Predictive Models for Railway Track Geometry Degradation

Iman Soleimanmeigouni

Operation and Maintenance Engineering
Predictive Models for Railway Track Geometry Degradation

Iman Soleimanmeigouni

Luleå 2019

Division of Operation and Maintenance Engineering
Department of Civil, Environmental and Natural Resources Engineering
Luleå University of Technology
SE-97187 Luleå, Sweden
Acknowledgements

The research presented in this thesis has been carried out at the Division of Operation, Maintenance and Acoustics of Luleå University of Technology (LTU). I gratefully acknowledge Luleå Railway Research Center (JVTC) and the Swedish Transport Administration (Trafikverket) for financing the project.

Furthermore, I would like to express my deep gratitude to my supervisors. Accordingly, I sincerely thank Professor Alireza Ahmadi for his great enthusiasm from the very first day of the research for my PhD until the completion of this research, for always supporting me and listening to my ideas and arguments, and for trusting me to follow my own chosen path of research to the end. I have been extremely lucky to have had a supervisor who has cared so much about my work and who has responded to my questions and queries so promptly. Moreover, I sincerely thank Professor Uday Kumar for providing me with excellent supervision and advice throughout the research for my PhD and for always being ready to provide assistance with a smile.

I would like to thank Dr Arne Nissen for his valuable suggestions and his contribution to my research, for providing the required data and information, and for sharing his experiences with me. I wish to thank Dr Iman Arastehkhhouy for sharing his expertise and time, and for being both a friend and a mentor. I would like to thank Dr Stephen Mayowa Famurewa for the fruitful discussions. Special thanks are extended to Dr Rickard Granström and Professor Peter Söderholm for spending their time and making the effort required to review the draft of the thesis and for making valuable comments which have led to improvements.

I have had the opportunity to participate in the SIMTRACK project and wish to thank the project partners for all the valuable and enjoyable discussions. In particular, I would like to thank Roland Bång, Martin Li, Jan Spännar, Mikael Åstrand, Professor Ramin Karim, Dr Adithya Thaduri, Dr Mohammad Hadadzadeh, and Hamid Khajehei for all the valuable and enjoyable discussions we have engaged in.

The research for my PhD involved me spending approximately two months at the City University of Hong Kong (CityU) at the Department of System Engineering and Engineering Management. Cordial thanks are extended to Professor Min Xie. It was a pleasure to participate in research in co-operation with him and his group. Special thanks are extended to Dr Xun Xiao for his continuous help and guidance during the research.
I also wish to thank Dr Christophe Letot from the University of Mons for his fruitful discussions, guidance and support.

I would like to thank all my colleagues and friends at the Division of Operation and Maintenance Engineering at LTU for their friendly and open-minded working environment. I would also to thank Veronica Jägare from JVTC for her support. The administrative support received from Cecilia Glover and Alexandra Lund Cipolla is also gratefully acknowledged.

Finally, my deepest gratitude is owed to my parents, Ghaffar and Roya, and to my brother, Pouya, for their love, support and understanding. I am forever indebted to my family for giving me the opportunities and experiences that have made me who I am. This journey would not have been possible if not for them, and I dedicate this milestone to them.

Iman Soleimanmeigouni

September, 2019

Luleå, Sweden
Abstract

Railways are a vital and effective means of mass transportation and play a vital role in modern transportation and social development. The benefits of the railway compared to other transportation modes are a high capacity, high efficiency and low pollution, and owing to these advantages, railways are nowadays experiencing a higher demand for the transportation of passengers and goods. This is in turn imposing higher demands on the railway capacity and service quality. As a result, infrastructure managers are being driven to develop new strategies and plans to fulfil new requirements, which include a higher level of resilience against failure, a more robust and available infrastructure, and cost reduction. This can be achieved by making efficient and effective maintenance decisions by applying RAMS (reliability, availability, maintainability, and safety) analysis and LCC (life cycle cost) assessment.

A major part of the railway maintenance burden is related to track geometry maintenance. Due to the forces induced on the track by traffic, the railway degrades over time, causing deviations from the designed vertical and horizontal alignment. When the track geometry degrades to an unacceptable level, this can cause catastrophic consequences, such as derailment. Maintenance actions are used to control the degradation of the track and restore the geometry condition of the track sections to an acceptable state.

With the current advancements in the field of technologies for railway track geometry measurement, a large amount of event data and condition monitoring data is available. Such technologies, along with advances in predictive analytics, are providing the possibility of predicting the track geometry condition in support of a predictive maintenance strategy. The aim of the research conducted for this thesis has been to develop methodologies and tools for the prediction of railway track geometry degradation, in order to facilitate and enhance the capability of making effective decisions for inspection and maintenance planning. To achieve the purpose of this research, literature studies, case studies and simulations have been conducted.

Firstly, a literature review was performed to identify the existing knowledge gaps and challenges for track geometry degradation modelling and maintenance planning. Secondly, a case study was conducted to analyse the effect of tamping on the track geometry condition. By considering the track geometry condition before tamping as the predictor, a probabilistic approach was utilised to model the recovery after tamping interventions. Thirdly, a two-level piecewise linear framework was developed to model the track geometry evolution over a spatial and temporal space. This model was implemented in a comprehensive case study. Fourthly, a data-driven analytical model was
developed to predict the occurrence of track geometry defects. This model enables infrastructure managers to predict the occurrence of severe isolated geometry defects. Finally, an integrated model was created to investigate the effect of different inspection intervals on the track geometry condition.

**Keywords:** Data-driven models, Degradation, Inspection, Maintenance, Predictive analytics, Tamping, Track geometry, Railway infrastructure, RAMS.
List of appended papers

Paper 1

Paper 2

Paper 3

Paper 4

Paper 5
The works carried out in the appended papers have been contributed by the thesis author as well as the other co-authors. The contributions of the author and the co-authors in the papers are highlighted in the table below:

1. Idea conception
2. Data processing and model development
3. Results and discussions
4. Article writing
5. Revision of important intellectual content

<table>
<thead>
<tr>
<th>Author/Paper</th>
<th>Paper 1</th>
<th>Paper 2</th>
<th>Paper 3</th>
<th>Paper 4</th>
<th>Paper 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iman Soleimanmeigouni</td>
<td>1,2,3,4</td>
<td>1,2,3,4</td>
<td>1,2,3,4</td>
<td>1,2,3,4</td>
<td>1,2,3,4</td>
</tr>
<tr>
<td>Alireza Ahmadi</td>
<td>1,2,3,4,5</td>
<td>1,2,3,5</td>
<td>1,2,3,5</td>
<td>1,3,5</td>
<td>1,3,5</td>
</tr>
<tr>
<td>Arne Nissen</td>
<td></td>
<td>2,3</td>
<td>2,3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Xun Xiao</td>
<td></td>
<td>2,3,4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iman Arasteh Khouy</td>
<td></td>
<td>1,2,3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christophe Letot</td>
<td></td>
<td>2,3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uday Kumar</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min Xie</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamid Khajehei</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
List of other papers

**Paper A**

**Paper B**

**Paper C**

**Paper D**

**Paper E**

**Paper F**
**Paper G**


**Paper H**


**Paper I**


**Paper J**


**Paper K**

## Acronyms and symbols

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_0$</td>
<td>Absolute amplitude of the longitudinal level defect after the latest tamping intervention</td>
</tr>
<tr>
<td>$A(t)$</td>
<td>Absolute value of the amplitude of the longitudinal level defect in time $t$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Degradation rate</td>
</tr>
<tr>
<td>ACF</td>
<td>Autocorrelation Function</td>
</tr>
<tr>
<td>AD</td>
<td>Anderson-Darling</td>
</tr>
<tr>
<td>ADASYN</td>
<td>Adaptive Synthetic sampling method</td>
</tr>
<tr>
<td>ADF</td>
<td>Augmented Dickey-Fuller</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>AL</td>
<td>Alert Limit</td>
</tr>
<tr>
<td>ARMA</td>
<td>Autoregressive Moving Average</td>
</tr>
<tr>
<td>Bessy</td>
<td>Trafikverket’s inspection report system</td>
</tr>
<tr>
<td>BIC</td>
<td>Bayesian Information Criterion</td>
</tr>
<tr>
<td>BIS</td>
<td>Trafikverket’s asset register system</td>
</tr>
<tr>
<td>BL</td>
<td>Box-Ljung</td>
</tr>
<tr>
<td>CM</td>
<td>Corrective Maintenance</td>
</tr>
<tr>
<td>CM$_e$</td>
<td>Emergency corrective maintenance</td>
</tr>
<tr>
<td>CM$_n$</td>
<td>Normal corrective maintenance</td>
</tr>
<tr>
<td>IAL</td>
<td>Immediate Action Limit</td>
</tr>
<tr>
<td>IL</td>
<td>Intervention Limit</td>
</tr>
<tr>
<td>KS</td>
<td>Kolmogorov-Smirnov</td>
</tr>
<tr>
<td>LCC</td>
<td>Life Cycle Cost</td>
</tr>
<tr>
<td>Optram</td>
<td>Trafikverket’s track geometry maintenance database</td>
</tr>
<tr>
<td>PACF</td>
<td>Partial Autocorrelation Function</td>
</tr>
<tr>
<td>PM</td>
<td>Preventive Maintenance</td>
</tr>
<tr>
<td>Q</td>
<td>Vertical force generated by a wheel</td>
</tr>
<tr>
<td>RAMS</td>
<td>Reliability, Availability, Maintainability, and Safety</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>RQ</td>
<td>Research Question</td>
</tr>
<tr>
<td>RUL</td>
<td>Remaining Useful Life</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SDLL</td>
<td>Standard Deviation of the Longitudinal Level</td>
</tr>
<tr>
<td>$t_{\text{tamp}}$</td>
<td>The latest tamping time</td>
</tr>
<tr>
<td>TQI</td>
<td>Track Quality Index</td>
</tr>
<tr>
<td>UH1</td>
<td>Underhåll 1 (in English: Maintenance 1). Lower bound for corrective maintenance</td>
</tr>
<tr>
<td>UH2</td>
<td>Underhåll 2 (in English: Maintenance 2). Upper bound for corrective maintenance</td>
</tr>
<tr>
<td>Y</td>
<td>Lateral force generated by a wheel</td>
</tr>
<tr>
<td>$y_P$</td>
<td>Distance [mm] between point P and a reference line, used to measure alignment (EN 13848-1)</td>
</tr>
<tr>
<td>$Z_P$</td>
<td>Limit [mm] of the range below the running surface within which the gauge is measured: “Zp is always 14 mm” (EN 13848-1)</td>
</tr>
<tr>
<td>$Z_{P'1}$</td>
<td>Deviation [mm] in the direction of consecutive running table levels on left hand rail, used to measure longitudinal level (EN 13848-1)</td>
</tr>
<tr>
<td>$Z_{P'2}$</td>
<td>Deviation [mm] in the direction of consecutive running table levels on right hand rail, used to measure longitudinal level (EN 13848-1)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Wavelength</td>
</tr>
<tr>
<td>$\omega$</td>
<td>Cut point</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Length of inspection interval</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>Gaussian random error term</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>$\Delta t$</td>
<td>Time interval</td>
</tr>
</tbody>
</table>
Table of Contents

