10: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Stations and Tunnels (Data A, Usage)

Stations and Tunnels	Neither	Stations	Tunnels
Neither	NA	0.00	0.50
Station	0.00	NA	0.50
Tunnel	0.25	0.50	NA

Table 8.11: Mean 1 Tailed K-S and Mann Whitney U Test for Stations and Tunnels (Data A, Usage)

Stations and Tunnels	Neither	Stations	Tunnels
Neither	NA	0.26	0.62
Station	0.74	NA	0.73
Tunnel	0.38	0.33	NA

Table 8.12: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Stations and Tunnels (Data B, Usage)

Stations and Tunnels	Neither	Stations	Tunnels
Neither	NA	0.40	0.17
Station	0.00	NA	0.00
Tunnel	0.25	0.42	NA

Table 8.13: Mean 1 Tailed K-S and Mann Whitney U Test for Stations and Tunnels (Data B, Usage)

8.2.2 Track Type

Data Amounts

Track Type		CWR	Jointed
Data A Usage	Datasets Used	7	0
	Data points Used	6663	0
Data B Usage	Datasets Used	35	11
	Data points Used	22989	549

Table 8.14: Data Amounts for Track Type

2-Tailed Hypothesis Tests

Track Type	CWR	Jointed
CWR	1.00	0.00
Jointed	0.00	1.00

Table 8.15: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Track Type (Data B, Usage)

Track Type	CWR	Jointed
CWR	0.00	1.00
Jointed	1.00	0.00

Table 8.16: Mean 2 Tailed K-S and Mann Whitney U Test for Track Type (Data B, Usage)

1-Tailed Hypothesis Tests

Track Type	CWR	Jointed
CWR	NA	0.00
Jointed	1.00	NA

Table 8.17: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Track Type (Data B, Usage)

Track Type	CWR	Jointed
CWR	NA	1.00
Jointed	0.00	NA

Table 8.18: Mean 1 Tailed K-S and Mann Whitney U Test for Track Type (Data B, Usage)

8.2.3 Track Construction

Data Amounts

Track Construction		A	B	C	D
Data A Usage	Datasets Used	7	4	1	0
	Data points Used	5490	1036	47	0
Data B Usage	Datasets Used	24	29	15	7
	Data points Used	6546	15085	1024	331

Table 8.19: Data Amounts for Track Construction

2-Tailed Hypothesis Tests

Track Construction	A	B
A	1.00	0.21
B	0.21	1.00

Table 8.20: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Track Construction (Data A, Usage)

Track Construction	A	B
A	0.00	0.75
B	0.75	0.00

Table 8.21: Mean 2 Tailed K-S and Mann Whitney U Test for Track Construction (Data A, Usage)

Track Construction	A	B	C	D
A	1.00	0.08	0.00	0.00
B	0.08	1.00	0.02	0.00
C	0.00	0.02	1.00	0.00
D	0.00	0.00	0.00	1.00

Table 8.22: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Track Construction (Data B, Usage)

Track Construction	A	B	C	D
A	0.00	0.67	1.00	1.00
B	0.67	0.00	0.90	1.00
C	1.00	0.90	0.00	1.00
D	1.00	1.00	1.00	0.00

Table 8.23: Mean 2 Tailed K-S and Mann Whitney U Test for Track Construction (Data B, Usage)

1-Tailed Hypothesis Tests

Track Construction	A	B
A	NA	0.11
B	0.92	NA

Table 8.24: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Track Construction (Data A, Usage)

Track Construction	A	B
A	NA	0.75
B	0.00	NA

Table 8.25: Mean 1 Tailed K-S and Mann Whitney U Test for Track Construction (Data A, Usage)

Track Construction	A	B	C	D
A	NA	0.06	0.00	0.00
B	0.94	NA	0.01	0.00
C	0.99	0.94	NA	0.00
D	1.00	1.00	0.99	NA

Table 8.26: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Track Construction (Data B, Usage)

Track Construction	A	B	C	D
A	NA	0.80	1.00	1.00
B	0.00	NA	0.97	1.00
C	0.00	0.00	NA	1.00
D	0.00	0.00	0.00	NA

Table 8.27: Mean 1 Tailed K-S and Mann Whitney U Test for Track Construction (Data B, Usage)

8.2.4 Track Category

Data Amounts

Track Category		1A	1	2	3	4	5	6
Data A Usage	Datasets Used	3	5	4	4	2	2	0
	Data points Used	341	1718	2160	1352	857	242	0
Data B Usage	Datasets Used	12	23	27	25	22	20	1
	Data points Used	1258	3414	4326	6945	4674	2575	26

Table 8.28: Data Amounts for Track Category

2-Tailed Hypothesis Tests

Track Category	1A	1	2	3	4	5
1A	1.00	0.00	0.00	0.00	0.00	0.00
1	0.00	1.00	0.18	0.02	0.00	0.00
2	0.00	0.18	1.00	0.10	0.00	0.00
3	0.00	0.02	0.10	1.00	0.00	0.00
4	0.00	0.00	0.00	0.00	1.00	0.01
5	0.00	0.00	0.00	0.00	0.01	1.00

Table 8.29: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Track Category (Data A, Usage)

Track Category	1A	1	2	3	4	5
1A	0.00	1.00	1.00	1.00	1.00	1.00
1	1.00	0.00	0.50	0.88	1.00	1.00
2	1.00	0.50	0.00	0.50	1.00	1.00
3	1.00	0.88	0.50	0.00	1.00	1.00
4	1.00	1.00	1.00	1.00	0.00	1.00
5	1.00	1.00	1.00	1.00	1.00	0.00

Table 8.30: Mean 2 Tailed K-S and Mann Whitney U Test for Track Category (Data A, Usage)

Track Category	1A	1	2	3	4	5
1A	1.00	0.30	0.00	0.00	0.00	0.00
1	0.30	1.00	0.08	0.00	0.00	0.00
2	0.00	0.08	1.00	0.06	0.00	0.00
3	0.00	0.00	0.06	1.00	0.05	0.00
4	0.00	0.00	0.00	0.05	1.00	0.01
5	0.00	0.00	0.00	0.00	0.01	1.00

Table 8.31: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Track Category (Data B, Usage)

Track Category	1A	1	2	3	4	5
1A	0.00	0.42	0.96	1.00	1.00	1.00
1	0.42	0.00	0.70	1.00	1.00	1.00
2	0.96	0.70	0.00	0.81	0.98	1.00
3	1.00	1.00	0.81	0.00	0.86	1.00
4	1.00	1.00	0.98	0.86	0.00	0.98
5	1.00	1.00	1.00	1.00	0.98	0.00

Table 8.32: Mean 2 Tailed K-S and Mann Whitney U Test for Track Category (Data B, Usage)

1-Tailed Hypothesis Tests

Track Category	1A	1	2	3	4	5
1A	NA	0.00	0.00	0.00	0.00	0.00
1	0.93	NA	0.09	0.01	0.00	0.00
2	1.00	0.94	NA	0.24	0.00	0.00
3	1.00	0.99	0.78	NA	0.00	0.00
4	1.00	1.00	1.00	1.00	NA	0.00
5	1.00	1.00	1.00	1.00	1.00	NA

Table 8.33: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Track Category (Data A, Usage)

Track Category	1A	1	2	3	4	5
1A	NA	1.00	0.00	1.00	1.00	1.00
1	0.00	NA	0.75	1.00	1.00	1.00
2	0.00	0.00	NA	0.75	1.00	1.00
3	0.00	0.00	0.00	NA	1.00	1.00
4	0.00	0.00	0.00	0.00	NA	1.00
5	0.00	0.00	0.00	0.00	0.00	NA

Table 8.34: Mean 1 Tailed K-S and Mann Whitney U Test for Track Category (Data A, Usage)

Track Category	1A	1	2	3	4	5
1A	NA	0.20	0.00	0.00	0.00	0.00
1	0.74	NA	0.04	0.00	0.00	0.00
2	0.97	0.93	NA	0.04	0.00	0.00
3	1.00	0.98	0.95	NA	0.05	0.00
4	1.00	1.00	0.99	0.95	NA	0.00
5	1.00	1.00	1.00	1.00	0.99	NA

Table 8.35: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Track Category (Data B, Usage)

Track Category	1A	1	2	3	4	5
1A	NA	0.50	1.00	1.00	1.00	1.00
1	0.00	NA	0.78	1.00	1.00	1.00
2	0.00	0.00	NA	0.90	0.98	1.00
3	0.00	0.00	0.00	NA	0.90	1.00
4	0.00	0.00	0.00	0.00	NA	0.98
5	0.00	0.00	0.00	0.00	0.00	NA

Table 8.36: Mean 1 Tailed K-S and Mann Whitney U Test for Track Category (Data B, Usage)

8.2.5 Track

Data Amounts

Track		PL	S&C
Data A Usage	Datasets Used	7	1
	Data points Used	6700	116
Data B Usage	Datasets Used	35	3
	Data points Used	23832	121

Table 8.37: Data Amounts for Track

2-Tailed Hypothesis Tests

Track	PL	S&C
PL	1.00	0.21
S&C	0.21	1.00

Table 8.38: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Track (Data B, Usage)

Track	PL	S&C
PL	0.00	0.33
S&C	0.33	0.00

Table 8.39: Mean 2 Tailed K-S and Mann Whitney U Test for Track (Data B, Usage)

1-Tailed Hypothesis Tests

Track	PL	S&C
PL	NA	0.74
S&C	0.23	NA

Table 8.40: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Track (Data B, Usage)

Track	PL	S&C
PL	NA	0.00
S&C	0.67	NA

Table 8.41: Mean 1 Tailed K-S and Mann Whitney U Test for Track (Data B, Usage)

8.2.6 Route Criticality

Data Amounts

Route Criticality		1	2	3	4	5
Data A Usage	Datasets Used	5	4	4	3	2
	Data points Used	1351	1274	2343	1317	433
Data B Usage	Datasets Used	21	25	26	26	20
	Data points Used	3020	3315	5356	7610	4252

Table 8.42: Data Amounts for Route Criticality

2-Tailed Hypothesis Tests

Route Criticality	1	2	3	4	5
1	1.00	0.39	0.02	0.00	0.00
2	0.39	1.00	0.00	0.00	0.00
3	0.02	0.00	1.00	0.28	0.00
4	0.00	0.00	0.28	1.00	0.00
5	0.00	0.00	0.00	0.00	1.00

Table 8.43: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Route Criticality (Data A, Usage)

Route Criticality	1	2	3	4	5
1	0.00	0.00	0.88	1.00	1.00
2	0.00	0.00	1.00	1.00	1.00
3	0.88	1.00	0.00	0.67	1.00
4	1.00	1.00	0.67	0.00	1.00
5	1.00	1.00	1.00	1.00	0.00

Table 8.44: Mean 2 Tailed K-S and Mann Whitney U Test for Route Criticality (Data A, Usage)

Route Criticality	1	2	3	4	5
1	1.00	0.25	0.14	0.00	0.00
2	0.25	1.00	0.10	0.03	0.00
3	0.14	0.10	1.00	0.10	0.00
4	0.00	0.03	0.10	1.00	0.03
5	0.00	0.00	0.00	0.03	1.00

Table 8.45: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Route Criticality (Data B, Usage)

Route Criticality	1	2	3	4	5
1	0.00	0.21	0.68	1.00	1.00
2	0.21	0.00	0.65	0.91	1.00
3	0.68	0.65	0.00	0.76	1.00
4	1.00	0.91	0.76	0.00	0.80
5	1.00	1.00	1.00	0.80	0.00

Table 8.46: Mean 2 Tailed K-S and Mann Whitney U Test for Route Criticality (Data B, Usage)

1-Tailed Hypothesis Tests

Route Criticality	1	2	3	4	5
1	NA	0.51	0.01	0.00	0.00
2	0.49	NA	0.00	0.00	0.00
3	0.94	0.99	NA	0.20	0.00
4	1.00	1.00	0.81	NA	0.00
5	1.00	1.00	1.00	0.98	NA

Table 8.47: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Route Criticality (Data A, Usage)

Route Criticality	1	2	3	4	5
1	NA	0.00	0.88	1.00	1.00
2	0.00	NA	1.00	1.00	1.00
3	0.00	0.00	NA	0.67	1.00
4	0.00	0.00	0.00	NA	1.00
5	0.00	0.00	0.00	0.00	NA

Table 8.48: Mean 1 Tailed K-S and Mann Whitney U Test for Route Criticality (Data A, Usage)

Route Criticality	1	2	3	4	5
1	NA	0.30	0.09	0.00	0.00
2	0.67	NA	0.07	0.01	0.00
3	0.94	0.93	NA	0.10	0.00
4	0.99	0.99	0.88	NA	0.02
5	1.00	1.00	1.00	0.98	NA

Table 8.49: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Route Criticality (Data B, Usage)

Route Criticality	1	2	3	4	5
1	NA	0.31	0.63	1.00	1.00
2	0.07	NA	0.73	0.91	1.00
3	0.00	0.00	NA	0.84	1.00
4	0.00	0.00	0.02	NA	0.85
5	0.00	0.00	0.00	0.00	NA

Table 8.50: Mean 1 Tailed K-S and Mann Whitney U Test for Route Criticality (Data B, Usage)

8.2.7 Embankments, Soil Cuttings and Rock Cuttings

Data Amounts

Geotechnical Infrastructure		None	Embankments	Soil Cuttings	Rock Cuttings
Data A Usage	Datasets Used	5	4	4	1
	Data points Used	2409	2083	1581	51
Data B Usage	Datasets Used	27	28	26	3
	Data points Used	7823	8206	5187	86