Acknowledgements ........................................................................................................... i
Abstract ................................................................................................................................. iii
List of appended papers .......................................................................................................... v
Distribution of works .............................................................................................................. vii
List of other papers ................................................................................................................ ix
Acronyms and symbols .......................................................................................................... xi
Table of Contents ................................................................................................................... xiii

Introduction ............................................................................................................................. 1
  1.1. Background ..................................................................................................................... 1
  1.2. Statement of the problem ............................................................................................... 3
  1.3. Purpose of the research .................................................................................................. 4
  1.4. Objectives ...................................................................................................................... 4
  1.5. Research questions ......................................................................................................... 5
  1.6. Scope and delimitation of the present research ............................................................... 5
  1.7. Structure of the thesis .................................................................................................... 5

Theoretical framework ............................................................................................................. 7
  2.1. Track geometry ............................................................................................................... 7
  2.2. Track geometry degradation ......................................................................................... 9
  2.3. Track quality index ....................................................................................................... 10
  2.4. Maintenance limits ....................................................................................................... 13
  2.5. Track geometry measurement process ......................................................................... 15
  2.6. Tamping process .......................................................................................................... 16
  2.7. Predictive models for track geometry maintenance ...................................................... 18

Research methodology ......................................................................................................... 21
  3.1. Research purpose ......................................................................................................... 21
  3.2. Research approach ....................................................................................................... 22
  3.3. Data collection and data analysis ................................................................................... 23
  3.4. Research process .......................................................................................................... 25

Summary of the appended papers .......................................................................................... 29
  4.1. Paper 1 ......................................................................................................................... 29
  4.2. Paper 2 ......................................................................................................................... 30
  4.3. Paper 3 ......................................................................................................................... 31
  4.4. Paper 4 ......................................................................................................................... 32
  4.5. Paper 5 ......................................................................................................................... 33
Results and discussion

5.1. Results and discussion related to RQ 1
5.2. Results and discussion related to RQ 2
5.3. Results and discussion related to RQ 3

Conclusions, contributions, and future research

6.1. Conclusions
6.2. Contributions
6.3. Suggested future works

References
Chapter 1

Introduction

A brief introduction is given in this chapter to make the reader acquainted with the problem area. Moreover, the research purpose, research questions and research delimitation, as well as the thesis structure, are presented.

1.1. Background

The reliable and durable performance of railway infrastructures is an essential component for the sustained economic growth and social development of modern society. In recent years, there has been an increase in rail freight and passenger transportation in European countries. Considering the Swedish rail transportation, the average annual growth is estimated to be 3% for passenger traffic and 1% for freight traffic up to 2050 (Trafikverket 2017). This will in turn impose higher demands on the railway capacity and service quality. A study on the Swedish railway sector has identified the key problems within the existing network and accordingly has discussed some future problems and challenges. The main future challenges for the Swedish railway sector are as follow: i) a backlog of track infrastructure maintenance, ii) capacity problems and iii) punctuality (Transport Analysis 2014). Building new railway infrastructures cannot solely solve the capacity deficiencies considering the dramatic traffic increases that are expected by 2050 (Trafikverket 2012). According to Trafikverket (2012), the Government has emphasized the application of a four-step principle in investigations of the transport system. This principle functions as a support for addressing needs and shortcomings in the transport system and a guide for choosing sustainable measures for the transportation system. The main idea of the four-step principle is firstly to use existing infrastructure instead of building new infrastructure (Trafikverket 2017). Since most of the transportation in the future will use the current transport infrastructure, the operation and
maintenance of the existing system are prioritized (Trafikverket 2012). As a result, infrastructure managers are being compelled to employ effective and efficient maintenance programmes to meet new requirements, which include a higher level of resilience against failure, a more robust and available infrastructure, and cost reduction.

Railway infrastructure is a complex system consisting of track, signalling systems, electrical systems and telecommunications. Among these elements of this complex system, railway track is especially important. The main function of the railway track is to guide the trains in a safe and economic manner with the desired ride comfort (Iwnicki 2006). A large part of the railway infrastructure maintenance burden concerns track maintenance. Track maintenance is the total process of maintenance and renewal needed to make sure that the track meets the safety and quality standards (Esveld 2001). Railway track maintenance is a vast subject and covers different components which must be maintained from time to time (Zaayman 2017). Sweden’s railway network consists of around 16,500 kilometres of track, and the track maintenance and reinvestment cost for 2017 was 2,358 million SEK (Trafikverket 2017).

Railway track will settle as a result of permanent deformation in the ballast and underlying soil (Iwnicki 2006). The settlement is caused by the static and dynamic forces induced by traffic, which cause deviations from the designed geometry. Track geometry is a key aspect of railway construction and characterizes the track condition (Esveld 2001, Jovanovic 2004). Track geometry can be defined as the three-dimensional geometry of the track. A good track geometry quality provides good ride comfort for passengers and prevents the track from wearing too quickly (Trafikverket 2017).

When a geometry failure occurs, the consequences can be significant, including economic loss, operation interruption, damage to the railway asset and the environment, and a possible loss of human lives. In addition, a degraded track geometry will affect subsequent loading and damage of the track (Karttunen, Kabo et al. 2012). In order to protect against such consequences, maintenance actions are scheduled to retain the geometry condition in or restore it to an acceptable state with respect to ride comfort and the safety limits. In order to implement an effective and efficient maintenance programme, infrastructure managers are required to develop a resource capacity plan to address the requirements for manpower, machinery, maintenance time slots, etc. In addition, a master maintenance schedule needs to be developed to address the operational requirements and resource availability, as well as to facilitate opportunistic maintenance during the execution phase. These require prediction of the track geometry condition and estimation of the maintenance needs within a medium- and a long-term time horizon.
The track geometry degradation mechanism is rather complex and its complexity arises from the numerous factors affecting it, such as dynamic train loads, the substructure, material properties and environmental conditions (Esveld 2001, Guler, Jovanovic et al. 2011). Considering the complexity of the degradation process and the interaction of this process with the maintenance process, predicting the effect of employing different maintenance strategies on the track geometry condition is a difficult task (Andrews, Prescott et al. 2014, Quiroga, Schnieder 2012). Accordingly, there is a need to develop proper predictive models to tackle important aspects of railway track geometry degradation (Andrade, Teixeira 2015). Using such predictive models would facilitate and enhance the capability of making effective decisions for inspection and maintenance planning.

1.2. Statement of the problem

There are different challenges to overcome when modelling track geometry degradation over a line section. The first challenge is modelling the temporal evolution of the track geometry condition by considering the effect of maintenance. The majority of research studies show that the track geometry degradation within a maintenance cycle can be described by a linear or an exponential model with time or tonnage as the explanatory variable (Esveld 2001, Lichtberger 2005, Andrade, Teixeira 2011, Caetano, Teixeira 2016, Guler 2014, Khouzani, Golroo et al. 2016, Famurewa, Xin et al. 2015). However, the degradation path always has a break point after a maintenance action. Although maintenance actions will improve the track geometry condition, they cannot rejuvenate the geometry condition to an as-good-as-new state. Tamping as the main geometry maintenance activity is an imperfect maintenance and usually causes two changes in the degradation path, i.e. a break point in the degradation path by improving the geometry condition and a change in the degradation rate (Audley, Andrews 2013, Martey, Attoh-Okine 2018, Famurewa, Junitti et al. 2016, Arasteh Khouy, Schunnnesson et al. 2013). Therefore, in order to predict the track geometry evolution in multiple maintenance cycles, the effect of tamping on track geometry degradation must be assessed. The second challenge is modelling the spatial variation in the degradation parameters, along with the possible spatial dependencies. Due to inhomogeneity in the track structure and sub-structure, the load distribution, and the environmental conditions, different track sections degrade at different rates (Esveld 2001, Andrade, Teixeira 2015, Guler 2014, Lee, Hwang et al. 2018). In addition, there may be a spatial correlation between the degradation parameters for adjacent track sections (Andrade, Teixeira 2013). The reason for this correlation might be that the adjacent track sections may have similar operational, environmental and structural conditions. The other challenge is modelling the occurrence of track geometry isolated defects. Since a track
quality index (TQI) based on standard deviation of geometry parameters aggregates the track geometry measurements to represent the overall condition of a track section, it may not provide complete information about severe isolated defects in the track section (Alemazkoor, Ruppert et al. 2018). Isolated defects are short irregularities in the track geometry that can dramatically increase the dynamic forces between the wheel and rail, which in turn will accelerate the growth or occurrence of internal rail defects. The occurrence of severe isolated defects can cause comfort problems for passengers, damage to track components and an increase in the risk of derailment. Moreover, these defects cause unplanned maintenance activities, which may impose operational interruption, decrease the track availability and increase the maintenance cost (Andrade, Teixeira 2018). The ability to predict the occurrence of severe isolated defects allows railway infrastructure managers more efficiently to maintain the railway track and remain within the ride comfort and safety limits (He, Li et al. 2014, Sharma, Cui et al. 2018, Cárdenas-Gallo, Sarmiento et al. 2017, Arasteh khouy 2013).

The track geometry condition must be monitored to enable prediction of the time until the maintenance limits are reached, and thus facilitate the planning of maintenance work (UIC 2008). The frequency of inspection is highly important as it affects the detectability of geometry defects, the maintenance costs and the risk of train derailment. Therefore, an important task for infrastructure managers is to choose the most adequate inspection interval for track geometry measurement vehicles.

1.3. Purpose of the research

The purpose of the research performed for this thesis has been to develop methodologies and tools for the prediction of railway track geometry degradation, in order to facilitate and enhance the capability of making effective decisions for inspection and maintenance planning.

1.4. Objectives

In order to achieve the research purpose, the following objectives were set and pursued:

i. to develop a methodology to predict the track geometry degradation in spatio-temporal space,

ii. to propose an approach for investigation of the effect of inspection intervals on the track geometry condition.
1.5. **Research questions**

In order to fulfil the above-stated objectives, the following research questions were formulated:

1. How does one model the track geometry degradation over multiple tamping cycles by considering the spatial variation in the degradation parameters?

2. How does one predict the occurrence of isolated geometry defects at the defect-based level and the section-based level?

3. How does one develop a model for assessing the effect of different inspection intervals on the track geometry condition?

1.6. **Scope and delimitation of the present research**

Based on the available resources and according to the research purpose and objectives, as well as industrial interests, the scope and limitation of this study are as follows.

- The longitudinal level, as the main parameter, is analysed for track geometry degradation modelling and maintenance planning. The other track geometry parameters, i.e. alignment, cant and twist, are partly discussed.

- Tamping, as the main geometry maintenance activity, is investigated in this thesis. The effect of other geometry maintenance activities, e.g. stone blowing, falls outside the scope of this thesis.

1.7. **Structure of the thesis**

The contents of this thesis are divided into six chapters as follows.

**Chapter 1: Introduction** – This chapter presents a brief background to the research performed for this thesis and deals with the importance of developing geometry degradation models and maintenance programmes. In addition, the problems to be addressed in this research area are discussed. Moreover, the purpose and objectives of the research, the research questions and the limitations of the research are described, explained and outlined. The chapter also explains the extent of the theoretical framework, which is described in more detail in Chapter 2.
Chapter 2: Theoretical framework – This chapter provides a description of the state of the art concerning the main concepts and theories relating to this research. Theories are presented which support track geometry degradation modelling and maintenance planning. The theoretical framework has been used to achieve an understanding of the research area.

Chapter 3: Research methodology – This chapter describes some aspects of the research methodology, e.g. approaches, purposes, and data collection and analysis. The chapter also states the reasons for making the research choices related to these aspects. The selection of research methodologies has been performed based on the research purpose and objectives, the research questions, which are described in Chapter 1, and the theoretical framework presented in Chapter 2.

Chapter 4: Summary of the appended papers – This chapter provides a summary of the five appended papers and highlights the important findings of each appended paper.

Chapter 5: Results and discussion – This chapter presents the results and discussions of the conducted research work. The discussions are structured based on the research questions formulated.

Chapter 6: Conclusions, contributions, and future research – This chapter draws conclusions from the results of the conducted research work. The chapter also provides a summary of the research contributions, as well as some suggestions for further research.

References: A list of references is provided.

Appended papers: This part of the thesis consists of five appended papers. The contents of these papers are summarized in different chapters of the thesis, e.g. Chapter 3, Research methodology, Chapter 4, Summary of the appended papers, Chapter 5, Results and discussion, and Chapter 6, Conclusions, contributions, and future research.
Chapter 2

Theoretical framework

This chapter provides the theoretical framework and the basic concepts used within the research performed for this thesis.

2.1. Track geometry

Track geometry refers to the position of each rail or the track centre line in three-dimensional space (American Railway Engineering and Maintenance-of-Way Association (AREMA) 2006). According to EN 13848-1 (2008), track geometry quality is defined as the “assessment of excursions from the mean or designed geometrical characteristics of specified parameters in the vertical and lateral planes which give rise to safety concerns or have a correlation with ride quality”. The main geometry parameters used to assess the track quality and plan maintenance activities are the longitudinal level, alignment, gauge, cant, and twist. The track geometry parameters are described by a relative rectangular co-ordinate system centred to the track (SS-EN 13848-1: 2004+A1 2008).

Longitudinal level

This parameter is defined as the vertical deviation ($z_p$) of consecutive running table levels on the top of the left or right rail from the mean vertical position (the reference line) (SS-EN 13848-1: 2004+A1 2008) (see Figure 2.1).
Alignment

This parameter is defined as the horizontal deviation ($y_p$) of consecutive positions of point $P$ (see Figure 2.2) on the left or the right rail from the mean horizontal position (the reference line) (SS-EN 13848-1: 2004+A1 2008) (see Figure 2.2).

Cant

The cant or cross level is defined as the height difference between the adjacent running tables which is computed from the angle between the
running surface and a horizontal reference plane (SS-EN 13848-1: 2004+A1 2008) (see Figure 2.3).

![Figure 2.3 Cant (SS-EN 13848-1: 2004+A1 2008)](image)

**Gauge**

The track gauge is the shortest distance between the two adjacent rails from point P, and is measured at a position which is \( z_p = 14 \) mm below the running surface and between the perpendicular lines intersecting the running surface at the inner sides of the rail heads (SS-EN 13848-1: 2004+A1 2008) (see Figure 2.4).

![Figure 2.4 Gauge (SS-EN 13848-1: 2004+A1 2008)](image)

**Twist**

The twist is the algebraic difference between two cant values measured at positions located at a specified distance from each other.

### 2.2. Track geometry degradation

As a result of the influence of dynamic track loads, the track geometry degrades and deviates from the designed vertical and horizontal alignments.
The mechanism governing this phenomenon is rather complex (Esveld 2001). Railway track settlement occurs as a result of permanent deformation in the ballast and underlying soil. The severity of the settlement is mainly related to the quality and behaviour of the ballast, the sub-ballast and the subgrade. Track settlement is caused by various mechanisms of the ballast and subgrade behaviour which mainly concern either the densification of ballast and subgrade or inelastic behaviour of the ballast and subgrade materials (Iwnicki 2006). If all the positions in the track settle equally, there will be no track irregularities. However, due to inhomogeneity in the support conditions, track structure and load distribution, the track settlements are not uniform. Therefore, different points in the track settle by different amounts, which causes track irregularities (Esveld 2001). It also must be mentioned that when the track geometry has degraded, this will cause larger forces when a train is passing compared to running on a high quality track. This is one reason for observing an exponential growth in the degradation, since the degradation rate is proportional on degradation level.

### 2.3. Track quality index

In order to evaluate the track geometry, EN 13848–5 (2008) stipulates three ranges of wavelengths (λ):

- **D1**: $3 \text{ m} < \lambda \leq 25 \text{ m}$,
- **D2**: $25 \text{ m} < \lambda \leq 70 \text{ m}$,
- **D3**: $70 \text{ m} < \lambda \leq 150 \text{ m}$.

In order to consider the short wavelength defects, the lower bound of D1 must go down to 1 m. The running behaviour of the train is influenced by wavelengths and the amplitude of the track irregularities. The vehicle responses to track irregularities with short wavelengths are different from the responses to irregularities with long wavelengths (Choi, Um et al. 2013). Long wavelength track irregularities increase low-frequency oscillations of the train, and short wavelength irregularities cause vibrations and noise in the train and in the environment (Iwnicki 2006). Since the short wavelength defects can become dangerous when their amplitude is high, special attention must be paid to them (EN 13848–5 2008).