Table 8.51: Data Amounts for Embankments, Soil Cuttings and Rock Cuttings

2-Tailed Hypothesis Tests

Geotechnical Infrastructure	None	Embankments	Soil Cuttings
None	1.00	0.21	0.34
Embankment	0.21	1.00	0.12
Soil Cutting	0.34	0.12	1.00

Table 8.52: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Embankments, Soil Cuttings and Rock Cuttings (Data A, Usage)

Geotechnical Infrastructure	None	Embankments	Soil Cuttings
None	0.00	0.38	0.50
Embankment	0.38	0.00	0.13
Soil Cutting	0.50	0.13	0.00

Table 8.53: Mean 2 Tailed K-S and Mann Whitney U Test for Embankments, Soil Cuttings and Rock Cuttings (Data A, Usage)

Geotechnical Infrastructure	None	Embankments	Soil Cuttings	Rock Cuttings
None	1.00	0.39	0.42	0.62
Embankment	0.39	1.00	0.38	0.53
Soil Cutting	0.42	0.38	1.00	0.44
Rock cutting	0.62	0.53	0.44	1.00

Table 8.54: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Embankments, Soil Cuttings and Rock Cuttings (Data B, Usage)

Geotechnical Infrastructure	None	Embankments	Soil Cuttings	Rock Cuttings
None	0.00	0.19	0.26	0.00
Embankment	0.19	0.00	0.19	0.00
Soil Cutting	0.26	0.19	0.00	0.00
Rock cutting	0.00	0.00	0.00	0.00

Table 8.55: Mean 2 Tailed K-S and Mann Whitney U Test for Embankments, Soil Cuttings and Rock Cuttings (Data B, Usage)

1-Tailed Hypothesis Tests

Geotechnical Infrastructure	None	Em-bank-ments	Soil Cut-tings
None	NA	0.54	0.25
Embankment	0.50	NA	0.06
Soil Cutting	0.70	0.89	NA

Table 8.56: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Embankments, Soil Cuttings and Rock Cuttings (Data A, Usage)

Geotechnical Infrastructure	None	Em-bank-ments	Soil Cut-tings
None	NA	0.25	0.50
Embankment	0.13	NA	0.38
Soil Cutting	0.00	0.00	NA

Table 8.57: Mean 1 Tailed K-S and Mann Whitney U Test for Embankments, Soil Cuttings and Rock Cuttings (Data A, Usage)

Geotechnical Infrastructure	None	Em-bank-ments	Soil Cut-tings	Rock Cut-tings
None	NA	0.48	0.51	0.33
Embankment	0.47	NA	0.55	0.56
Soil Cutting	0.47	0.42	NA	0.47
Rock cutting	0.72	0.48	0.55	NA

Table 8.58: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Embankments, Soil Cuttings and Rock Cuttings (Data B, Usage)

Geotechnical Infrastructure	None	Embankments	Soil Cuttings	Rock Cuttings
None	NA	0.09	0.14	0.00
Embankment	0.13	NA	0.08	0.00
Soil Cutting	0.16	0.19	NA	0.00
Rock cutting	0.00	0.17	0.17	NA

Table 8.59: Mean 1 Tailed K-S and Mann Whitney U Test for Embankments, Soil Cuttings and Rock Cuttings (Data B, Usage)

8.2.8 Curvature

Data Amounts

Curvature $[m^{-1}]$		0	0-0.0005	0.0005-0.001	0.001-0.002
Data A Usage	Datasets Used	5	4	3	1
	Data points Used	2755	756	568	71
Data B Usage	Datasets Used	29	20	20	8
	Data points Used	11211	2078	1442	245

Table 8.60: Data Amounts for Curvature

2-Tailed Hypothesis Tests

Curvature $[m^{-1}]$	0	0-0.0005	0.0005-0.001
0	1.00	0.39	0.37
0-0.0005	0.39	1.00	0.44
0.001-0.002	0.37	0.44	1.00

Table 8.61: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Curvature (Data A, Usage)

Curvature $[m^{-1}]$	0	0-0.0005	0.0005-0.001
0	0.00	0.50	0.17
0-0.0005	0.50	0.00	0.17
0.001-0.002	0.17	0.17	0.00

Table 8.62: Mean 2 Tailed K-S and Mann Whitney U Test for Curvature (Data A, Usage)

Curvature $[m^{-1}]$	0	0-0.0005	0.0005-0.001	0.001-0.002
0	1.00	0.31	0.38	0.32
0-0.0005	0.31	1.00	0.31	0.10
0.0005-0.001	0.38	0.31	1.00	0.27
0.001-0.002	0.32	0.10	0.27	1.00

Table 8.63: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Curvature (Data B, Usage)

Curvature $[m^{-1}]$	0	0-0.0005	0.0005-0.001	0.001-0.002
0	0.00	0.23	0.08	0.44
0-0.0005	0.23	0.00	0.15	0.50
0.0005-0.001	0.08	0.15	0.00	0.31
0.001-0.002	0.44	0.50	0.31	0.00

Table 8.64: Mean 2 Tailed K-S and Mann Whitney U Test for Curvature (Data B, Usage)

1-Tailed Hypothesis Tests

Curvature $[m^{-1}]$	0	0-0.0005	0.0005-0.001
0	NA	0.81	0.68
0-0.0005	0.21	NA	0.39
0.001-0.002	0.18	0.50	NA

Table 8.65: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Curvature (Data A, Usage)

Curvature $[m^{-1}]$	0	0-0.0005	0.0005-0.001
0	NA	0.00	0.00
0-0.0005	0.50	NA	0.17
0.001-0.002	0.33	0.00	NA

Table 8.66: Mean 1 Tailed K-S and Mann Whitney U Test for Curvature (Data A, Usage)

Curvature $[m^{-1}]$	0	0-0.0005	0.0005-0.001	0.001-0.002
0	NA	0.77	0.39	0.22
0-0.0005	0.21	NA	0.22	0.05
0.0005-0.001	0.55	0.77	NA	0.20
0.001-0.002	0.73	0.89	0.81	NA

Table 8.67: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Curvature (Data B, Usage)

Curvature [m^{-1}]	0	0-0.0005	0.0005-0.001	0.001-0.002
0	NA	0.00	0.10	0.50
0-0.0005	0.28	NA	0.20	0.69
0.0005-0.001	0.03	0.00	NA	0.44
0.001-0.002	0.00	0.00	0.00	NA

Table 8.68: Mean 1 Tailed K-S and Mann Whitney U Test for Curvature (Data B, Usage)

8.2.9 Cant

Data Amounts

Cant [cm]		0	0-40	40-80	80-210
Data A Usage	Datasets Used	5	2	3	2
	Data points Used	2747	353	626	402
Data B Usage	Datasets Used	29	19	20	17
	Data points Used	11201	1126	1841	913

Table 8.69: Data Amounts for Cant

2-Tailed Hypothesis Tests

Cant [cm]	0	0-40	40-80	80-210
0	1.00	0.54	0.51	0.04
0-40	0.54	1.00	0.24	0.02
40-80	0.51	0.24	1.00	0.11
80-210	0.04	0.02	0.11	1.00

Table 8.70: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Cant (Data A, Usage)

Cant [cm]	0	0-40	40-80	80-210
0	0.00	0.00	0.33	0.75
0-40	0.00	0.00	0.50	1.00
40-80	0.33	0.50	0.00	0.00
80-210	0.75	1.00	0.00	0.00

Table 8.71: Mean 2 Tailed K-S and Mann Whitney U Test for Cant (Data A, Usage)

Cant [cm]	0	0-40	40-80	80-210
0	1.00	0.48	0.36	0.40
0-40	0.48	1.00	0.41	0.45
40-80	0.36	0.41	1.00	0.44
80-210	0.40	0.45	0.44	1.00

Table 8.72: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Cant (Data B, Usage)

Cant [cm]	0	0-40	40-80	80-210
0	0.00	0.05	0.15	0.18
0-40	0.05	0.00	0.08	0.13
40-80	0.15	0.08	0.00	0.18
80-210	0.18	0.13	0.13	0.00

Table 8.73: Mean 2 Tailed K-S and Mann Whitney U Test for Cant (Data B, Usage)

1-Tailed Hypothesis Tests

Cant [cm]	0	0-40	40-80	80-210
0	NA	0.39	0.77	0.82
0-40	0.55	NA	0.87	1.00
40-80	0.28	0.12	NA	0.81
80-210	0.02	0.01	0.05	NA

Table 8.74: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Cant (Data A, Usage)

Cant [cm]	0	0-40	40-80	80-210
0	NA	0.00	0.00	0.00
0-40	0.00	NA	0.00	0.00
40-80	0.33	0.50	NA	0.00
80-210	0.75	1.00	0.75	NA

Table 8.75: Mean 1 Tailed K-S and Mann Whitney U Test for Cant (Data A, Usage)

Cant [cm]	0	0-40	40-80	80-210
0	NA	0.50	0.54	0.57
0-40	0.52	NA	0.53	0.62
40-80	0.42	0.44	NA	0.57
80-210	0.37	0.35	0.43	NA

Table 8.76: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Cant (Data B, Usage)

Cant [cm]	0	0-40	40-80	80-210
0	NA	0.13	0.05	0.12
0-40	0.00	NA	0.03	0.00
40-80	0.15	0.16	NA	0.06
80-210	0.21	0.19	0.21	NA

Table 8.77: Mean 1 Tailed K-S and Mann Whitney U Test for Cant (Data B, Usage)

8.2.10 Maximum Axle Load

Data Amounts

Maximum Axle Load [Tonnes]		0-22	23-25	26
Data A Usage	Datasets Used	3	4	5
	Data points Used	792	2046	3874
Data B Usage	Datasets Used	20	28	29
	Data points Used	5581	8150	9938

Table 8.78: Data Amounts for Maximum Axle Load

2-Tailed Hypothesis Tests

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	1.00	0.11	0.25
23-25	0.11	1.00	0.19
26	0.25	0.19	1.00

Table 8.79: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Maximum Axle Load (Data A, Usage)

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	0.00	0.00	0.67
23-25	0.00	0.00	0.75
26	0.67	0.75	0.00

Table 8.80: Mean 2 Tailed K-S and Mann Whitney U Test for Maximum Axle Load (Data A, Usage)

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	1.00	0.15	0.07
23-25	0.15	1.00	0.27
26	0.07	0.27	1.00

Table 8.81: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Maximum Axle Load (Data B, Usage)

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	0.00	0.75	0.85
23-25	0.75	0.00	0.45
26	0.85	0.45	0.00

Table 8.82: Mean 2 Tailed K-S and Mann Whitney U Test for Maximum Axle Load (Data B, Usage)

1-Tailed Hypothesis Tests

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	NA	0.28	0.88
23-25	0.66	NA	0.88
26	0.13	0.10	NA

Table 8.83: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Maximum Axle Load (Data A, Usage)

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	NA	0.33	0.00
23-25	0.00	NA	0.00
26	0.67	0.75	NA

Table 8.84: Mean 1 Tailed K-S and Mann Whitney U Test for Maximum Axle Load (Data A, Usage)

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	NA	0.89	0.93
23-25	0.08	NA	0.70
26	0.07	0.27	NA

Table 8.85: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Maximum Axle Load (Data B, Usage)

Maximum Axle Load [Tonnes]	0-22	23-25	26
0-22	NA	0.00	0.00
23-25	0.80	NA	0.04
26	0.90	0.46	NA

Table 8.86: Mean 1 Tailed K-S and Mann Whitney U Test for Maximum Axle Load (Data B, Usage)

8.2.11 Electrification

Data Amounts

Electrification		None	OLE	$3^{rd}/4^{th}$ Rail
Data A Usage	Datasets Used	4	5	2
	Data points Used	4108	2094	529
Data A Usage	Datasets Used	29	28	22
	Data points Used	14063	6083	3621

Table 8.87: Data Amounts for Electrification

2-Tailed Hypothesis Tests

Electrification	None	OLE	$3^{rd}/4^{th}$ Rail
None	1.00	0.00	0.00
OLE	0.00	1.00	0.00
3rd/4th Rail	0.00	0.00	1.00

Table 8.88: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Electrification (Data A, Usage)

Electrification	None	OLE	$3^{rd}/4^{th}$ Rail
None	0.00	1.00	1.00
OLE	1.00	0.00	1.00
3rd/4th Rail	1.00	1.00	0.00

Table 8.89: Mean 2 Tailed K-S and Mann Whitney U Test for Electrification (Data A, Usage)

Electrification	None	OLE	$3^{rd}/4^{th}$ Rail
None	1.00	0.02	0.04
OLE	0.02	1.00	0.18
3rd/4th Rail	0.04	0.18	1.00

Table 8.90: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Electrification (Data B, Usage)

Electrification	None	OLE	$3^{rd}/4^{th}$ Rail
None	0.00	0.93	0.95
OLE	0.93	0.00	0.43
3rd/4th Rail	0.95	0.43	0.00

Table 8.91: Mean 2 Tailed K-S and Mann Whitney U Test for Electrification (Data B, Usage)

1-Tailed Hypothesis Tests

Electrification	None	OLE	$3^{rd}/4^{th}$ Rail
None	NA	1.00	0.99
OLE	0.00	NA	0.00
3rd/4th Rail	0.00	0.98	NA

Table 8.92: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Electrification (Data A, Usage)