Different track geometry irregularities have different effects on the running behaviour of trains, the ride comfort and the safety parameters. Cant irregularities cause a vehicle body to roll, sway and twist, and longitudinal level irregularities can cause a vehicle body to pitch and bounce. The longitudinal level irregularities correlate with the vertical force, which is
related to the vertical acceleration felt by the passenger. Track alignment and gauge irregularities can lead to large lateral wheel and axle forces, which in turn may cause derailment or damage to the track structure (Choi et al. 2013). From a safety point of view, the derailment coefficient can be used as a key parameter to evaluate the possibility of train derailment. Considering Y and Q as the lateral and vertical forces, respectively, generated by a wheel, the ratio (Y/Q) characterizes the risk of derailment, corresponding to a flange climb derailment situation (EN 14363 2016). The value of the derailment coefficient is related to different track geometry irregularities. Longitudinal level irregularities are weakly correlated with running safety parameters such as derailment coefficients. However, longitudinal level irregularities strongly affect the vertical body acceleration, which is related to the ride comfort (Choi et al. 2013). The twist and gauge are particularly important as they have a considerable effect on the wheel loading and vehicle stability, which are related to derailment (EN 13848–5 2008).

According to the level of aggregation, different TQIs can be used to characterize the track geometry quality. Depending on the purpose of the analysis, the following three levels of aggregation of track geometry data are used: (1) the detailed level, which supports the analyses for local maintenance interventions and short-term maintenance planning, (2) the intermediate level, which supports the analyses for medium-term maintenance and renewal planning, and (3) the overview level, which supports the analyses needed for strategic decisions (EN 13848–6 2012). The three indicators mainly used to represent the track geometry quality are as follows: the standard deviation over a specific track length, the mean value, and the extreme value of isolated defects (EN 13848–5 2008). Isolated track geometry defects are mainly used to characterize the track geometry quality on the detailed aggregation level, while the standard deviation is commonly applied to represent the track geometry quality on the intermediate and overview aggregation levels (EN 13848–6 2012).

According to EN 13848–1: 2004+A1 (2008), as the minimum requirement, the track geometry parameters must be prescribed through the indicators provided in Table 2.1.

The standard deviation is the TQI most commonly used by European railway networks (EN 13848–5 2008, Sadeghi 2010). This indicator is used to show the variation or dispersion of track geometry measurement data. The standard deviation of a signal over a track section is calculated as follows:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}}$$

where N denotes the number of samples, $x_i$ is the current value of a signal, $\bar{x}$ is the mean value of a signal, and $SD$ denotes the standard deviation.
Commonly, the standard deviation is calculated for the longitudinal level and alignment irregularities in the wavelength range D1. It is recommended that one should calculate the standard deviation for the left and the right rail separately. The standard deviation is usually calculated for track sections with a length of 100 m or 200 m, but it can also be calculated for longer track sections, e.g. sections with a length of 1 km. It is also calculated for other parameters, including twist, gauge and cant (EN 13848-6 2012). The standard deviation is used to represent the variation of track geometry measurement data. A low standard deviation indicates that the geometry measurements are close to the mean and a high standard deviation indicates that the geometry measurements have a high variation around the mean value. It could be argued that a higher standard deviation reflects more irregularities in the track.

<table>
<thead>
<tr>
<th>Geometry parameter</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longitudinal level</td>
<td>Isolated defects that exceed a prescribed threshold. The standard deviation over a defined length, in the wavelength range 3 to 25 m.</td>
</tr>
<tr>
<td>Alignment</td>
<td>Isolated defects that exceed a prescribed threshold. The standard deviation over a defined length in the wavelength range 3 to 25 m.</td>
</tr>
<tr>
<td>Cross level</td>
<td>The absolute value.</td>
</tr>
<tr>
<td>Twist</td>
<td>Isolated defects that exceed a prescribed threshold. The standard deviation over a defined length.</td>
</tr>
<tr>
<td>Gauge</td>
<td>The identification of individual defects which exceed a prescribed threshold. The measured track gauge. The difference between the measured track gauge and the nominal track gauge. The mean track gauge over a specified distance. The variation of the track gauge over a specified distance.</td>
</tr>
</tbody>
</table>

By assigning different weights, the standard deviations of different track geometry parameters can be integrated for the purpose of assessment of the overall track geometry quality of a track section. In addition, in some cases the standard deviation of a combination of track geometry parameters can be used to represent the track geometry quality. The reason for using this kind of track quality index is that the level of the combined signal may give a better
picture of the vehicle behaviour than the picture obtained by comparing the individual signals.

Since the standard deviation aggregates the track geometry measurements to represent the overall condition of track sections, it may not provide complete information about severe isolated defects in the track sections. Isolated defects significantly influence the running safety parameters (EN 13848-6 2012). The number of severe isolated defects which have exceeded the specific maintenance limits, e.g. the intervention limit or alarm limit, in a certain length of track can represent the track geometry quality. The number of severe isolated defects in a given track section is commonly counted for the longitudinal level and alignment irregularities in the wavelength range D1, and for the twist, gauge, and cant (EN 13848-6 2012).

2.4. Maintenance limits

According to EN 13848-5 (2008), there are three limits for maintenance actions. Firstly, the immediate action limit (IAL) or safety limit refers to the value which, if exceeded, due to the potential risk of derailment, requires that a speed reduction or line closure be imposed before a corrective maintenance action is conducted. Secondly, the intervention limit (IL) or corrective maintenance limit refers to the value which, if exceeded, requires a corrective maintenance action before the immediate action limit is reached. Thirdly, the alert limit (AL) or preventive maintenance limit refers to the value which, if exceeded, requires that the track geometry be analysed for the planning of future maintenance actions.

The European standard EN 13848-5 (2008) provides the IALs, ILs and ALs for isolated defects and gives ALs for standard deviations. Generally, the track quality indicators based on the standard deviation of track geometry parameters are used to plan and perform preventive maintenance actions. On the other hand, the execution of corrective maintenance actions is based on the severity of isolated defects. Whenever the amplitude of an isolated defect exceeds the intervention limit or immediate action limit, corrective maintenance should be conducted on the track. Figure 2.5 shows the various maintenance action zones corresponding to limits associated with the aggregated TQIs and indicator of isolated defects. From this figure the following action zones can defined.

- No action zone: When the TQI is less than the AL and indicator of isolated defect is less than IL, no maintenance action will be performed.
• PM zone: When the TQI exceeds the AL and the indicator of isolated defect is less than IL, the track section(s) is monitored and considered for preventive maintenance actions.

• CM zone: When the indicator of an isolated defect is between the IL and IAL, CM is carried out on the track geometry without any operational restriction (i.e. speed reduction or line closure). Since this action is not based on any prior plan, its execution is more expensive than preventive maintenance.

• Safety zone: When the indicator of an isolated defect exceeds the IAL, corrective maintenance along with operational actions (speed reduction or line closure) is carried out on the track.

Figure 2.5 Different maintenance zones based on limits

The IALs are normative and take into account the track-vehicle interaction and the risk of unexpected events, whereas the ILs and the ALs are informative and are mainly linked with the maintenance policy. In fact, the ILs and ALs provided in EN 13848-5 (2008) reflect the common practice adopted by most European infrastructure managers. In alignment with the European standard EN 13848-5 (2008), Trafikverket (2015) have defined four main limits, namely the planning limit, the UH1 and UH2 limits, and the critical limit, as can be seen in Figure 2.6. The guidelines issued by Trafikverket (2015) express the intervention limit as a range rather than a discrete value. They state that track irregularities that exceed the UH1 limit must be assessed for conducting maintenance before the UH2 limit is exceeded. For track irregularities exceeding the UH2 limit, maintenance actions must be planned without unnecessary delay. Therefore, track irregularities must be corrected before the UH2 limit is reached. Table 2.2 shows the relations between the limits specified in EN 13848-5 (2008) and Trafikverket’s guidelines (Trafikverket 2015).
Table 2.2 Maintenance limits defined in the European standard EN 13848-5 and Trafikverket’s guidelines

<table>
<thead>
<tr>
<th>European standard EN 13848-5</th>
<th>Trafikverket’s guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert limit</td>
<td>Planning limit</td>
</tr>
<tr>
<td>Intervention limit</td>
<td>UH1 limit (lower bound for corrective maintenance)</td>
</tr>
<tr>
<td></td>
<td>UH2 limit (upper bound for corrective maintenance)</td>
</tr>
<tr>
<td>Immediate action limit</td>
<td>Critical limit</td>
</tr>
</tbody>
</table>

Based on the above-mentioned limits, geometry defects can be classified according to their severity into three groups, i.e. UH1, UH2 and critical defects. UH1, UH2 and critical defects occur when track irregularities exceed the UH1, UH2 and critical limits, respectively. Examples of these three categories of defects are shown in Figure 2.6.

![Figure 2.6 Examples of UH1, UH2 and critical defects of the longitudinal level](image)

2.5. Track geometry measurement process

The main objectives of measuring the track geometry are to locate the occurrence of severe isolated defects and prioritize the corrective maintenance actions required to rectify them, to plan the preventive maintenance actions, to evaluate the effectiveness and quality of conducted maintenance activities, and to establish the effectiveness and adequacy of financial expenditure...
The purpose of track geometry measurement vehicle is to measure and record the track geometry in real time and to display the measurements obtained (American Railway Engineering and Maintenance-of-Way Association (AREMA) 2006). The vehicle measures various geometry parameters, e.g. the longitudinal level, alignment, cant, gauge, twist, curve radius, and gradient at a specific sampling interval. The measured data are available in three forms, i.e. on-board network viewing of measurements, on-board real-time reports, and off-board post-processed reports. On-board real-time reports can be used to address severe isolated defects and comprise strip charts and exceptions reports. The strip chart is a graphical representation of the overall track geometry condition in one picture and displays all the track geometry measurements and the corresponding maintenance limits. The strip chart shows when a particular measurement approaches or exceeds the maintenance limits. When the track geometry measurements exceed the maintenance limits, a geometry defect will appear in the exceptions report. The purpose of the exceptions report is to provide a list of all the geometry defects found by the track geometry measurement vehicle during the inspection of the track. The basic information provided in the exceptions report is the defect name, the position of the defect, the defect length, defect magnitude, maintenance limit, and maximum allowable speed. Using the strip charts and the exceptions reports will aid the process of conducting corrective maintenance activities to avoid the occurrence of severe isolated defects which may cause catastrophic consequences. In addition to on-board reports, the geometry measurements are post-processed into different standard reports. These reports can be used for further analysis of the track geometry condition to support track maintenance decision making to achieve a robust and cost-effective maintenance plan. Analysing the track geometry measurements and aggregating them into a TQI can aid the planning of track geometry maintenance activities. By performing a statistical analysis on the evolution of the TQI, the trends in the track geometry degradation can be identified. In addition, the measured data can be used to identify potentially hazardous track geometry conditions.