Electrification	None	OLE	$3^{rd}/4^{th}$ Rail
None	NA	0.00	0.00
OLE	1.00	NA	1.00
3rd/4th Rail	1.00	0.00	NA

Table 8.93: Mean 1 Tailed K-S and Mann Whitney U Test for Electrification (Data A, Usage)

Electrification	None	OLE	3rd/4th Rail
None	NA	0.98	0.97
OLE	0.01	NA	0.16
3rd/4th Rail	0.02	0.78	NA

Table 8.94: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Electrification (Data B, Usage)

Electrification	None	OLE	3rd/4th Rail
None	NA	0.00	0.00
OLE	0.96	NA	0.57
3rd/4th Rail	0.95	0.00	NA

Table 8.95: Mean 1 Tailed K-S and Mann Whitney U Test for Electrification (Data B, Usage)

8.2.12 Rail Type

Data Amounts

Rail Type		95lb Bull-head	98lb Flat-bot-tom	109lb Flat-bot-tom	110lb Flat-bot-tom	113lb Flat-bot-tom	UIC 60
Data A Usage	Datasets Used	0	0	1	0	4	5
	Data points Used	0	0	27	0	2978	3468
Data B Usage	Datasets Used	7	5	9	18	31	12
	Data points Used	259	154	369	1154	17926	2999

Table 8.96: Data Amounts for Rail Type

2-Tailed Hypothesis Tests

Rail Type	113lb Flat-bot-tom	UIC 60
113lb Flatbottom	1.00	0.22
UIC 60	0.22	1.00

Table 8.97: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Type (Data A, Usage)

Rail Type	113lb Flat-bot-tom	UIC 60
113lb Flatbottom	0.00	0.75
UIC 60	0.75	0.00

Table 8.98: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Type (Data A, Usage)

Rail Type	95lb Bull-head	98lb Flat-bot-tom	109lb Flat-bot-tom	110lb Flat-bot-tom	113lb Flat-bot-tom	UIC 60
95lb Bullhead	1.00	0.06	0.00	0.00	0.00	0.00
98lb Flatbottom	0.06	1.00	0.16	0.14	0.00	0.00
109lb Flatbottom	0.00	0.16	1.00	0.18	0.09	0.00
110lb Flatbottom	0.00	0.14	0.18	1.00	0.22	0.02
113lb Flatbottom	0.00	0.00	0.09	0.22	1.00	0.06
UIC 60	0.00	0.00	0.00	0.02	0.06	1.00

Table 8.99: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Type (Data B, Usage)

Rail Type	95lb Bull-head	98lb Flat-bot-tom	109lb Flat-bot-tom	110lb Flat-bot-tom	113lb Flat-bot-tom	UIC 60
95lb Bullhead	0.00	0.70	1.00	1.00	1.00	1.00
98lb Flatbottom	0.70	0.00	0.30	0.80	1.00	1.00
109lb Flatbottom	1.00	0.30	0.00	0.44	0.78	1.00
110lb Flatbottom	1.00	0.80	0.44	0.00	0.61	0.86
113lb Flatbottom	1.00	1.00	0.78	0.61	0.00	0.83
UIC 60	1.00	1.00	1.00	0.86	0.83	0.00

Table 8.100: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Type (Data B, Usage)

1-Tailed Hypothesis Tests

Rail Type	113lb Flat-bot-tom	UIC 60
113lb Flatbottom	NA	0.86
UIC 60	0.12	NA

Table 8.101: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Rail Type (Data A, Usage)

Rail Type	113lb Flat-bot-tom	UIC 60
113lb Flatbottom	NA	0.00
UIC 60	0.75	NA

Table 8.102: Mean 1 Tailed K-S and Mann Whitney U Test for Rail Type (Data A, Usage)

Rail Type	95lb Bull-head	98lb Flat-bot-tom	109lb Flat-bot-tom	110lb Flat-bot-tom	113lb Flat-bot-tom	UIC 60
95lb Bullhead	NA	0.98	1.00	1.00	1.00	1.00
98lb Flatbottom	0.03	NA	0.75	0.87	0.98	1.00
109lb Flatbottom	0.00	0.23	NA	0.69	0.93	0.99
110lb Flatbottom	0.00	0.07	0.28	NA	0.87	0.99
113lb Flatbottom	0.00	0.00	0.04	0.13	NA	0.97
UIC 60	0.00	0.00	0.00	0.01	0.03	NA

Table 8.103: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Rail Type (Data B, Usage)

Rail Type	95lb Bull-head	98lb Flat-bot-tom	109lb Flat-bot-tom	110lb Flat-bot-tom	113lb Flat-bot-tom	UIC 60
95lb Bullhead	NA	0.00	0.00	0.00	0.00	0.00
98lb Flatbottom	0.80	NA	0.10	0.00	0.00	0.00
109lb Flatbottom	1.00	0.40	NA	0.11	0.00	0.00
110lb Flatbottom	1.00	0.80	0.50	NA	0.00	0.00
113lb Flatbottom	1.00	1.00	0.78	0.69	NA	0.00
UIC 60	1.00	1.00	1.00	0.86	0.88	NA

Table 8.104: Mean 1 Tailed K-S and Mann Whitney U Test for Rail Type (Data B, Usage)

8.2.13 Rail Types Reduced Groups

Second Layer Analysis

Data Amounts

Rail Type Group		Group 1	Group 2	Group 3
Data A Usage	Datasets Used	0	9	10
	Data points Used	0	2884	3374
Data B Usage	Datasets Used	8	86	19
	Data points Used	346	18958	2807
Group 1	95lb Bullhead, 98lb Flatbottom			
Group 2	109, 110lb, 113lb Flatbottom			
Group 3	UIC 60			

Table 8.105: Data Amounts for Rail Types Grouped

2-Tailed Hypothesis Tests

Rail Type Group		Group 2	Group 3
Group 2		1.00	0.20
Group 3		0.20	1.00
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.106: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Types Grouped (Data A, Usage)

Rail Type Group	Group 2		Group 3
Group 2	0.00		0.57
Group 3	0.57		0.00
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.107: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Types Grouped (Data A, Usage)

Rail Type Group	Group 1	Group 2	Group 3
Group 1	1.00	0.00	NaN
Group 2	0.00	1.00	0.13
Group 3	NaN	0.13	1.00
Group 1	95lb Bullhead, 98lb Flatbottom		
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.108: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Types Grouped (Data B, Usage)

Rail Type Group	Group 1	Group 2	Group 3
Group 1	0.00	1.00	NaN
Group 2	1.00	0.00	0.69
Group 3	NaN	0.69	0.00
Group 1	95lb Bullhead, 98lb Flatbottom		
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.109: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Types Grouped (Data B, Usage)

1-Tailed Hypothesis Tests

Rail Type Group	Group 2	Group 3
Group 2	NA	0.79
Group 3	0.19	NA
Group 2	109, 110lb, 113lb Flatbottom	
Group 3	UIC 60	

Table 8.110: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Types Grouped (Data A, Usage)

Rail Type Group	Group 2	Group 3
Group 2	NA	0.00
Group 3	0.57	NA
Group 2	109, 110lb, 113lb Flatbottom	
Group 3	UIC 60	

Table 8.111: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Types Grouped (Data A, Usage)

Rail Type Group	Group 1	Group 2	Group 3
Group 1	NA	1.00	NaN
Group 2	0.00	NA	0.70
Group 3	NaN	0.29	NA
Group 1	95lb Bullhead, 98lb Flatbottom		
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.112: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Types Grouped (Data B, Usage)

Rail Type Group	Group 1	Group 2	Group 3
Group 1	NA	0.00	NaN
Group 2	1.00	NA	0.22
Group 3	NaN	0.56	NA
Group 1	95lb Bullhead, 98lb Flatbottom		
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.113: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Types Grouped (Data B, Usage)

Third Layer Analysis

Data Amounts

Rail Type Group		Group 1	Group 2	Group 3
Data A Usage	Datasets Used	0	18	13
	Data points Used	0	2839	3299
Data B Usage	Datasets Used	7	151	23
	Data points Used	222	17904	2646
Group 1	95lb Bullhead, 98lb Flatbottom			
Group 2	109, 110lb, 113lb Flatbottom			
Group 3	UIC 60			

Table 8.114: Data Amounts for Rail Types Grouped

2-Tailed Hypothesis Tests

Rail Type Group		Group 2	Group 3
Group 2		1.00	0.27
Group 3		0.27	1.00
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.115: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Types Grouped (Data A, Usage)

Rail Type Group		Group 2	Group 3
Group 2		0.00	0.39
Group 3		0.39	0.00
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.116: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Types Grouped (Data A, Usage)

Rail Type Group		Group 1	Group 2	Group 3
Group 1		1.00	0.31	NaN
Group 2		0.31	1.00	0.22
Group 3		NaN	0.22	1.00
Group 1	95lb Bullhead, 98lb Flatbottom			
Group 2	109, 110lb, 113lb Flatbottom			
Group 3	UIC 60			

Table 8.117: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Rail Types Grouped (Data B, Usage)

Rail Type Group	Group 1	Group 2	Group 3
Group 1	0.00	0.30	NaN
Group 2	0.30	0.00	0.60
Group 3	NaN	0.60	0.00
Group 1	95lb Bullhead, 98lb Flatbottom		
Group 2	109, 110lb, 113lb Flatbottom		
Group 3	UIC 60		

Table 8.118: Mean 2 Tailed K-S and Mann Whitney U Test for Rail Types Grouped (Data B, Usage)

8.2.14 Passenger % Usage

Data Amounts

Passenger % Usage		0-20	20-40	40-60	60-80	80-100
Data A Usage	Datasets Used	4	2	4	4	4
	Data points Used	1016	605	1258	1089	2599
Data B Usage	Datasets Used	20	22	24	24	30
	Data points Used	2622	1724	2712	3429	12856

Table 8.119: Data Amounts for Passenger pcnt Usage

2-Tailed Hypothesis Tests

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	1.00	0.00	0.00	0.01	0.01
20-40	0.00	1.00	0.04	0.41	0.50
40-60	0.00	0.04	1.00	0.32	0.20
60-80	0.01	0.41	0.32	1.00	0.55
80-100	0.01	0.50	0.20	0.55	1.00

Table 8.120: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Passenger pcnt Usage (Data A, Usage)

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	0.00	1.00	1.00	0.88	1.00
20-40	1.00	0.00	0.75	0.00	0.25
40-60	1.00	0.75	0.00	0.25	0.38
60-80	0.88	0.00	0.25	0.00	0.00
80-100	1.00	0.25	0.38	0.00	0.00

Table 8.121: Mean 2 Tailed K-S and Mann Whitney U Test for Passenger pcnt Usage (Data A, Usage)

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	1.00	0.25	0.17	0.03	0.18
20-40	0.25	1.00	0.30	0.24	0.27
40-60	0.17	0.30	1.00	0.23	0.22
60-80	0.03	0.24	0.23	1.00	0.29
80-100	0.18	0.27	0.22	0.29	1.00

Table 8.122: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Passenger pcnt Usage (Data B, Usage)

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	0.00	0.43	0.55	0.93	0.60
20-40	0.43	0.00	0.25	0.40	0.23
40-60	0.55	0.25	0.00	0.45	0.38
60-80	0.93	0.40	0.45	0.00	0.48
80-100	0.60	0.23	0.38	0.48	0.00

Table 8.123: Mean 2 Tailed K-S and Mann Whitney U Test for Passenger pcnt Usage (Data B, Usage)

1-Tailed Hypothesis Tests

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	1.00	1.00	0.97	0.97
20-40	0.00	NA	0.77	0.41	0.29
40-60	0.00	0.02	NA	0.16	0.10
60-80	0.00	0.25	0.83	NA	0.35
80-100	0.00	0.32	0.85	0.61	NA

Table 8.124: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Passenger pcnt Usage (Data A, Usage)

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	0.00	0.00	0.00	0.00
20-40	1.00	NA	0.00	0.00	0.25
40-60	1.00	0.75	NA	0.50	0.50
60-80	1.00	0.00	0.00	NA	0.13
80-100	1.00	0.00	0.00	0.00	NA

Table 8.125: Mean 1 Tailed K-S and Mann Whitney U Test for Passenger pcnt Usage (Data A, Usage)

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	0.87	0.87	0.99	0.85
20-40	0.13	NA	0.46	0.77	0.41
40-60	0.13	0.46	NA	0.67	0.42
60-80	0.02	0.15	0.29	NA	0.24
80-100	0.09	0.41	0.51	0.73	NA

Table 8.126: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Passenger pcnt Usage (Data B, Usage)

Passenger % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	0.00	0.00	0.00	0.00
20-40	0.48	NA	0.13	0.00	0.14
40-60	0.63	0.23	NA	0.05	0.23
60-80	0.95	0.50	0.48	NA	0.48
80-100	0.65	0.20	0.29	0.04	NA

Table 8.127: Mean 1 Tailed K-S and Mann Whitney U Test for Passenger pcnt Usage (Data B, Usage)

8.2.15 Axle > 50 % Usage

Data Amounts

Axle > 50 % Usage		0-20	20-40	40-60	60-80	80-100
Data A Usage	Datasets Used	4	4	5	4	3
	Data points Used	1124	1466	1874	1292	843
Data B Usage	Datasets Used	22	25	24	24	21
	Data points Used	3750	4353	4470	4931	5680