### 2.6. Tamping process

Tamping is a main maintenance action applied to restore the geometrically correct track position (Lichtberger 2005). The main factors affecting the tamping quality are the frequency and amplitude of the tine vibrations, the tamping pressure, the squeezing speed and the tamping time, as well as the tamping depth (Lichtberger 2005, Zaayman 2017). A schematic illustration of the tamping process is provided in Figure 2.7.
As shown in Figure 2.7, the tamping process consists of the following four main steps.

- **Step 1**: The tamping machine moves to the tamping site, where it then stands still with the tamping tines straddling the sleeper to be tamped.
- **Step 2**: In conjunction with the measuring system, the lifting and lining units lift the sleepers to a predetermined height while correcting the vertical alignment defects, and position the track to correct the horizontal alignment.
- **Step 3**: The tamping units are lowered and the vibrating tamping tines are inserted into the ballast, stopping at a predetermined depth. The vibration significantly reduces the force required for the penetration of the tamping tines into the ballast.
- **Step 4**: The tamping tines perform a squeezing motion and compact the ballast under the sleepers in the space created by the lifting process (Zaayman 2017).

Tamping is nonsynchronous with directional vibration. This means that the individual tamping tines are moved with the same force, independently of the path. This guarantees a uniform and homogenous tamping of the track. Tamping is normally conducted to remedy track geometry defects with a wavelength of 3 to 25 m in the smoothing mode and defects with a wavelength greater than 25 m in the design mode. Short wavelength geometry defects up
to 3 m are mainly related to corrugation or a discontinuity in the rail caused by a weld or joint, and therefore these defects are corrected using grinding or weld straightening (Esveld 2001).

Tamping machines can be classified into different categories based on their speed, which is mainly influenced by the number of sleepers that can be tamped simultaneously. Other categories used for the classification of tamping machines are continuous action tamping machines and intermittent action tamping machines. Intermittent action tamping machines tamp a group of sleepers, moving along for a couple of metres and then stopping to perform tamping on the next group of sleepers. Continuous action tamping machines are equipped with a tamping trolley placed under the machine itself, in such a way that the trolley can move freely backward and forward while the tamping machine moves at a constant speed. This results in a higher tamping speed, a higher level of comfort for the operators in the tamping machine, and a lower energy cost (Esveld 2001).

2.7. Predictive models for track geometry maintenance

Maintenance can be defined as the combination of all the technical, administrative and managerial actions to be performed during the life cycle of an item to retain it in, or restore it to, a state in which it can perform the required function (CEN 2010). Maintenance is an efficient way to guarantee a desirable level of reliability during the useful life of an asset. Maintenance strategies can be divided into two main categories, i.e. corrective maintenance and preventive maintenance. The oldest maintenance strategy historically is corrective maintenance, which is also called reactive maintenance and takes place after a defect or failure occurs (Jardine, Lin et al. 2006). In traditional preventive maintenance, maintenance activities are scheduled at discrete time intervals to prevent failures, based on either calendar time or usage. The main advantage of preventive maintenance is that it can be planned beforehand and performed at proper times (Budai-Balke 2009). However, traditional preventive maintenance schedules maintenance activities regardless of the asset condition, which can lead to a high maintenance cost and lower achieved availability (Lu, Tu et al. 2007). Therefore, to increase the effectiveness of preventive maintenance, condition-based maintenance strategies are being implemented. In a condition-based maintenance strategy, maintenance decisions are made based on the actual performance of the asset as determined through condition monitoring, inspection and testing (CEN 2010). Eventually, in the predictive maintenance strategy, the failure times are predicted to schedule maintenance activities proactively. According to (CEN 2010), predictive maintenance is condition-based maintenance.
undertaken following a prediction based on repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item. Condition-based maintenance can be applied to perform both diagnostics and prognostics. Diagnostics consists of the detection, isolation and identification of faults when they occur. Prognostics focuses on the prediction of faults before they occur (Jardine et al. 2006). Generally, prognostic methods can be classified into data-driven, physics-based and hybrid approaches (AN, KIM et al. 2015). Data-driven approaches extract information from training data to identify the characteristics of the current degraded state and to predict the future condition of the asset. Data-driven approaches can be divided into two categories, i.e. artificial intelligence approaches and statistical approaches (Jardine et al. 2006, AN et al. 2015). The prediction performance of the statistical approaches is dependent on a number of factors, e.g. the quality of the model, the data quality and the degree to which the degradation process is deterministic, stochastic or chaotic. Physics-based approaches rely on the assumption that a physical model describes the degradation process, and they need mechanistic knowledge and theory relevant to the monitored asset. The hybrid approach combines the data-driven and physics-based approaches to improve the prediction performance (AN et al. 2015). Using prognostic approaches, the remaining useful life (RUL) of the asset can be predicted, which in turn can be used to predict the maintenance needs.

In order to minimize transport disruptions and their cost and keep safety and ride comfort at a desirable level, railway infrastructure managers want to shift from a reactive maintenance strategy to a proactive maintenance strategy. In addition, with the current advancements in the field of railway track geometry measurement technologies, a large amount of event data and condition monitoring data is available, and degradation monitoring has become more practicable as a support for a condition-based maintenance strategy. Furthermore, these data can be used for the purpose of diagnosis and prognosis of the track geometry condition. Using diagnostic and prognostic approaches, a better understanding of track geometry degradation can be achieved, which in turn can improve the life cycle management of the track system (Xin, Famurewa et al. 2016). Track geometry degradation models provide basic information for making decisions as to when and where a maintenance activity or an inspection should be performed. Descriptions of the track geometry degradation process must be as close as possible to the real degradation of the track.

The track is one of the critical linear assets of the railway infrastructure. In order to plan maintenance activities for a linear asset, it is necessary to define the location of a point or a section along the asset (Bergquist, Söderholm 2015). In contrast to point assets, the length of linear assets is critical for maintenance planning. In the case of this type of asset, since both the location
and the time for monitoring the degradation of the linear asset are important, the monitoring scheme requires spatiotemporal information (Bergquist, Söderholm 2015). From a statistical perspective, the condition monitoring data of linear assets are usually spatially auto-correlated. Therefore, in order to model the track geometry degradation, the spatial variation in the track geometry degradation parameters and the possible spatial dependency must be considered.
Chapter 3

Research methodology

This chapter describes some aspects of the research methodology, e.g. approaches, purposes, data collection, and analysis.

3.1. Research purpose

Research can be defined as a scientific and systematic search for solutions to a specific problem (Kothari 2011). Research consists of creative and systematic work carried out to increase the stock of knowledge – including knowledge of humans, culture and society – and the use of available knowledge to devise new applications (OECD (Organization for Economic Cooperation and Development) 2015). Research methodology refers to all the principles, procedures and practices that govern research (Marczyk, DeMatteo et al. 2005). Research methodology comprises not only the research methods applied in the context of the research to solve a certain problem, but also the logic behind the methods (Kothari 2011).

The purpose of research is to find answers to questions by applying scientific procedures (Kothari 2011). Depending on its purpose, research can be classified into exploratory, descriptive and explanatory research (see Table 3.1 for details).

The research methodology selected to fulfil the purpose of the research conducted for this thesis is a combination of exploratory, descriptive and explanatory approaches. In the initial stage, an exploratory research approach was used to become familiar with track geometry degradation modelling and maintenance planning, to identify the existing challenges and opportunities, and to obtain new insight into the research field. The knowledge gained from
the exploratory research was used to identify the research gaps and to formulate RQ1, RQ2 and RQ3. In addition, the exploratory research provided the knowledge required to select and develop the case studies and to identify the types of available data. In the second stage, descriptive and explanatory approaches were used, involving the collection of field data and the employment of statistical models. The descriptive and explanatory approaches were employed to determine how to model track geometry degradation, how to predict tamping recovery, how to predict the occurrence of isolated defects, and how to investigate the effect of different inspection intervals on the track geometry condition.