Table 8.128: Data Amounts for Axle > 50 pcnt Usage

2-Tailed Hypothesis Tests

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	1.00	0.04	0.06	0.08	0.00
20-40	0.04	1.00	0.31	0.32	0.01
40-60	0.06	0.31	1.00	0.11	0.02
60-80	0.08	0.32	0.11	1.00	0.07
80-100	0.00	0.01	0.02	0.07	1.00

Table 8.129: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Axle > 50 pcnt Usage (Data A, Usage)

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	0.00	0.75	0.75	0.75	1.00
20-40	0.75	0.00	0.25	0.25	1.00
40-60	0.75	0.25	0.00	0.50	0.83
60-80	0.75	0.25	0.50	0.00	0.67
80-100	1.00	1.00	0.83	0.67	0.00

Table 8.130: Mean 2 Tailed K-S and Mann Whitney U Test for Axle > 50 pcnt Usage (Data A, Usage)

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	1.00	0.32	0.16	0.20	0.12
20-40	0.32	1.00	0.08	0.16	0.11
40-60	0.16	0.08	1.00	0.33	0.04
60-80	0.20	0.16	0.33	1.00	0.05
80-100	0.12	0.11	0.04	0.05	1.00

Table 8.131: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Axle > 50 pcnt Usage (Data B, Usage)

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	0.00	0.24	0.67	0.62	0.55
20-40	0.24	0.00	0.66	0.60	0.70
40-60	0.67	0.66	0.00	0.25	0.81
60-80	0.62	0.60	0.25	0.00	0.81
80-100	0.55	0.70	0.81	0.81	0.00

Table 8.132: Mean 2 Tailed K-S and Mann Whitney U Test for Axle > 50 pcnt Usage (Data B, Usage)

1-Tailed Hypothesis Tests

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	0.27	0.41	0.29	0.00
20-40	0.64	NA	0.62	0.40	0.00
40-60	0.37	0.33	NA	0.37	0.01
60-80	0.73	0.38	0.49	NA	0.03
80-100	0.97	0.80	0.90	0.87	NA

Table 8.133: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Axle > 50 pcnt Usage (Data A, Usage)

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	0.63	0.38	0.50	1.00
20-40	0.25	NA	0.00	0.25	1.00
40-60	0.50	0.38	NA	0.50	1.00
60-80	0.25	0.25	0.25	NA	0.67
80-100	0.00	0.00	0.00	0.00	NA

Table 8.134: Mean 1 Tailed K-S and Mann Whitney U Test for Axle > 50 pcnt Usage (Data A, Usage)

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	0.47	0.77	0.68	0.10
20-40	0.48	NA	0.89	0.74	0.10
40-60	0.22	0.04	NA	0.31	0.02
60-80	0.32	0.18	0.68	NA	0.07
80-100	0.84	0.78	0.98	0.91	NA

Table 8.135: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Axle > 50 pcnt Usage (Data B, Usage)

Axle > 50 % Usage	0-20	20-40	40-60	60-80	80-100
0-20	NA	0.19	0.14	0.14	0.60
20-40	0.12	NA	0.00	0.05	0.70
40-60	0.55	0.80	NA	0.32	0.93
60-80	0.50	0.64	0.09	NA	0.83
80-100	0.03	0.03	0.00	0.05	NA

Table 8.136: Mean 1 Tailed K-S and Mann Whitney U Test for Axle > 50 pcnt Usage (Data B, Usage)

8.2.16 Dirty % Usage

Data Usage

Dirty % Usage		0-10	10-20	20-30	30-40
Data A Usage	Datasets Used	7	2	2	2
	Data points Used	5535	673	197	226
Data B Usage	Datasets Used	34	20	9	4
	Data points Used	21354	1558	284	244

Table 8.137: Data Usage for Dirty pcnt Usage

2-Tailed Hypothesis Tests

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	1.00	0.00	0.10	0.00
10-20	0.00	1.00	0.08	0.01
20-30	0.10	0.08	1.00	0.00
30-40	0.00	0.01	0.00	1.00

Table 8.138: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Dirty pcnt Usage (Data A, Usage)

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	0.00	1.00	0.50	1.00
10-20	1.00	0.00	0.25	1.00
20-30	0.50	0.25	0.00	1.00
30-40	1.00	1.00	1.00	0.00

Table 8.139: Mean 2 Tailed K-S and Mann Whitney U Test for Dirty pcnt Usage (Data A, Usage)

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	1.00	0.33	0.37	0.20
10-20	0.33	1.00	0.36	0.45
20-30	0.37	0.36	1.00	0.20
30-40	0.20	0.45	0.20	1.00

Table 8.140: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Dirty pcnt Usage (Data B, Usage)

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	0.00	0.35	0.33	0.75
10-20	0.35	0.00	0.17	0.00
20-30	0.33	0.17	0.00	0.50
30-40	0.75	0.00	0.50	0.00

Table 8.141: Mean 2 Tailed K-S and Mann Whitney U Test for Dirty pcnt Usage (Data B, Usage)

1-Tailed Hypothesis Tests

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	NA	0.00	0.05	0.00
10-20	0.99	NA	0.92	0.00
20-30	0.81	0.04	NA	0.00
30-40	0.99	0.93	0.99	NA

Table 8.142: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Dirty pcnt Usage (Data A, Usage)

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	NA	1.00	0.50	1.00
10-20	0.00	NA	0.00	1.00
20-30	0.00	0.75	NA	1.00
30-40	0.00	0.00	0.00	NA

Table 8.143: Mean 1 Tailed K-S and Mann Whitney U Test for Dirty pcnt Usage (Data A, Usage)

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	NA	0.34	0.31	0.15
10-20	0.53	NA	0.53	0.23
20-30	0.64	0.49	NA	0.10
30-40	0.64	0.59	0.82	NA

Table 8.144: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Dirty pcnt Usage (Data B, Usage)

Dirty % Usage	0-10	10-20	20-30	30-40
0-10	NA	0.30	0.33	0.75
10-20	0.10	NA	0.22	0.13
20-30	0.11	0.11	NA	0.50
30-40	0.00	0.13	0.00	NA

Table 8.145: Mean 1 Tailed K-S and Mann Whitney U Test for Dirty pcnt Usage (Data B, Usage)

8.2.17 Superficial Geology

First Layer Analysis

Data Amounts

Superficial Geology		None	Clay	Diamic-ton	Gravel	Peat	Sand
Data A Usage	Datasets Used	4	4	4	1	1	4
	Data points Used	2306	1136	1009	32	34	737
Data B Usage	Datasets Used	28	20	21	4	6	20
	Data points Used	7276	3654	3551	90	157	3032

Table 8.146: Data Amounts for Superficial Geology

2-Tailed Hypothesis Tests

Superficial Geology	None	Clay	Diamic-ton	Sand
None	1.00	0.05	0.29	0.23
Clay	0.05	1.00	0.02	0.16
Diamicton	0.29	0.02	1.00	0.08
Sand	0.23	0.16	0.08	1.00

Table 8.147: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	0.00	0.63	0.25	0.75
Clay	0.63	0.00	0.88	0.13
Diamicton	0.25	0.88	0.00	0.63
Sand	0.75	0.13	0.63	0.00

Table 8.148: Mean 2 Tailed K-S and Mann Whitney U Test for Superficial Geology (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Gravel	Peat	Sand
None	1.00	0.36	0.16	0.55	0.17	0.30
Clay	0.36	1.00	0.20	0.35	0.17	0.33
Diamicton	0.16	0.20	1.00	0.16	0.17	0.15
Gravel	0.55	0.35	0.16	1.00	0.07	0.40
Peat	0.17	0.17	0.17	0.07	1.00	0.12
Sand	0.30	0.33	0.15	0.40	0.12	1.00

Table 8.149: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Data B, Usage)

Superficial Geology	None	Clay	Diamic-ton	Gravel	Peat	Sand
None	0.00	0.33	0.50	0.25	0.67	0.33
Clay	0.33	0.00	0.45	0.13	0.50	0.35
Diamicton	0.50	0.45	0.00	0.13	0.42	0.70
Gravel	0.25	0.13	0.13	0.00	0.50	0.25
Peat	0.67	0.50	0.42	0.50	0.00	0.58
Sand	0.33	0.35	0.70	0.25	0.58	0.00

Table 8.150: Mean 2 Tailed K-S and Mann Whitney U Test for Superficial Geology (Data B, Usage)

1-Tailed Hypothesis Tests

Superficial Geology	None	Clay	Diamic-ton	Sand
None	NA	0.98	0.17	0.89
Clay	0.02	NA	0.01	0.70
Diamicton	0.72	0.95	NA	0.96
Sand	0.14	0.30	0.04	NA

Table 8.151: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	NA	0.00	0.25	0.00
Clay	0.88	NA	1.00	0.13
Diamicton	0.00	0.00	NA	0.00
Sand	0.75	0.25	0.75	NA

Table 8.152: Mean 1 Tailed K-S and Mann Whitney U Test for Superficial Geology (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Gravel	Peat	Sand
None	NA	0.37	0.09	0.47	0.08	0.68
Clay	0.56	NA	0.14	0.60	0.16	0.66
Diamicton	0.86	0.81	NA	0.74	0.20	0.87
Gravel	0.53	0.40	0.27	NA	0.03	0.83
Peat	0.87	0.78	0.70	0.98	NA	0.86
Sand	0.29	0.29	0.11	0.21	0.06	NA

Table 8.153: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Data B, Usage)

Superficial Geology	None	Clay	Diamic-ton	Gravel	Peat	Sand
None	NA	0.28	0.57	0.25	0.75	0.00
Clay	0.13	NA	0.53	0.25	0.58	0.00
Diamicton	0.02	0.03	NA	0.00	0.58	0.00
Gravel	0.00	0.00	0.50	NA	0.75	0.00
Peat	0.00	0.00	0.08	0.00	NA	0.00
Sand	0.43	0.40	0.78	0.38	0.67	NA

Table 8.154: Mean 1 Tailed K-S and Mann Whitney U Test for Superficial Geology (Data B, Usage)

Third Layer Analysis

Data Amounts

Superficial Geology		None	Clay	Diamic-ton	Sand
Data A Usage	Datasets Used	17	12	9	10
	Data points Used	1409	660	499	386
Data B Usage	Datasets Used	83	44	32	28
	Data points Used	3045	1340	981	844

Table 8.155: Data Amounts for Superficial Geology (Bedrock, Speed)

2-Tailed Hypothesis Tests

Superficial Geology	None	Clay	Diamic-ton	Sand
None	1.00	0.28	0.21	0.29
Clay	0.28	1.00	0.27	0.21
Diamicton	0.21	0.27	1.00	0.27
Sand	0.29	0.21	0.27	1.00

Table 8.156: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Bedrock, Speed) (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	0.00	0.40	0.56	0.39
Clay	0.40	0.00	0.31	0.45
Diamicton	0.56	0.31	0.00	0.44
Sand	0.39	0.45	0.44	0.00

Table 8.157: Mean 2 Tailed K-S and Mann Whitney U Test for Superficial Geology (Bedrock, Speed) (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	1.00	0.37	0.29	0.25
Clay	0.37	1.00	0.32	0.28
Diamicton	0.29	0.32	1.00	0.20
Sand	0.25	0.28	0.20	1.00

Table 8.158: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Bedrock, Speed) (Data B, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	0.00	0.24	0.35	0.41
Clay	0.24	0.00	0.26	0.36
Diamicton	0.35	0.26	0.00	0.53
Sand	0.41	0.36	0.53	0.00

Table 8.159: Mean 2 Tailed K-S and Mann Whitney U Test for Superficial Geology (Bedrock, Speed) (Data B, Usage)

1-Tailed Hypothesis Tests

Superficial Geology	None	Clay	Diamic-ton	Sand
None	NaN	0.60	0.39	0.67
Clay	0.34	NaN	0.36	0.66
Diamicton	0.50	0.59	NaN	0.73
Sand	0.27	0.31	0.25	NaN

Table 8.160: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Bedrock, Speed) (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	NaN	0.20	0.38	0.11
Clay	0.25	NaN	0.06	0.10
Diamicton	0.19	0.25	NaN	0.06
Sand	0.28	0.40	0.44	NaN

Table 8.161: Mean 1 Tailed K-S and Mann Whitney U Test for Superficial Geology (Bedrock, Speed) (Data A, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	NaN	0.41	0.47	0.71
Clay	0.54	NaN	0.41	0.71
Diamicton	0.53	0.52	NaN	0.78
Sand	0.27	0.26	0.19	NaN

Table 8.162: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Superficial Geology (Bedrock, Speed) (Data B, Usage)

Superficial Geology	None	Clay	Diamic-ton	Sand
None	NaN	0.17	0.31	0.04
Clay	0.17	NaN	0.21	0.00
Diamicton	0.08	0.15	NaN	0.03
Sand	0.43	0.43	0.58	NaN

Table 8.163: Mean 1 Tailed K-S and Mann Whitney U Test for Superficial Geology (Bedrock, Speed) (Data B, Usage)

8.2.18 Artificial Geology

Data Amounts

Artificial Geology		None	Made Ground
Data A Usage	Datasets Used	7	2
	Data points Used	6280	269
Data B Usage	Datasets Used	34	15
	Data points Used	21888	868

Table 8.164: Data Amounts for Artificial Geology

2-Tailed Hypothesis Tests

Artificial Geology	None	Made Ground
None	1.00	0.00
Made Ground	0.00	1.00

Table 8.165: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Artificial Geology (Data A, Usage)

Artificial Geology	None	Made Ground
None	0.00	1.00
Made Ground	1.00	0.00

Table 8.166: Mean 2 Tailed K-S and Mann Whitney U Test for Artificial Geology (Data A, Usage)