Table 3.1 Three types of research based on the research purpose (Neuman 2013)

<table>
<thead>
<tr>
<th>Exploratory</th>
<th>Descriptive</th>
<th>Explanatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Become familiar with the basic facts, setting and concerns.</td>
<td>Provide a detailed, highly accurate picture.</td>
<td>Test a theory’s predictions or principle.</td>
</tr>
<tr>
<td>Create a general mental picture of conditions.</td>
<td>Locate new data that contradict past data.</td>
<td>Elaborate and enrich a theory’s explanation.</td>
</tr>
<tr>
<td>Formulate and focus on questions for future research.</td>
<td>Create a set of categories or classify types.</td>
<td>Extend a theory to new issues or topics.</td>
</tr>
<tr>
<td>Generate new ideas, conjectures or hypotheses.</td>
<td>Clarify a sequence of steps or stages.</td>
<td>Support or refute an explanation or prediction.</td>
</tr>
<tr>
<td>Determine the feasibility of conducting research.</td>
<td>Document a causal process or mechanism.</td>
<td>Link issues or topics with a general principle.</td>
</tr>
</tbody>
</table>

3.2. Research approach

Research can be either applied or fundamental (basic). The main aim of applied research is to find a solution to an existing practical problem. The central aim of fundamental research is to increase the existing scientific knowledge so that future research activities can be based on the increased body of knowledge (Kothari 2011).

The problems dealt with in the present research were defined based on the needs and requirements of railway companies. This thesis concerns applied
research whose purpose is to develop methodologies and tools for the prediction of railway track geometry degradation, in order to facilitate and enhance the capability of making effective decisions for inspection and maintenance planning.

Research can also be categorized as quantitative and qualitative. The main characteristics of these two types of research are presented in Table 3.2.

<table>
<thead>
<tr>
<th>Quantitative</th>
<th>Qualitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure objective facts</td>
<td>Construct social reality, cultural meaning</td>
</tr>
<tr>
<td>Focus on variables</td>
<td>Focus on interactive processes, events</td>
</tr>
<tr>
<td>Reliability is key</td>
<td>Authenticity is key</td>
</tr>
<tr>
<td>Value-free</td>
<td>Values are present and explicit</td>
</tr>
<tr>
<td>Independent of context</td>
<td>Situationally constrained</td>
</tr>
<tr>
<td>Many cases, subjects</td>
<td>Few cases, subjects</td>
</tr>
<tr>
<td>Statistical analysis</td>
<td>Thematic analysis</td>
</tr>
<tr>
<td>Researcher is detached</td>
<td>Researcher is involved</td>
</tr>
</tbody>
</table>

The present research uses both the quantitative and the qualitative approach. The qualitative approach was used to explore drivers, challenges and issues in railway track geometry degradation modelling and maintenance planning. The quantitative approach was used to predict track geometry degradation and the effect of tamping on the track geometry condition, and to investigate the effect of different inspection intervals on the track geometry condition.

### 3.3. Data collection and data analysis

Data can be defined as facts which can be communicated and stored (Spender 1996). Data can be categorized as primary data and secondary data. Primary data refer to those data which are collected by the researcher for the purpose of study. Secondary data refer to those data which were already collected by someone else before being used by the researcher (Kothari 2011). In the present research, mostly secondary data were used for the purpose of the research. Data can also be divided into theoretical and empirical data. Both theoretical data and empirical data were collected in the present research to address the research questions. The theoretical data were collected from a literature review and the empirical data were collected from interviews, as well as historical data concerning operation and maintenance.
An effective review of prior and relevant literature plays a vital role in any scientific research by creating a firm foundation for advancing knowledge (Webster, Watson 2002). The review conducted for the present research covered different theories and practices used in track geometry degradation modelling and maintenance planning. The documents and research papers were collected from different databases, e.g. Scopus, Emerald, Google Scholar, and Web of Science. The documents studied in the review include books, peer-reviewed journal papers, papers in conference proceedings, Licentiate and PhD theses, standards (e.g. Series EN 13848 “Railway applications – Track – Track geometry quality”), and technical reports. Different keywords and their combinations were used in the literature review, for example railway track geometry, degradation modelling, maintenance planning, predictive models, machine learning, tamping, track quality index, and isolated defect.

The objective of the interviews held during the present research was to consider the opinions and experiments of the personnel and experts involved in railway track geometry maintenance, to complement the literature review and data analysis. The outcomes of the literature review and the data analysis conducted by the author were considered as the basis for the interviews. The main issues discussed in the interviews were practical issues concerning track geometry maintenance planning, data cleaning and pre-processing, and interpretation of the results of the data analysis. The interviewees were experienced practitioners in the field of railway maintenance from Trafikverket and Infranord. Moreover, discussions were held with other experts in the field of railway maintenance. In addition, the practitioners involved from Trafikverket actively took part in enhancement of the proposed methodologies, e.g. by discussing, reading and making comments, and providing valuable and applicable documents and data.

The collected data can be categorized into event data and condition monitoring data. The event data include information about the activities performed on the track (e.g. installation, overhaul, minor repair, and preventive maintenance) and their causes. The condition monitoring data are the geometry measurements which are related to the health condition of the railway track. Historical data on track geometry degradation and maintenance were collected from Trafikverket’s databases, i.e. Optram (track geometry measurement database), BESSY (inspection report system) and BIS (asset register system). Optram is a maintenance decision-support system which can graphically show the results of track geometry measurements. Only measurement data gathered after 2007 are available in this database. The system also provides functionality for analysis and displays data trends. BESSY¹ is an inspection report system which contains information on

¹ Trafikverket’s inspection report system
inspections and the types of actions performed after inspection comments (Nissen 2009). BIS² is an asset register database which contains information on infrastructure and facilities, agreements, the history of tamping (including data such as the location of tamped sections, the length of tamped sections, the tamping dates, etc.), the history of grinding, and curves. The historical data on track geometry degradation and maintenance were collected for line section 414, which is the part of the Main Western Line in Sweden (Västra Stambanan) running between Järna and Katrineholm Central Station.

Through data analysis, information is extracted from data. Data analysis includes the aspects of examining, categorizing, tabulating, and recombining the evidence to address the propositions of a study (Yin 2017). The track geometry data, maintenance history data and other facts collected were analysed using different methods to discover useful information for track geometry degradation modelling and maintenance planning. Since data always contain errors, data cleaning is an important step in data processing. Therefore, in order to remove corrupt and inaccurate records from the data set, data cleaning was performed. This process was carried out by checking the track geometry data and consulting railway track maintenance experts. A preliminary analysis was also performed to describe the key features of the data, provide an overview of the information content of the data, and prepare the data for further analysis and modelling. A variety of models are available in the literature to analyse track geometry data for a better understanding and interpretation of track geometry degradation (see Paper 1). In the present research, statistical analysis was performed to discover underlying patterns and trends in the track geometry data, to test the hypotheses, and to determine the relationship between the predictors and the response variables. Statistical methods were used to develop models for the prediction of track geometry degradation, tamping recovery, and the occurrence of isolated defects.

3.4. Research process

The research process consists of a series of steps and their desired sequencing to carry out the undertaken research effectively (Kothari 2011). An overview of the research process used in the present research is presented in Figure 3.1. The different steps in the research process are described and discussed in the following.

² Baninformationssystem (in English: Track Information System)
Step 1. Defining the research problems: The problems dealt with in the applied research documented in this thesis were formulated in cooperation with Trafikverket. This contributed to an understanding of what was needed in practice and what the requirements of usefulness were. In this step, an extensive literature study was performed to identify the issues and challenges, as well as the gaps in the current knowledge, concerning track geometry degradation modelling and maintenance planning. The outcome of this literature study resulted in the formulation of three sharpened research questions and a sharpened research purpose. Thereafter, a tentative research methodology was constructed. The outcome of this step is mainly summarized in Chapter 1 of this thesis.

Step 2. Specifying the research purpose: The stated research purpose was formulated according to the stated research problem and based on the stated requests of Trafikverket. Since there were interactions between the various steps, the research purpose was modified as new ideas and good reasoning...
were presented through other steps. Hence, the purpose of the research was not fixed from the very beginning of the research, but open to revision, and was subjected to continuous improvement.

**Step 3. Specifying the limitation and constraint conditions:** Based on the available resources (such as time), and according to the research purpose and objectives and the industrial interests, the scope and limitation of the research were defined. Based on the empirical evidence, the research literature reviewed and some rational and logical reasoning gained from the other steps of the research, the limitation and constraint conditions were changed later.

**Step 4. Identification of the alternative methodologies:** This step dealt with the need for further investigation of the alternative methodologies which were applicable to addressing the problems associated with track geometry degradation and restoration modelling, prediction of the probability of the occurrence of isolated defects, and investigation of the effect of different inspection intervals on the track geometry condition. The alternative methodologies that were investigated further in this step had been identified in the first and second steps. Creating appropriate alternative methodologies was a creative and explorative process that included considering the consequences of the various methodology choices.

**Step 5. Construction of the methodology of our choice:** A further literature study was performed concerning track geometry degradation, the effect of tamping on the track geometry, the occurrence of isolated defects, and inspection optimization. In this stage, methodologies were developed for modelling track geometry degradation, tamping recovery, the occurrence of isolated defects, and the effect of the inspection intervals on the track geometry condition. The aim was to develop methodologies that would address the associated research questions and fulfil the purpose and objectives of the research according to the limitation and scope.

**Step 6. Verification of the proposed methodology:** The aim of this step was to obtain an idea of the applicability, feasibility and effectiveness of the proposed methodology. The verification step was conducted with the collaboration of railway maintenance experts through an iterative process, by matching the developed methodologies to specific cases. In addition, colleagues of the author at Luleå University of Technology gave comments on the research design and worked with the appended papers at seminars, which strengthened the verification.