Artificial Geology	None	Made Ground
None	1.00	0.21
Made Ground	0.21	1.00

Table 8.167: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Artificial Geology (Data B, Usage)

Artificial Geology	None	Made Ground
None	0.00	0.27
Made Ground	0.27	0.00

Table 8.168: Mean 2 Tailed K-S and Mann Whitney U Test for Artificial Geology (Data B, Usage)

1-Tailed Hypothesis Tests

Artificial Geology	None	Made Ground
None	NA	0.98
Made Ground	0.00	NA

Table 8.169: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Artificial Geology (Data A, Usage)

Artificial Geology	None	Made Ground
None	NA	0.00
Made Ground	1.00	NA

Table 8.170: Mean 1 Tailed K-S and Mann Whitney U Test for Artificial Geology (Data A, Usage)

Artificial Geology	None	Made Ground
None	NA	0.70
Made Ground	0.29	NA

Table 8.171: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Artificial Geology (Data B, Usage)

Artificial Geology	None	Made Ground
None	NA	0.07
Made Ground	0.27	NA

Table 8.172: Mean 1 Tailed K-S and Mann Whitney U Test for Artificial Geology (Data B, Usage)

8.2.19 Bedrock Geology Grouped

Data Amounts

Bedrock Geology Group		Group 1	Group 2	Group 3	Group 4
Data A Usage	Datasets Used	2	4	1	5
	Data points Used	538	1481	81	3452
Data B Usage	Datasets Used	19	22	12	30
	Data points Used	1906	5328	571	11966
Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone				
Group 2	Sand, Chalk and Clay				
Group 3	Halite and Pebbly Sand				
Group 4	Mudstone, Sandstone and Slate				

Table 8.173: Data Amounts for Bedrock Geology Grouped

2-Tailed Hypothesis Tests

Bedrock Geology	Group 1	Group 2	Group 4
Group 1	1.00	0.00	0.00
Group 2	0.00	1.00	0.07
Group 4	0.00	0.07	1.00

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 4	Mudstone, Sandstone and Slate

Table 8.174: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Grouped Bedrock Geology (Data A, Usage)

Bedrock Geology	Group 1	Group 2	Group 4
Group 1	0.00	1.00	1.00
Group 2	1.00	0.00	0.75
Group 4	1.00	0.75	0.00

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 4	Mudstone, Sandstone and Slate

Table 8.175: Mean 2 Tailed K-S and Mann Whitney U Test for Grouped Bedrock Geology (Data A, Usage)

Bedrock Geology	Group 1	Group 2	Group 3	Group 4
Group 1	1.00	0.01	0.19	0.20
Group 2	0.01	1.00	0.17	0.05
Group 3	0.19	0.17	1.00	0.27
Group 4	0.20	0.05	0.27	1.00

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 3	Halite and Pebbly Sand
Group 4	Mudstone, Sandstone and Slate

Table 8.176: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Grouped Bedrock Geology (Data B, Usage)

Bedrock Geology	Group 1	Group 2	Group 3	Group 4
Group 1	0.00	0.95	0.55	0.45
Group 2	0.95	0.00	0.42	0.86
Group 3	0.55	0.42	0.00	0.38
Group 4	0.45	0.86	0.38	0.00

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 3	Halite and Pebbly Sand
Group 4	Mudstone, Sandstone and Slate

Table 8.177: Mean 2 Tailed K-S and Mann Whitney U Test for Grouped Bedrock Geology (Data B, Usage)

1-Tailed Hypothesis Tests

Bedrock Geology	Group 1	Group 2	Group 4
Group 1	NA	0.99	0.91
Group 2	0.00	NA	0.03
Group 4	0.00	0.98	NA

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 4	Mudstone, Sandstone and Slate

Table 8.178: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Grouped Bedrock Geology (Data A, Usage)

Bedrock Geology	Group 1	Group 2	Group 4
Group 1	NA	0.00	0.00
Group 2	1.00	NA	0.75
Group 4	1.00	0.00	NA

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 4	Mudstone, Sandstone and Slate

Table 8.179: Mean 2 Tailed K-S and Mann Whitney U Test for Grouped Bedrock Geology (Data A, Usage)

Bedrock Geology	Group 1	Group 2	Group 3	Group 4
Group 1	NA	0.93	0.70	0.80
Group 2	0.04	NA	0.22	0.06
Group 3	0.19	0.73	NA	0.36
Group 4	0.15	0.94	0.58	NA

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 3	Halite and Pebbly Sand
Group 4	Mudstone, Sandstone and Slate

Table 8.180: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Grouped Bedrock Geology (Data B, Usage)

Bedrock Geology	Group 1	Group 2	Group 3	Group 4
Group 1	NA	0.00	0.00	0.00
Group 2	0.95	NA	0.42	0.86
Group 3	0.59	0.08	NA	0.38
Group 4	0.53	0.00	0.13	NA

Group 1	Argillaceous, Dolomitic, Limestone, Psammite and Siltstone
Group 2	Sand, Chalk and Clay
Group 3	Halite and Pebbly Sand
Group 4	Mudstone, Sandstone and Slate

Table 8.181: Mean 2 Tailed K-S and Mann Whitney U Test for Grouped Bedrock Geology (Data B, Usage)

8.2.20 Bedrock Geology

Data Amounts

Bedrock Geology		Argilla-ceous	Chalk	Clay	Dolomitic	Halite	Lime-stone	Mud-stone	Pebbly Sand	Psam-mite	Sand	Sand-stone	Siltstone	Slate
Data A Usage	Datasets Used	1	3	4	1	0	2	5	1	0	1	4	1	1
	Data points Used	38	747	569	22	0	340	2542	67	0	130	879	125	22
Data B Usage	Datasets Used	3	20	19	4	4	12	30	7	5	14	20	14	5
	Data points Used	71	2437	1730	93	95	603	8188	318	151	1014	3374	486	178

Table 8.182: Data Amounts for Bedrock Geology

2-Tailed Hypothesis Tests

Bedrock Geology	Chalk	Clay	Lime-stone	Mud-stone	Sand-stone
Chalk	1.00	0.24	0.00	0.05	0.27
Clay	0.24	1.00	0.00	0.10	0.10
Limestone	0.00	0.00	1.00	0.00	0.01
Mudstone	0.05	0.10	0.00	1.00	0.11
Sandstone	0.27	0.10	0.01	0.11	1.00

Table 8.183: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Bedrock Geology (Data A, Usage)

Bedrock Geology	Chalk	Clay	Lime-stone	Mud-stone	Sand-stone
Chalk	0.00	0.33	1.00	0.67	0.67
Clay	0.33	0.00	1.00	0.75	0.63
Limestone	1.00	1.00	0.00	1.00	1.00
Mudstone	0.67	0.75	1.00	0.00	0.25
Sandstone	0.67	0.63	1.00	0.25	0.00

Table 8.184: Mean 2 Tailed K-S and Mann Whitney U Test for Bedrock Geology (Data A, Usage)

Bedrock Geology	Argilla-ceous	Chalk	Clay	Dolomitic	Halite	Lime-stone	Mud-stone	Pebbly Sand	Psam-mite	Sand	Sand-stone	Siltstone	Slate
Argillaceous	1.00	0.07	0.09	0.52	0.37	0.32	0.19	0.03	0.52	0.02	0.08	0.56	0.72
Chalk	0.07	1.00	0.39	0.00	0.34	0.09	0.04	0.24	0.00	0.34	0.05	0.08	0.09
Clay	0.09	0.39	1.00	0.00	0.43	0.25	0.12	0.21	0.01	0.30	0.17	0.14	0.07
Dolomitic	0.52	0.00	0.00	1.00	0.16	0.55	0.19	0.09	0.37	0.03	0.07	0.27	0.02
Halite	0.37	0.34	0.43	0.16	1.00	0.03	0.16	0.54	0.15	0.18	0.16	0.33	0.30
Limestone	0.32	0.09	0.25	0.55	0.03	1.00	0.23	0.05	0.39	0.11	0.17	0.25	0.30
Mudstone	0.19	0.04	0.12	0.19	0.16	0.23	1.00	0.31	0.15	0.33	0.35	0.35	0.44
Pebbly Sand	0.03	0.24	0.21	0.09	0.54	0.05	0.31	1.00	0.07	0.38	0.37	0.36	0.25
Psammite	0.52	0.00	0.01	0.37	0.15	0.39	0.15	0.07	1.00	0.09	0.03	0.21	0.35
Sand	0.02	0.34	0.30	0.03	0.18	0.11	0.33	0.38	0.09	1.00	0.30	0.22	0.30
Sandstone	0.08	0.05	0.17	0.07	0.16	0.17	0.35	0.37	0.03	0.30	1.00	0.24	0.37
Siltstone	0.56	0.08	0.14	0.27	0.33	0.25	0.35	0.36	0.21	0.22	0.24	1.00	0.50
Slate	0.72	0.09	0.07	0.02	0.30	0.30	0.44	0.25	0.35	0.30	0.37	0.50	1.00

Table 8.185: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Bedrock Geology (Data B, Usage)

Bedrock Geology	Argilla-ceous	Chalk	Clay	Dolomitic	Halite	Lime-stone	Mud-stone	Pebbly Sand	Psam-mite	Sand	Sand-stone	Siltstone	Slate
Argillaceous	0.00	0.67	0.50	0.00	0.00	0.00	0.33	0.75	0.00	1.00	0.67	0.17	0.00
Chalk	0.67	0.00	0.18	1.00	0.00	0.79	0.88	0.50	1.00	0.36	0.85	0.75	0.80
Clay	0.50	0.18	0.00	1.00	0.00	0.54	0.66	0.36	1.00	0.43	0.53	0.69	0.80
Dolomitic	0.00	1.00	1.00	0.00	0.00	0.00	0.38	0.00	0.00	0.75	0.75	0.33	1.00
Halite	0.00	0.00	0.00	0.00	0.00	0.88	0.13	0.00	0.00	0.50	0.13	0.25	0.17
Limestone	0.00	0.79	0.54	0.00	0.88	0.00	0.29	0.58	0.50	0.50	0.46	0.25	0.40
Mudstone	0.33	0.88	0.66	0.38	0.13	0.29	0.00	0.14	0.40	0.43	0.20	0.11	0.00
Pebbly Sand	0.75	0.50	0.36	0.00	0.00	0.58	0.14	0.00	0.50	0.43	0.14	0.29	0.38
Psammite	0.00	1.00	1.00	0.00	0.00	0.50	0.40	0.50	0.00	0.60	0.80	0.30	0.33
Sand	1.00	0.36	0.43	0.75	0.50	0.50	0.43	0.43	0.60	0.00	0.39	0.50	0.40
Sandstone	0.67	0.85	0.53	0.75	0.13	0.46	0.20	0.14	0.80	0.39	0.00	0.39	0.00
Siltstone	0.17	0.75	0.69	0.33	0.25	0.25	0.11	0.29	0.30	0.50	0.39	0.00	0.00
Slate	0.00	0.80	0.80	1.00	0.17	0.40	0.00	0.38	0.33	0.40	0.00	0.00	0.00

Table 8.186: Mean 2 Tailed K-S and Mann Whitney U Test for Bedrock Geology (Data B, Usage)

8.2.21 Maximum Speed

Data Amounts

Maximum Speed [MPH]		0-30	35-40	45-50	55-60	65-70	75-80	85-95	100-110	115-125
Data A Usage	Datasets Used	1	1	1	3	3	4	4	4	4
	Data points Used	64	66	122	946	691	1089	1253	1403	953
Data B Usage	Datasets Used	9	11	19	21	20	24	23	24	19
	Data points Used	364	511	1401	4033	3323	4397	3560	2579	2600

Table 8.187: Data Amounts for Maximum Speed

2-Tailed Hypothesis Tests

Maximum Speed [MPH]	55-60	65-70	75-80	85-95	100-110	115-125
55-60	1.00	0.00	0.00	0.00	0.00	0.00
65-70	0.00	1.00	0.20	0.19	0.31	0.00
75-80	0.00	0.20	1.00	0.30	0.46	0.03
85-95	0.00	0.19	0.30	1.00	0.29	0.05
100-110	0.00	0.31	0.46	0.29	1.00	0.01
115-125	0.00	0.00	0.03	0.05	0.01	1.00

Table 8.188: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Maximum Speed (Data A, Usage)

Maximum Speed [MPH]	55-60	65-70	75-80	85-95	100-110	115-125
55-60	0.00	1.00	1.00	1.00	1.00	1.00
65-70	1.00	0.00	0.33	0.67	0.33	1.00
75-80	1.00	0.33	0.00	0.38	0.25	0.83
85-95	1.00	0.67	0.38	0.00	0.25	0.67
100-110	1.00	0.33	0.25	0.25	0.00	1.00
115-125	1.00	1.00	0.83	0.67	1.00	0.00

Table 8.189: Mean 2 Tailed K-S and Mann Whitney U Test for Maximum Speed (Data A, Usage)

Maximum Speed [MPH]	0-30	35-40	45-50	55-60	65-70	75-80	85-95	100-110	115-125
0-30	1.00	0.35	0.25	0.20	0.01	0.00	0.00	0.00	0.00
35-40	0.35	1.00	0.41	0.26	0.01	0.00	0.00	0.00	0.00
45-50	0.25	0.41	1.00	0.30	0.10	0.06	0.01	0.00	0.00
55-60	0.20	0.26	0.30	1.00	0.07	0.04	0.01	0.00	0.00
65-70	0.01	0.01	0.10	0.07	1.00	0.31	0.09	0.00	0.00
75-80	0.00	0.00	0.06	0.04	0.31	1.00	0.28	0.16	0.01
85-95	0.00	0.00	0.01	0.01	0.09	0.28	1.00	0.38	0.07
100-110	0.00	0.00	0.00	0.00	0.00	0.16	0.38	1.00	0.19
115-125	0.00	0.00	0.00	0.00	0.00	0.01	0.07	0.19	1.00

Table 8.190: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Maximum Speed (Data B, Usage)

Maximum Speed [MPH]	0-30	35-40	45-50	55-60	65-70	75-80	85-95	100-110	115-125
0-30	0.00	0.33	0.33	0.38	0.94	1.00	1.00	1.00	1.00
35-40	0.33	0.00	0.05	0.32	0.91	1.00	1.00	1.00	1.00
45-50	0.33	0.05	0.00	0.45	0.76	0.87	0.89	0.95	1.00
55-60	0.38	0.32	0.45	0.00	0.83	0.85	0.93	1.00	1.00
65-70	0.94	0.91	0.76	0.83	0.00	0.35	0.73	0.95	1.00
75-80	1.00	1.00	0.87	0.85	0.35	0.00	0.40	0.41	0.92
85-95	1.00	1.00	0.89	0.93	0.73	0.40	0.00	0.32	0.82
100-110	1.00	1.00	0.95	1.00	0.95	0.41	0.32	0.00	0.59
115-125	1.00	1.00	1.00	1.00	1.00	0.92	0.82	0.59	0.00

Table 8.191: Mean 2 Tailed K-S and Mann Whitney U Test for Maximum Speed (Data B, Usage)

8.2.22 Speed Reduced Groups

Data Amounts

Maximum Speed [MPH]		0-60	65-70	75-110	115-125
Data A Usage	Datasets Used	3	3	6	4
	Data points Used	1221	691	3790	953
Data B Usage	Datasets Used	21	20	30	19
	Data points Used	6504	3323	10767	2600

Table 8.192: Data Amounts for Maximum Speed Reduced

2-Tailed Hypothesis Tests

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	1.00	0.00	0.00	0.00
65-70	0.00	1.00	0.19	0.00
75-110	0.00	0.19	1.00	0.10
115-125	0.00	0.00	0.10	1.00

Table 8.193: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Maximum Speed Reduced (Data A, Usage)

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	0.00	1.00	1.00	1.00
65-70	1.00	0.00	0.67	1.00
75-110	1.00	0.67	0.00	0.75
115-125	1.00	1.00	0.75	0.00

Table 8.194: Mean 2 Tailed K-S and Mann Whitney U Test for Maximum Speed Reduced (Data A, Usage)

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	1.00	0.06	0.00	0.00
65-70	0.06	1.00	0.06	0.00
75-110	0.00	0.06	1.00	0.05
115-125	0.00	0.00	0.05	1.00

Table 8.195: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Maximum Speed Reduced (Data B, Usage)

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	0.00	0.83	1.00	1.00
65-70	0.83	0.00	0.80	1.00
75-110	1.00	0.80	0.00	0.82
115-125	1.00	1.00	0.82	0.00

Table 8.196: Mean 2 Tailed K-S and Mann Whitney U Test for Maximum Speed Reduced (Data B, Usage)

1-Tailed Hypothesis Tests

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	NaN	1.00	1.00	1.00
65-70	0.00	NaN	0.81	0.99
75-110	0.00	0.09	NaN	0.97
115-125	0.00	0.00	0.05	NaN

Table 8.197: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Maximum Speed Reduced (Data A, Usage)

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	NaN	0.00	0.00	0.00
65-70	1.00	NaN	0.00	0.00
75-110	1.00	0.67	NaN	0.00
115-125	1.00	1.00	0.75	NaN

Table 8.198: Mean 1 Tailed K-S and Mann Whitney U Test for Maximum Speed Reduced (Data A, Usage)

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	NaN	0.90	1.00	1.00
65-70	0.06	NaN	0.95	0.98
75-110	0.00	0.03	NaN	0.93
115-125	0.00	0.00	0.02	NaN

Table 8.199: Mean 1 Tailed K-S and Mann Whitney U Test p-values for Maximum Speed Reduced (Data B, Usage)

Maximum Speed [MPH]	0-60	65-70	75-110	115-125
0-60	NaN	0.00	0.00	0.00
65-70	0.85	NaN	0.00	0.00
75-110	1.00	0.88	NaN	0.00
115-125	1.00	1.00	0.84	NaN

Table 8.200: Mean 1 Tailed K-S and Mann Whitney U Test for Maximum Speed Reduced (Data B, Usage)

8.2.23 Sleepers

Data Amounts

Sleepers		F19	F23	F24	F27	F28	F40	G44	G47	G49	HH10	W402	W500	W560	W600	Metal	Hard-wood	Soft-wood	Tim-ber
Material Type		C	C	C	C	C	C	C	C	C	S	S	S	S	S	S	T	T	T
Data A Usage	Sets	0	1	0	1	0	1	7	2	2	1	0	0	2	0	0	1	0	0
	Points	0	26	0	54	0	21	4208	726	323	38	0	0	725	0	0	44	0	0
Data B Usage	Sets	4	4	14	25	21	16	14	2	3	2	3	2	4	3	7	9	11	2
	Points	106	135	1023	7459	1876	2045	3572	119	309	407	286	180	844	126	1255	280	463	86

Material Type		
	C	Concrete
	S	Steel
	T	Timber

Table 8.201: Data Amounts for Sleepers

2-Tailed Hypothesis Tests

Sleepers	G44	G47	G49	W560
G44	1.00	0.12	0.08	0.00
G47	0.12	1.00	0.19	0.00
G49	0.08	0.19	1.00	0.01
W560	0.00	0.00	0.01	1.00

Table 8.202: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers (Data A, Usage)

Sleepers	G44	G47	G49	W560
G44	0.00	0.25	0.50	1.00
G47	0.25	0.00	0.50	1.00
G49	0.50	0.50	0.00	1.00
W560	1.00	1.00	1.00	0.00

Table 8.203: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers (Data A, Usage)

Sleepers	F19	F23	F24	F27	F28	F40	G44	G47	G49	HH10	W402	W500	W560	W600	Metal	Hard-wood	Soft-wood	Tim-ber
F19	1.00	NaN	0.16	0.09	0.00	0.00	0.39	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.04	0.06	0.01	NaN
F23	NaN	1.00	0.23	0.17	0.00	0.00	0.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.21	0.65	0.17	NaN
F24	0.16	0.23	1.00	0.26	0.13	0.11	0.16	0.42	0.73	0.02	0.19	0.42	0.04	0.50	0.08	0.25	0.01	0.00
F27	0.09	0.17	0.26	1.00	0.06	0.04	0.06	0.44	0.10	0.00	0.07	0.19	0.03	0.42	0.25	0.28	0.02	0.00
F28	0.00	0.00	0.13	0.06	1.00	0.19	0.26	0.60	0.28	0.00	0.01	0.18	0.00	0.34	0.02	0.10	0.00	0.00
F40	0.00	0.00	0.11	0.04	0.19	1.00	0.10	0.02	0.29	0.13	0.01	0.03	0.09	0.06	0.14	0.12	0.01	0.00
G44	0.39	0.00	0.16	0.06	0.26	0.10	1.00	0.33	0.06	0.00	0.00	0.00	0.00	0.01	0.00	0.03	0.00	0.00
G47	NaN	NaN	0.42	0.44	0.60	0.02	0.33	1.00	0.13	0.00	0.01	0.08	0.00	0.58	0.00	0.35	0.00	NaN
G49	NaN	NaN	0.73	0.10	0.28	0.29	0.06	0.13	1.00	0.00	0.12	0.13	0.00	0.23	0.00	0.13	0.00	NaN
HH10	NaN	NaN	0.02	0.00	0.00	0.13	0.00	0.00	0.00	1.00	0.07	0.01	0.38	0.00	0.73	0.40	0.00	NaN
W402	NaN	NaN	0.19	0.07	0.01	0.01	0.00	0.01	0.12	0.07	1.00	0.36	0.03	0.31	0.05	0.18	0.00	NaN
W500	NaN	NaN	0.42	0.19	0.18	0.03	0.00	0.08	0.13	0.01	0.36	1.00	0.00	0.35	0.00	0.21	0.00	NaN
W560	NaN	NaN	0.04	0.03	0.00	0.09	0.00	0.00	0.00	0.38	0.03	0.00	1.00	0.02	0.40	0.26	0.00	NaN
W600	NaN	NaN	0.50	0.42	0.34	0.06	0.01	0.58	0.23	0.00	0.31	0.35	0.02	1.00	0.08	0.29	0.00	NaN
Metal	0.04	0.21	0.08	0.25	0.02	0.14	0.00	0.00	0.00	0.73	0.05	0.00	0.40	0.08	1.00	0.38	0.03	NaN
Hardwood	0.06	0.65	0.25	0.28	0.10	0.12	0.03	0.35	0.13	0.40	0.18	0.21	0.26	0.29	0.38	1.00	0.08	0.00
Softwood	0.01	0.17	0.01	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.08	1.00	0.08
Timber	NaN	NaN	0.00	0.00	0.00	0.00	0.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.00	0.08	1.00

Table 8.204: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers (Data B, Usage)

Sleepers	F19	F23	F24	F27	F28	F40	G44	G47	G49	HH10	W402	W500	W560	W600	Metal	Hard-wood	Soft-wood	Tim-ber
F19	0.00	NaN	0.50	0.75	1.00	1.00	0.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.50	0.50	1.00	NaN
F23	NaN	0.00	0.33	0.38	1.00	1.00	1.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.00	0.00	0.75	NaN
F24	0.50	0.33	0.00	0.39	0.75	0.73	0.64	0.00	0.00	0.75	0.33	0.00	0.67	0.00	0.60	0.17	0.94	1.00
F27	0.75	0.38	0.39	0.00	0.87	0.88	0.82	0.00	0.25	1.00	0.67	0.50	0.67	0.00	0.50	0.17	0.91	1.00
EF28	1.00	1.00	0.75	0.87	0.00	0.38	0.55	0.00	0.17	1.00	0.83	0.50	1.00	0.33	0.86	0.78	1.00	1.00
F40	1.00	1.00	0.73	0.88	0.38	0.00	0.75	0.75	0.50	0.50	1.00	0.75	0.67	0.50	0.57	0.69	0.88	1.00
G44	0.00	1.00	0.64	0.82	0.55	0.75	0.00	0.50	0.67	1.00	1.00	1.00	1.00	0.83	1.00	0.86	1.00	1.00
G47	NaN	NaN	0.00	0.00	0.00	0.75	0.50	0.00	0.00	1.00	1.00	0.50	1.00	0.00	1.00	0.50	1.00	NaN
G49	NaN	NaN	0.00	0.25	0.17	0.50	0.67	0.00	0.00	1.00	0.25	0.50	1.00	0.25	1.00	0.50	1.00	NaN
HH10	NaN	NaN	0.75	1.00	1.00	0.50	1.00	1.00	1.00	0.00	0.50	1.00	0.00	1.00	0.00	0.50	1.00	NaN
W402	NaN	NaN	0.33	0.67	0.83	1.00	1.00	1.00	0.25	0.50	0.00	0.50	0.83	0.50	0.67	0.33	1.00	NaN
W500	NaN	NaN	0.00	0.50	0.50	0.75	1.00	0.50	0.50	1.00	0.50	0.00	1.00	0.50	1.00	0.25	1.00	NaN
W560	NaN	NaN	0.67	0.67	1.00	0.67	1.00	1.00	1.00	0.00	0.83	1.00	0.00	0.83	0.00	0.33	1.00	NaN
W600	NaN	NaN	0.00	0.00	0.33	0.50	0.83	0.00	0.25	1.00	0.50	0.50	0.83	0.00	0.67	0.33	1.00	NaN
Metal	0.50	0.00	0.60	0.50	0.86	0.57	1.00	1.00	1.00	0.00	0.67	1.00	0.00	0.67	0.00	0.30	0.83	NaN
Hardwood	0.50	0.00	0.17	0.17	0.78	0.69	0.86	0.50	0.50	0.50	0.33	0.25	0.33	0.33	0.30	0.00	0.79	1.00
Softwood	1.00	0.75	0.94	0.91	1.00	0.88	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.83	0.79	0.00	0.50
Timber	NaN	NaN	1.00	1.00	1.00	1.00	1.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.00	0.50	0.00

Table 8.205: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers (Data B, Usage)

Second Layer Analysis

Data Amounts

Sleepers		F19	F23	F24	F27	F28	F40	G44	G47	G49	HH10	W402	W500	W560	W600	Metal	Hard-wood	Soft-wood	Tim-ber
Material Type		C	C	C	C	C	C	C	C	C	S	S	S	S	S	S	T	T	T
Data A Usage	Sets	0	1	0	1	0	0	12	5	5	1	0	0	6	0	0	1	0	0
	Points	0	23	0	38	0	0	4064	701	304	23	0	0	718	0	0	25	0	0
Data B Usage	Sets	1	2	23	59	33	28	23	2	5	5	5	4	7	3	12	3	10	2
	Points	26	49	929	7105	1601	1863	3374	74	259	387	223	151	766	80	1166	73	364	60

Material Type		
	C	Concrete
	S	Steel
	T	Timber

Table 8.206: Data Amounts for Sleepers (Second Layer)