**Step 7. Evaluation of the proposed methodology:** In this process, the weaknesses and gaps of the methodologies were identified through interviews with experienced practitioners, together with some documents from their companies, which contributed to improving the proposed methodologies. From these evaluations, conclusions were drawn regarding the applicability and
effectiveness of the proposed methodologies in practice. This step played an important role for the continuous process of improving the proposed methodologies.

*Step 8. Application of the proposed methodology:* Based on the evaluation phase, the suggested improvements were applied, and the improved methodology was verified again to assure its applicability and effectiveness.
Chapter 4

Summary of the appended papers

This chapter provides a summary of the five papers appended to the thesis, and describes their contribution towards answering the research questions (see Table 4.1). Further information can be found in the appended papers.

Table 4.1 Relationship between the appended papers and research questions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>How does one model the track geometry degradation over multiple tamping cycles by considering the spatial variation in the degradation parameters?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>RQ 2</td>
<td>How does one predict the occurrence of isolated geometry defects at the defect-based level and the section-based level?</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>RQ 3</td>
<td>How does one develop a model for assessing the effect of different inspection intervals on the track geometry condition?</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

4.1. Paper 1

Paper 1 presents a review of the literature on track geometry degradation modelling and maintenance planning, as well as suggesting possible knowledge gaps from the point of view of both researchers and practitioners.
This paper synthesises 102 published documents into a summary of what is known and has been validated, and what is not known, to provide a comprehensive picture of the state of the art concerning different issues in this field, as a major step towards further developments. In particular, the following three aspects are addressed in Paper 1: (1) track quality indicators, (2) track geometry degradation modelling, and (3) maintenance planning and scheduling. Subsequently, various challenges and opportunities with regard to these aspects are identified, to catalyse the development of research on these topics in the future. To select a proper index for characterizing track geometry quality, two main issues must be considered, namely the level of aggregation and the combination of track geometry parameters. By synthesizing the information in the literature into a summary, the following four main issues were identified as important for modelling the track geometry evolution of a line section over multiple maintenance cycles: (1) the temporal changes in the geometry degradation within a maintenance cycle, (2) the effect of shock events on the degradation path, (3) the effect of tamping on the geometry degradation pattern, and (4) the section-to-section variation in the degradation parameters, along with possible spatial dependence. Regarding maintenance planning, Paper 1 shows that in recent years there has been an obvious trend towards using simulation-based methods for identifying the effect of different maintenance strategies on the track geometry condition. This will change the current practices adopted in the field of maintenance planning and lead to more effective and efficient track geometry maintenance decisions. The main objective of simulation-based methods is to find the optimal inspection interval, maintenance limits and renewal strategy with respect to direct and indirect maintenance costs. The last phase of track geometry maintenance planning is creating the optimum maintenance scheduling, with a view to finding the right times for performing maintenance activities. Various approaches using different objective functions are proposed in the literature for track maintenance scheduling. The following tasks should be carried out to obtain a proper track maintenance scheduling: defining an objective function with respect to the RAMS parameters and the track life cycle cost; applying robust degradation and recovery models; considering the practical constraints of maintenance activities; and taking advantage of the concept of opportunistic maintenance in developing a scheduling model.

4.2. Paper 2

Paper 2 presents a study in which the effect of tamping on the track geometry degradation was analysed using data collected from line section 414 between Järna and Katrineholm Central Station. It should be noted that whilst tamping will improve the track geometry condition, it cannot
rejuvenate the geometry condition to an as-good-as-new state. Tamping is an
imperfect maintenance action and will cause two changes in the track
groupy\e condition and a change in the degradation rate after tamping. Hence,
estimation of the effect of tamping on the track geometry degradation is a vital
input for any track geometry predictive model. A preliminary analysis showed
that even for track sections with a similar track geometry condition before
tamping, there was a significant variation in the recovery values after
tamping interventions. In order to predict the recovery value after tamping
intervention, a probabilistic model was developed in the study documented in
Paper 2. The results show that the Weibull distribution and lognormal
distribution are the best-fitted distributions for the recovery values after
tamping for the standard deviation of the longitudinal level and the standard
deviation of the alignment, respectively. In addition, the results show that the
recovery values simulated by the proposed model are very close to the real
recovery values after tamping interventions. In order to measure the linear
relation between the recovery of the longitudinal level and the alignment
irregularities, the Pearson correlation coefficient was applied. The results of
the correlation analysis show a weak correlation between the recovery of the
longitudinal level and the alignment irregularities. This means that tamping
will not affect the different track geometry parameters equally. In addition,
the effect of tamping on the degradation rate was also examined in the study
presented in Paper 2. The Wiener process and linear regression were used to
model the track geometry degradation. According to the results of the case
study, there is no significant difference between the degradation rates
estimated by the Wiener process and those obtained using linear regression.
The paired t-test was applied on the degradation rates before and after
tamping interventions to evaluate the effect of tamping on the geometry
degradation rate. The results show that tamping will increase the geometry
degradation rate on average, which is related to the destructive effect of
tamping on ballast.

4.3. Paper 3

Paper 3 proposes a two-level piecewise linear model to characterize the
track geometry degradation and restoration with possible spatial dependence.
At the first level, the track geometry degradation of each track section is
modelled using a piecewise linear model with break points at the maintenance
times. Then, two multivariable linear regression models are developed to link
different covariates with the responses, i.e. the recovery values and changes
in the degradation rates after a tamping action. At the second level,
autoregressive moving average (ARMA) models are used to capture the spatial
dependence in the parameters of the regression lines indexed by their
locations. Since the hierarchical model tends to be complex and difficult for the statistical estimation, a simple ad hoc procedure is performed to estimate all the unknown parameters. To illustrate the model, a comprehensive case study is presented using data from line section 414 between Järna and Katrineholm Central Station. A preliminary analysis of the evolution of the standard deviation of the longitudinal level shows that for a small number of maintenance cycles there is an unusual trend in the degradation path. This phenomenon is attributed to the occurrence of shock events and is discussed in the paper. The results of the case study show that there is a spatial autocorrelation between the degradation parameters of track sections. In addition, an analysis of the results of the case study shows that the correlation structures of the processes estimated by the ARMA models are very close to those of the true processes.

4.4. Paper 4

Paper 4 presents a study in which data-driven models were developed to predict the occurrence of severe isolated geometry defects using the data collected from line section 414 between Järna and Katrineholm Central Station. In order to model the track geometry degradation, the evolution of the amplitude of the longitudinal level defects within a maintenance cycle was modelled using linear regression. However, if a shock event has occurred in the degradation path, using one linear model to study the data leaves the data poorly explained by the model. Therefore, for those sections where a shock event has occurred, two separate lines are fitted using the piecewise linear function to model the geometry degradation. In addition, the tamping effectiveness in terms of rectifying the longitudinal level defects was analysed. In order to predict the probability of the occurrence of UH2 defects, a section-based model was developed using binary logistic regression, and in this model the kurtosis of the longitudinal level is also considered as an explanatory variable. The results of the case study documented in Paper 4 show that the linear model is an appropriate choice for modelling the degradation pattern of longitudinal level defects. Furthermore, it has been found that the standard deviation and kurtosis of the longitudinal level are statistically significant predictors of the occurrence of UH2 defects in a given track section. The applied binary logistic regression model shows a satisfactory performance in predicting the occurrence of UH2 defects over the test set, which further justifies the proposed approach.
4.5. Paper 5

Paper 5 proposes an integrated statistical model to investigate the effect of the inspection intervals on the track geometry condition. The track geometry degradation, the occurrence of isolated defects and the tamping recovery are modelled and integrated for long-term prediction of the track geometry condition under different inspection intervals. Using the proposed model, a prediction is made of the percentage of time spent by the track in the different track geometry states under various inspection intervals. By considering the maintenance limits, the following three geometry states were defined.

- State 1: There is no UH2 defect or critical defect in the track section.
- State 2: There is at least one UH2 defect in the track section, but there is no critical defect in that section.
- State 3: There is at least one critical defect in the track section.

In order to characterize the track geometry degradation and restoration, a piecewise exponential model is applied which considers break points at the maintenance times. A multivariable linear regression model is used to link different covariates to the tamping recovery. In addition, ordinal logistic regression is applied to predict the probability of the occurrence of geometry defects which have exceeded the upper bound of the intervention limit and the immediate action limit (critical limit) in a given track section. Owing to the complexity in the degradation process and the maintenance process, predicting the effect of employing different inspection intervals on the track geometry condition is a difficult task. Therefore, the Monte Carlo simulation technique is used to estimate the percentage of time spent in the different track geometry states. Prediction of the percentage of time spent by a track section in different geometry states is very important, as it is a primary input for predicting the probability of the occurrence of temporary speed reductions and the risk of derailment. The results of the Monte Carlo simulations performed in the study presented in Paper 5 show that the inspection frequency has a significant effect on the length of time spent by a track section in different geometry states. Based on the results obtained, changing the inspection interval from one month to six months increases the percentage of time spent by a track section in state 2 and state 3 by 3.22% and 0.25%, respectively.

---

3 UH2 defects
Chapter 5

Results and discussion

This chapter discusses and draws conclusions from the results of the conducted research work. The structure of the chapter is based on the research questions (RQs) formulated.

5.1. Results and discussion related to RQ 1

RQ 1: How does one model the track geometry degradation over multiple tamping cycles by considering the spatial variation in the degradation parameters?