2-Tailed Hypothesis Tests

Sleepers	G44	G47	G49	W560
G44	1.00	0.41	0.11	0.01
G47	0.41	1.00	0.28	0.01
G49	0.11	0.28	1.00	0.20
W560	0.01	0.01	0.20	1.00

Table 8.207: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers (Data A, Usage)

Sleepers	G44	G47	G49	W560
G44	0.00	0.20	0.50	0.80
G47	0.20	0.00	0.50	1.00
G49	0.50	0.50	0.00	0.60
W560	0.80	1.00	0.60	0.00

Table 8.208: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers (Data A, Usage)

Sleepers	F19	F23	F24	F27	F28	F40	G44	G47	G49	HH10	W402	W500	W560	W600	Metal	Hard-wood	Soft-wood	Tim-ber
F19	1.00	NaN	0.00	0.00	0.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
F23	NaN	1.00	NaN	0.41	0.01	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.00	NaN
F24	0.00	NaN	1.00	0.31	0.17	0.43	0.44	NaN	0.44	0.03	0.08	0.69	0.12	0.35	0.22	NaN	0.00	0.00
F27	0.00	0.41	0.31	1.00	0.06	0.24	0.09	NaN	NaN	0.13	0.11	0.73	0.04	0.63	0.21	0.16	0.00	0.00
F28	0.00	0.01	0.17	0.06	1.00	0.22	0.49	0.08	0.48	0.13	0.14	0.22	0.09	0.25	0.01	0.42	0.00	0.00
F40	NaN	NaN	0.43	0.24	0.22	1.00	0.13	0.20	0.41	0.11	0.35	0.17	0.04	0.31	0.03	0.21	NaN	NaN
G44	NaN	NaN	0.44	0.09	0.49	0.13	1.00	0.02	0.45	0.00	0.39	0.07	0.00	0.06	0.02	0.09	0.00	NaN
G47	NaN	NaN	NaN	NaN	0.08	0.20	0.02	1.00	0.01	0.07	NaN	0.01	0.30	NaN	0.31	0.03	0.00	NaN
G49	NaN	NaN	0.44	NaN	0.48	0.41	0.45	0.01	1.00	0.22	0.35	0.23	0.30	0.38	0.03	0.32	0.00	NaN
HH10	NaN	NaN	0.03	0.13	0.13	0.11	0.00	0.07	0.22	1.00	0.29	0.08	0.39	0.13	0.34	0.26	0.00	NaN
W402	NaN	NaN	0.08	0.11	0.14	0.35	0.39	NaN	0.35	0.29	1.00	0.04	0.22	0.28	0.07	0.02	NaN	NaN
W500	NaN	NaN	0.69	0.73	0.22	0.17	0.07	0.01	0.23	0.08	0.04	1.00	0.00	0.61	0.05	0.45	0.00	NaN
W560	NaN	NaN	0.12	0.04	0.09	0.04	0.00	0.30	0.30	0.39	0.22	0.00	1.00	0.01	0.22	0.14	0.00	NaN
W600	NaN	NaN	0.35	0.63	0.25	0.31	0.06	NaN	0.38	0.13	0.28	0.61	0.01	1.00	0.27	0.94	NaN	NaN
Metal	NaN	NaN	0.22	0.21	0.01	0.03	0.02	0.31	0.03	0.34	0.07	0.05	0.22	0.27	1.00	0.42	0.00	NaN
Hardwood	NaN	NaN	NaN	0.16	0.42	0.21	0.09	0.03	0.32	0.26	0.02	0.45	0.14	0.94	0.42	1.00	0.05	NaN
Softwood	NaN	0.00	0.00	0.00	0.00	NaN	0.00	0.00	0.00	0.00	NaN	0.00	0.00	NaN	0.00	0.05	1.00	0.22
Timber	NaN	NaN	0.00	0.00	0.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.22	1.00

Table 8.209: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers (Data B, Usage)

Sleepers	F19	F23	F24	F27	F28	F40	G44	G47	G49	HH10	W402	W500	W560	W600	Metal	Hard-wood	Soft-wood	Tim-ber
F19	0.00	NaN	1.00	1.00	1.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
F23	NaN	0.00	NaN	0.50	1.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.00	NaN
F24	1.00	NaN	0.00	0.32	0.54	0.19	0.33	NaN	0.00	1.00	0.17	0.00	0.67	0.00	0.50	NaN	1.00	1.00
F27	1.00	0.50	0.32	0.00	0.75	0.42	0.70	NaN	NaN	0.00	0.50	0.00	0.50	0.00	0.36	0.00	1.00	1.00
F28	1.00	1.00	0.54	0.75	0.00	0.39	0.22	0.50	0.00	0.70	0.80	0.00	0.71	0.67	0.96	0.33	1.00	1.00
F40	NaN	NaN	0.19	0.42	0.39	0.00	0.75	0.00	0.33	0.67	0.20	0.33	0.70	0.00	0.80	0.00	NaN	NaN
G44	NaN	NaN	0.33	0.70	0.22	0.75	0.00	0.75	0.25	1.00	0.50	0.50	1.00	0.67	0.79	0.00	1.00	NaN
G47	NaN	NaN	NaN	NaN	0.50	0.00	0.75	0.00	1.00	0.50	NaN	1.00	0.50	NaN	0.50	0.50	1.00	NaN
G49	NaN	NaN	0.00	NaN	0.00	0.33	0.25	1.00	0.00	0.75	0.00	0.17	0.60	0.00	0.80	0.00	1.00	NaN
HH10	NaN	NaN	1.00	0.00	0.70	0.67	1.00	0.50	0.75	0.00	0.50	0.63	0.10	0.50	0.30	0.50	1.00	NaN
W402	NaN	NaN	0.17	0.50	0.80	0.20	0.50	NaN	0.00	0.50	0.00	0.50	0.67	0.25	0.70	1.00	NaN	NaN
W500	NaN	NaN	0.00	0.00	0.00	0.33	0.50	1.00	0.17	0.63	0.50	0.00	1.00	0.00	0.63	0.00	1.00	NaN
W560	NaN	NaN	0.67	0.50	0.71	0.70	1.00	0.50	0.60	0.10	0.67	1.00	0.00	1.00	0.50	0.50	1.00	NaN
W600	NaN	NaN	0.00	0.00	0.67	0.00	0.67	NaN	0.00	0.50	0.25	0.00	1.00	0.00	0.67	0.00	NaN	NaN
Metal	NaN	NaN	0.50	0.36	0.96	0.80	0.79	0.50	0.80	0.30	0.70	0.63	0.50	0.67	0.00	0.33	1.00	NaN
Hardwood	NaN	NaN	NaN	0.00	0.33	0.00	0.00	0.50	0.00	0.50	1.00	0.00	0.50	0.00	0.33	0.00	0.50	NaN
Softwood	NaN	1.00	1.00	1.00	1.00	NaN	1.00	1.00	1.00	1.00	NaN	1.00	1.00	NaN	1.00	0.50	0.00	0.50
Timber	NaN	NaN	1.00	1.00	1.00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0.50	0.00

Table 8.210: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers (Data B, Usage)

8.2.24 Sleeper Reduced Groups

Second Layer Analysis

Data Amounts

Sleeper Group		Group 1	Group 2	Group 3
Data A Usage	Datasets Used	6	13	1
	Data points Used	836	5160	70
Data B Usage	Datasets Used	70	60	29
	Data points Used	9745	7967	3097
Group 1	F23, F24, W600, F27, W560			
Group 2	F28, G49, G44, G47, F40			
Group 3	Timber, W402, HH10, Metal, W500			

Table 8.211: Data Amounts for Sleepers Grouped

2-Tailed Hypothesis Tests

Sleeper Group	Group 1	Group 2
Group 1	1.00	0.08
Group 2	0.08	1.00
Group 1	F23, F24, W600, F27, W560	
Group 2	F28, G49, G44, G47, F40	

Table 8.212: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers Grouped (Data A, Usage)

Sleeper Group	Group 1		Group 2
Group 1	0.00		0.83
Group 2	0.83		0.00
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		

Table 8.213: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers Grouped (Data A, Usage)

Sleeper Group	Group 1	Group 2	Group 3
Group 1	1.00	0.11	0.19
Group 2	0.11	1.00	0.03
Group 3	0.19	0.03	1.00
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.214: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers Grouped (Data B, Usage)

Sleeper Group	Group 1	Group 2	Group 3
Group 1	0.00	0.72	0.50
Group 2	0.72	0.00	0.87
Group 3	0.50	0.87	0.00
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.215: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers Grouped (Data B, Usage)

1-Tailed Hypothesis Tests

Sleeper Group	Group 1		Group 2
Group 1	NA		0.94
Group 2	0.04		NA
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		

Table 8.216: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers Grouped (Data A, Usage)

Sleeper Group	Group 1		Group 2
Group 1	NA		0.00
Group 2	0.83		NA
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		

Table 8.217: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers Grouped (Data A, Usage)

Sleeper Group	Group 1	Group 2	Group 3
Group 1	NA	0.90	0.28
Group 2	0.09	NA	0.03
Group 3	0.69	0.94	NA
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.218: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers Grouped (Data B, Usage)

Sleeper Group	Group 1	Group 2	Group 3
Group 1	NA	0.00	0.52
Group 2	0.74	NA	0.91
Group 3	0.11	0.00	NA
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.219: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers Grouped (Data B, Usage)

Third Layer Analysis

Data Amounts

Sleeper Group		Group 1	Group 2	Group 3
Data A Usage	Datasets Used	3	18	7
	Data points Used	146	4877	757
Data B Usage	Datasets Used	73	72	30
	Data points Used	10400	7417	1887
Group 1	F23, F24, W500, W600, F27, W560			
Group 2	F28, G49, G44, G47, F40			
Group 3	Timber, W402, HH10, Metal, W500			

Table 8.220: Data Amounts for Sleepers Grouped

2-Tailed Hypothesis Tests

Sleeper Group	Group 1	Group 2	Group 3
Group 1	1.00	0.00	0.37
Group 2	0.00	1.00	0.09
Group 3	0.37	0.09	1.00
Group 1	F23, F24, W500, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.221: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers Grouped (Data A, Usage)

Sleeper Group	Group 1	Group 2	Group 3
Group 1	0.00	1.00	0.00
Group 2	1.00	0.00	0.86
Group 3	0050	0.86	0.00
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.222: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers Grouped (Data A, Usage)

Sleeper Group	Group 1	Group 2	Group 3
Group 1	1.00	0.11	0.17
Group 2	0.11	1.00	0.09
Group 3	0.17	0.09	1.00
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.223: Mean 2 Tailed K-S and Mann Whitney U Test p-values for Sleepers Grouped (Data B, Usage)

Sleeper Group	Group 1	Group 2	Group 3
Group 1	0.00	0.74	0.59
Group 2	0.74	0.00	0.83
Group 3	0.59	0.83	0.00
Group 1	F23, F24, W600, F27, W560		
Group 2	F28, G49, G44, G47, F40		
Group 3	Timber, W402, HH10, Metal, W500		

Table 8.224: Mean 2 Tailed K-S and Mann Whitney U Test for Sleepers Grouped (Data B, Usage)

8.2.25 Track Geometry Degradation Model

Table 8.225: Vertical geometry degradation on large concrete sleepers with track speeds 5-60MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull A	2.126e-4	1.698e-4	1.800e-4	1.472e-4	7.973e-5	1.559e-4	1.664e-4	NA
Weibull β	5.635e-1	1.304e0	9.726e-1	1.767e0	4.304e0	1.821e0	1.340e0	NA
Min	5.014e-5	5.013e-5	5.043e-5	5.318e-5	5.175e-5	5.285e-5	5.051e-5	6.996e-5
Q1	6.888e-5	7.355e-5	6.815e-5	7.463e-5	5.667e-5	6.815e-5	7.085e-5	6.996e-5
Median	1.040e-4	1.032e-4	9.093e-5	9.634e-5	6.352e-5	1.021e-4	9.716e-5	6.996e-5
Q3	1.724e-4	1.823e-4	1.898e-4	1.739e-4	9.006e-5	2.138e-4	1.943e-4	6.996e-5
Max	1.147e-1	1.189e-3	3.477e-3	3.984e-4	1.043e-4	2.860e-4	8.338e-4	6.996e-5
Mean	9.174e-4	1.546e-4	1.832e-4	1.297e-4	7.251e-5	1.375e-4	1.508e-4	6.996e-5
Number	150	196	108	26	10	12	81	1

Table 8.226: Vertical geometry degradation on small concrete and large steel sleepers with track speeds 5-60MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.933E-4	2.145E-4	2.222E-4	2.061E-4	2.211E-4	2.367E-4	2.124E-4	NA
Weibull β	1.029E+0	8.184E-1	1.049E+0	1.165E+0	7.253E-1	1.004E+0	6.490E-1	NA
Min	5.134E-5	5.001E-5	5.021E-5	5.036E-5	5.033E-5	5.039E-5	5.001E-5	5.652E-5
Q1	6.682E-5	7.053E-5	7.671E-5	7.540E-5	7.674E-5	7.110E-5	7.322E-5	5.652E-5
Median	1.011E-4	1.068E-4	1.271E-4	1.184E-4	1.212E-4	1.239E-4	1.098E-4	5.652E-5
Q3	1.608E-4	1.929E-4	2.495E-4	1.945E-4	1.941E-4	2.591E-4	1.754E-4	5.652E-5
Max	1.460E-3	1.853E-2	3.306E-3	1.587E-3	3.145E-2	2.916E-3	1.295E-1	5.652E-5
Mean	1.905E-4	2.577E-4	2.168E-4	1.933E-4	3.458E-4	2.362E-4	5.083E-4	5.652E-5
Number	86	371	278	193	193	219	428	1

Table 8.227: Vertical geometry degradation on small steel and timber sleepers with track speeds 5-60MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.683E-4	2.449E-4	2.128E-4	6.852E-4	4.508E-4	5.928E-4	4.757E-4	NA
Weibull β	9.376E-1	8.984E-1	1.032E+0	6.843E-1	1.105E+0	1.108E+0	8.192E-1	NA
Min	5.186E-5	5.009E-5	5.043E-5	5.014E-5	5.519E-5	5.400E-5	5.018E-5	NA
Q1	6.617E-5	7.817E-5	7.302E-5	1.725E-4	1.572E-4	1.669E-4	9.216E-5	NA
Median	9.074E-5	1.240E-4	1.224E-4	4.178E-4	2.668E-4	3.820E-4	2.565E-4	NA
Q3	1.659E-4	2.291E-4	2.033E-4	7.839E-4	5.568E-4	7.972E-4	5.136E-4	NA
Max	4.391E-3	5.278E-3	1.840E-3	7.507E-2	2.674E-3	3.297E-3	4.742E-3	NA
Mean	1.760E-4	2.632E-4	2.093E-4	1.126E-3	4.319E-4	5.687E-4	5.443E-4	NA
Number	122	590	164	179	87	149	139	0

Table 8.228: Vertical geometry degradation on large concrete sleepers with track speeds 65-70MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.430E-4	1.692E-4	1.202E-4	1.177E-4	1.499E-4	9.852E-5	1.584E-4	NA
Weibull β	1.537E+0	1.410E+0	1.903E+0	2.481E+0	2.122E+0	2.740E+0	6.454E-1	NA
Min	5.122E-5	5.044E-5	5.024E-5	5.077E-5	5.857E-5	5.067E-5	5.071E-5	8.876E-5
Q1	6.520E-5	7.022E-5	6.374E-5	6.554E-5	6.414E-5	5.873E-5	6.136E-5	8.876E-5
Median	9.428E-5	1.012E-4	9.134E-5	9.976E-5	1.149E-4	7.727E-5	7.772E-5	8.928E-5
Q3	1.409E-4	1.987E-4	1.206E-4	1.327E-4	1.983E-4	1.157E-4	1.056E-4	8.980E-5
Max	6.220E-4	4.570E-4	3.510E-4	2.246E-4	2.527E-4	1.487E-4	1.002E-2	8.980E-5
Mean	1.272E-4	1.522E-4	1.059E-4	1.041E-4	1.320E-4	8.732E-5	3.168E-4	8.928E-5
Number	68	63	36	16	10	7	44	2

Table 8.229: Vertical geometry degradation on small concrete and large steel sleepers with track speeds 65-70MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.361E-4	1.402E-4	1.640E-4	1.559E-4	1.547E-4	1.410E-4	1.339E-4	NA
Weibull β	1.616E+0	1.566E+0	9.848E-1	1.212E+0	1.584E+0	1.395E+0	1.548E+0	NA
Min	5.132E-5	5.060E-5	5.009E-5	5.055E-5	5.036E-5	5.017E-5	5.044E-5	5.916E-5
Q1	6.454E-5	6.543E-5	6.890E-5	6.819E-5	6.893E-5	6.352E-5	6.959E-5	7.651E-5
Median	8.353E-5	8.753E-5	1.027E-4	8.643E-5	9.800E-5	9.066E-5	9.360E-5	1.286E-4
Q3	1.407E-4	1.547E-4	1.594E-4	1.424E-4	1.836E-4	1.431E-4	1.361E-4	1.503E-4
Max	4.098E-4	5.352E-4	4.610E-3	1.155E-3	4.779E-4	7.463E-4	7.832E-4	1.575E-4
Mean	1.205E-4	1.244E-4	1.656E-4	1.442E-4	1.373E-4	1.268E-4	1.190E-4	1.151E-4
Number	53	164	168	118	87	104	179	3

Table 8.230: Vertical geometry degradation on small steel and timber sleepers with track speeds 65-70MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.413E-4	1.601E-4	1.948E-4	1.652E-4	6.280E-4	2.690E-4	1.182E-4	NA
Weibull β	1.946E+0	1.514E+0	1.011E+0	1.587E+0	7.753E-1	9.743E-1	2.001E+0	NA
Min	5.027E-5	5.054E-5	5.077E-5	5.149E-5	5.112E-5	5.624E-5	5.031E-5	NA
Q1	7.342E-5	7.807E-5	6.832E-5	8.490E-5	9.873E-5	6.596E-5	6.855E-5	NA
Median	1.053E-4	1.054E-4	1.175E-4	9.933E-5	3.023E-4	1.411E-4	8.028E-5	NA
Q3	1.580E-4	1.744E-4	1.490E-4	1.629E-4	1.405E-3	3.217E-4	1.191E-4	NA
Max	3.447E-4	7.898E-4	1.723E-3	3.918E-4	2.426E-3	1.140E-3	2.753E-4	NA
Mean	1.244E-4	1.426E-4	1.936E-4	1.465E-4	7.358E-4	2.726E-4	1.040E-4	NA
Number	44	125	45	19	7	9	26	0

Table 8.231: Vertical geometry degradation on large concrete sleepers with track speeds 75-110MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.263E-4	1.434E-4	1.292E-4	1.556E-4	1.550E-4	1.043E-4	1.450E-4	1.090E-4
Weibull β	1.470E+0	1.235E+0	1.485E+0	1.381E+0	1.249E+0	3.316E+0	1.287E+0	2.360E+0
Min	5.016E-5	5.036E-5	5.034E-5	5.126E-5	5.187E-5	5.854E-5	5.012E-5	5.131E-5
Q1	6.313E-5	6.491E-5	6.406E-5	6.544E-5	6.440E-5	7.284E-5	6.160E-5	6.889E-5
Median	8.332E-5	8.727E-5	8.478E-5	9.893E-5	9.145E-5	8.020E-5	8.229E-5	8.078E-5
Q3	1.245E-4	1.464E-4	1.255E-4	1.566E-4	1.418E-4	1.199E-4	1.443E-4	1.145E-4
Max	1.029E-3	1.555E-3	6.751E-4	6.677E-4	7.529E-4	1.547E-4	1.002E-3	2.079E-4
Mean	1.127E-4	1.317E-4	1.152E-4	1.401E-4	1.423E-4	9.345E-5	1.322E-4	9.624E-5
Number	352	318	192	69	44	10	217	12

Table 8.232: Vertical geometry degradation on small concrete and large steel sleepers with track speeds 75-110MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.897E-4	1.500E-4	1.327E-4	1.325E-4	1.251E-4	1.361E-4	1.189E-4	1.607E-4
Weibull β	1.041E+0	1.488E+0	1.568E+0	1.659E+0	1.712E+0	1.323E+0	1.656E+0	2.272E+0
Min	5.010E-5	5.008E-5	5.049E-5	5.029E-5	5.004E-5	5.001E-5	5.003E-5	6.416E-5
Q1	6.976E-5	6.972E-5	6.704E-5	6.318E-5	6.610E-5	6.375E-5	6.346E-5	7.596E-5
Median	1.142E-4	1.007E-4	8.684E-5	8.721E-5	8.695E-5	8.380E-5	8.235E-5	1.298E-4
Q3	1.692E-4	1.569E-4	1.343E-4	1.428E-4	1.258E-4	1.252E-4	1.175E-4	2.057E-4
Max	2.511E-3	8.831E-4	5.727E-4	5.098E-4	4.914E-4	8.782E-4	5.696E-4	2.517E-4
Mean	1.857E-4	1.338E-4	1.177E-4	1.171E-4	1.104E-4	1.232E-4	1.052E-4	1.415E-4
Number	92	141	245	231	184	209	321	10

Table 8.233: Vertical geometry degradation on small steel and timber sleepers with track speeds 75-110MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.276E-4	1.777E-4	1.326E-4	1.871E-4	1.900E-4	3.121E-4	1.151E-4	NA
Weibull β	2.346E+0	1.457E+0	1.852E+0	9.318E-1	9.714E-1	1.242E+0	1.741E+0	NA
Min	5.278E-5	5.010E-5	5.033E-5	5.059E-5	5.061E-5	5.376E-5	5.170E-5	NA
Q1	6.205E-5	7.334E-5	6.776E-5	6.518E-5	7.296E-5	8.242E-5	6.396E-5	NA
Median	1.116E-4	1.144E-4	1.008E-4	9.153E-5	9.698E-5	2.378E-4	8.110E-5	NA
Q3	1.451E-4	2.092E-4	1.472E-4	1.253E-4	1.376E-4	3.994E-4	1.081E-4	NA
Max	2.644E-4	8.148E-4	4.042E-4	1.322E-3	1.134E-3	9.674E-4	3.314E-4	NA
Mean	1.126E-4	1.592E-4	1.168E-4	1.954E-4	1.932E-4	2.898E-4	1.014E-4	NA
Number	36	166	73	38	26	13	19	0

Table 8.234: Vertical geometry degradation on large concrete sleepers with track speeds 110-125MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	1.089E-4	1.621E-4	1.996E-4	2.695E-4	1.546E-4	2.724E-4	1.086E-4	NA
Weibull β	1.753E+0	1.095E+0	1.163E+0	1.104E+0	9.954E-1	1.640E+0	1.861E+0	NA
Min	5.013E-5	5.044E-5	5.039E-5	5.169E-5	5.219E-5	7.316E-5	5.028E-5	5.362E-5
Q1	6.284E-5	5.838E-5	6.376E-5	6.772E-5	5.249E-5	8.641E-5	5.788E-5	5.362E-5
Median	7.369E-5	8.014E-5	1.129E-4	1.368E-4	7.341E-5	1.414E-4	7.134E-5	5.362E-5
Q3	1.018E-4	1.271E-4	2.570E-4	2.701E-4	1.873E-4	2.956E-4	1.166E-4	5.362E-5
Max	3.518E-4	9.457E-4	8.403E-4	8.257E-4	5.205E-4	1.022E-3	2.873E-4	5.362E-5
Mean	9.593E-5	1.554E-4	1.878E-4	2.587E-4	1.550E-4	2.620E-4	9.567E-5	5.362E-5
Number	48	123	58	22	5	6	19	1

Table 8.235: Vertical geometry degradation on small concrete and large steel sleepers with track speeds 110-125MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	4.220E-5	1.619E-4	1.624E-4	9.611E-5	6.630E-5	1.247E-4	7.562E-5	NA
Weibull β	9.378E-1	2.028E+0	1.206E+0	2.845E+0	4.568E+0	2.043E+0	5.093E+0	NA
Min	1.847E-5	6.153E-5	5.380E-5	5.145E-5	5.043E-5	5.606E-5	5.571E-5	NA
Q1	1.950E-5	7.007E-5	5.512E-5	5.375E-5	5.219E-5	5.683E-5	5.935E-5	NA
Median	2.894E-5	1.156E-4	8.387E-5	6.699E-5	5.936E-5	9.480E-5	6.157E-5	NA
Q3	4.969E-5	2.206E-4	2.075E-4	8.024E-5	6.024E-5	1.395E-4	8.669E-5	NA
Max	7.950E-5	2.524E-4	4.283E-4	9.124E-5	9.018E-5	2.250E-4	9.251E-5	NA
Mean	3.792E-5	1.425E-4	1.511E-4	8.499E-5	6.100E-5	1.097E-4	6.957E-5	NA
Number	7	5	5	5	7	7	6	0

Table 8.236: Vertical geometry degradation on small steel and timber sleepers with track speeds 110-125MPH, Weibull parameters and descriptive statistics [nm/ Equivalent Million Gross Tonnage]

		After Tamping					After Stoneblowing	
History:	R	T0-1	T2-3	T4-5	T5-6	T6-7	S0-1	S2-3
Weibull η	NA	NA	NA	NA	NA	NA	NA	NA
Weibull β	NA	NA	NA	NA	NA	NA	NA	NA
Min	6.806E-5	NA	NA	NA	NA	NA	NA	NA
Q1	6.806E-5	NA	NA	NA	NA	NA	NA	NA
Median	6.806E-5	NA	NA	NA	NA	NA	NA	NA
Q3	6.806E-5	NA	NA	NA	NA	NA	NA	NA
Max	6.806E-5	NA	NA	NA	NA	NA	NA	NA
Mean	6.806E-5	NA	NA	NA	NA	NA	NA	NA
Number	1	0	0	0	0	0	0	0

8.3 Railway Track Asset Management Model

Table 8.237: Maintenance methods for different rail faults (Stacked Probabilities)

Fault Type	f	Rerail (fmt = 3)	Or Weld (fmt = 2)	Or Grind and Other (fmt = 1)
Squat	1	0.328	0.954	1.000
Tache Ovale	2	0.722	0.963	1.000
Bolt Hole	3	0.900	0.919	1.000
Weld	6	0.519	0.904	1.000
Other	7	0.641	0.918	1.000
RCF	8	0.508	0.786	1.000
Wheelburn	4	0.267	0.874	1.000
Lipping	10	0.029	0.134	1.000
Side wear	11	0.044	0.338	1.000
Head wear	12	0.213	0.752	1.000
Corrugation	5	0.706	0.765	1.000
Unknown	9	0.464	0.630	1.000