RQ 1 is mainly answered in Papers 1, 2 and 3, and partly in Paper 5. Paper 1 presents the results of a constructive literature review conducted to evaluate the available literature and to identify the issues, challenges and current knowledge gaps concerning track geometry degradation modelling. By synthesizing the information in the literature into a summary, three main issues were identified as important for modelling the track geometry evolution of a line section over multiple maintenance cycles: (1) the temporal changes in the geometry degradation within a maintenance cycle, (2) the effect of tamping on the geometry degradation pattern, and (3) the section-to-section variation in the degradation parameters along with the possible spatial dependencies.

As discussed in Paper 1, track degradation models can be classified into mechanistic and data-driven models. Mechanistic models are mainly based on laboratory studies and their purpose is to explain the physical reactions in the track which result in track degradation. The main advantage of the mechanistic approach is that it leads to a good understanding of the relationship between track settlement and vehicle loading. However, the drawback of the mechanistic approach is its inability to cope with the uncertainty concerning the track degradation due to the heterogeneous factors.
influencing that degradation. As has been widely mentioned in the literature, there is section-to-section variation in the degradation parameters, due to the variability of the track structure, traffic conditions, environmental conditions, and maintenance history. Data-driven models use geometry measurement data to learn the track degradation behaviour. Owing to the ability of data-driven models to capture the uncertainty in the degradation modelling process, these models have been widely used for track geometry degradation modelling. Examples of such models are linear models, exponential models, machine learning models, and stochastic processes. Linear and exponential models, with time or tonnage as the explanatory variable, are the most popular models in the literature due to their simplicity and ability to represent the underlying degradation path. In order to model the geometry evolution of a track section over multiple tamping cycles, one needs to consider the effect of tamping on the degradation pattern. Whilst tamping will improve the track geometry condition, it cannot rejuvenate the geometry condition to an as-good-as-new state. In order to predict the recovery value after a tamping intervention, linear regression with the track geometry condition before tamping as the explanatory variable has been widely used in the literature. However, the outcome of the literature review presented in Paper 1 shows that in some cases there is a high variation in the recovery value after a tamping intervention, even for track sections with a similar track geometry condition before tamping. It has been observed in the literature that conducting tamping interventions damages the track ballast and in the long term increases the degradation rate. This means that conducting more tamping interventions is not necessarily the best solution to apply in the long term to control and manage the geometry degradation.

The effect of tamping on the degradation parameters was investigated in the study presented in Paper 2, which dealt with a real case in a line section of the Swedish railway network. The Wiener process was applied to model the geometry degradation. A stationary Wiener process is particularly well suited to modelling the evolution of a degradation mechanism characterized by a linear increase over time with random noise. It must be noted that, although the Wiener process is not monotonously increasing, the mean degradation is linearly increasing over time. The drift coefficient of the Wiener process is closely related to the progression of the degradation and characterizes the geometry degradation rate. Moreover, in the study documented in Paper 2, linear regression was used to model the track geometry degradation. The degradation rates obtained by the two models are used to analyse the effect of tamping on the degradation pattern. The results show that the values of the drift coefficient for the Wiener process and the slope for the linear regression model are very close to each other. This means that there is no significant difference between the degradation rates obtained using the Wiener process and the linear regression model. Figure 5.1 displays a graphical comparison
of the results obtained using the Wiener process and linear regression to model the track geometry degradation.

![Graph showing track geometry degradation modelling using linear regression and the Wiener process](image)

*Figure 5.1 Track geometry degradation modelling using linear regression and the Wiener process*

As can be seen in Figure 5.1, the mean degradation values (the standard deviation of the longitudinal level (SDLL)) obtained using the Wiener process and linear regression have a similar path. However, the two models differ in their representation of the uncertainty in the degradation path. In the regression model the basic idea is that degradation follows a deterministic trajectory where there is some random noise at each observation point. In such a model the deviation from the trajectory is the same all the time. Using Wiener process can characterize a non-monotonic degradation process, and provide a good description of track geometry degradation due to an increased or reduced intensity of the use. For a Wiener process with positive drift, the first-passage-time of the degradation path to a fixed failure threshold level follows an inverse Gaussian distribution.

In order to evaluate the effect of tamping on the geometry degradation rate, the paired t-test was applied on the degradation rates before and after tamping interventions. The results of the paired t-tests applied on the degradation rates show that there is an increase in the mean of the
degradation rates after a tamping intervention (see Paper 2 for more details). Hence, it is concluded that tamping will increase the geometry degradation rate on average. This result is in alignment with previous studies which have discussed the destructive effect of tamping on the ballast.

In order to predict the recovery value after tamping, a probabilistic approach was applied in the study presented in Paper 2. The recovery is measured in terms of the change in the standard deviation of geometry parameters. In order to obtain an accurate estimation of the recovery after tamping, it is necessary to use the measurement data in a short interval after the tamping interventions. Therefore, the cases, in which there are long periods of time without measurement after a tamping intervention, are removed from the constructed dataset. In a preliminary analysis performed in that study, it was observed that even for track sections with a similar track geometry condition before tamping, there is a significant variation in the recovery value after a tamping intervention (see Figure 5.2 a). In another attempt in Paper 2, the tamping interventions were classified into partial and complete tamping interventions. By identifying the tamping types, it was observed that a significant part of the variation in the recovery value was due to the tamping type used (see Figure 5.2 b).

![Figure 5.2](image)

*Figure 5.2 Recovery values after tamping versus the condition before tamping: (a) before identifying the tamping types and (b) after identifying the tamping types*

There are two main reasons for the difference in the tamping recovery values after partial and complete tamping interventions. Firstly, partial tamping actions are mainly performed using handheld vibratory tampers, tractors and lightweight machines to remedy isolated defects. These methods usually achieve a lower quality than tamping machines. Secondly, when a longer fraction of a track section is tamped, there is a greater expectation that
a better recovery will be achieved after a tamping intervention. Therefore, we distinguished between partial and complete tamping interventions to predict the recovery value. In order to find the best-fitted distributions for the recovery values after tamping interventions, the Anderson-Darling (AD) and Kolmogorov-Smirnov (KS) tests were applied. In the study documented in Paper 2, the underlying dependence of the recovery value on the track geometry condition before a tamping intervention was characterized by considering the distribution parameters as a function of the track geometry condition before tamping. Based on the results of the AD and KS tests for the standard deviation of the longitudinal level, the three-parameter Weibull distribution was found to be the best choice for representing the recovery values after tamping (see Paper 2 for more details). According to the results, the shape parameter was not dependent on the standard deviation of the longitudinal level before a tamping intervention. Figure 5.3 displays a comparison between the simulated and the real recovery values after tamping interventions versus the standard deviation of the longitudinal level before tamping.

![Figure 5.3 Recovery values after tamping interventions versus the standard deviation of the longitudinal level before tamping](image)

As is obvious in Figure 5.3, the recovery values simulated by the proposed model are very close to the real recovery values after tamping interventions. In addition, this figure shows that both the mean and the variance of the
recovery values after tamping interventions linearly increase with the standard deviation of the longitudinal level before tamping. The results of the AD and KS tests for the standard deviation of the alignment show that the three-parameter lognormal distribution is the best-fitted distribution for the recovery values after tamping interventions. Similarly to the results for the longitudinal level, the results show that the shape parameter is not dependent on the standard deviation of the alignment before tamping interventions. Figure 5.4 shows a comparison between the recovery values simulated using the proposed model and the real recovery values.

\[ \text{Figure 5.4 Recovery values after tamping interventions versus the standard deviation of the alignment before tamping} \]

As can be seen in Figure 5.4, the recovery values simulated by the proposed model are very close to the real recovery values. In addition, this figure shows that both the mean and the variance of the recovery values after tamping interventions increase with the standard deviation of the alignment before tamping. In order to measure the linear relation between the recovery of the longitudinal level and the alignment irregularities, the Pearson correlation coefficient was applied. The results show that there is a weak correlation between the recovery of the longitudinal level and that of the alignment. This means that tamping will not affect the different track geometry parameters equally. Hence, the study concludes that to obtain an accurate understanding of the tamping effectiveness, both the lateral and the vertical geometry parameters should be involved in the analysis.
To characterize the track geometry degradation and restoration with possible spatial dependencies, a two-level piecewise linear model is proposed in Paper 3. Since the standard deviation of the longitudinal level used as the degradation characteristic is strictly positive, the original data can be transformed into a log-scale. Since the log transformation is monotone, it does not affect the time when a degradation characteristic reaches the maintenance thresholds. The log-scale provides a great deal of flexibility and many benefits in statistical modelling and analysis (see Paper 3 for more details). Therefore, the geometry degradation was modelled under log-scale in the study presented in Paper 3. All the issues dealt with in the rest of this section were investigated under the log-scale, unless stated otherwise. The linear model proposed in Paper 3 was applied to represent the geometry degradation in a maintenance cycle. It should be noted that the proposed model is equivalent to an exponential model with a multiplicative lognormal error term in the original scale. Compared to linear or exponential models with an additive normal error term which can predict a negative degradation value, the proposed model guarantees the prediction of a positive degradation value at any given time.

The study documented in Paper 3 shows that in addition to gradual degradation, there can be an abrupt change in the degradation path in which the degradation level dramatically increases over time. We call this phenomenon a shock event. Figure 5.5 displays the effect of a shock event on the geometry degradation path.

![Figure 5.5 A degradation path with a shock event](image)

Although the occurrence of a change in the degradation path in a maintenance cycle is not a common case and may occur in a small number of sections, considering such changes when modelling track geometry degradation is very important. As can be seen in Figure 5.5, track sections
with a shock event normally have a significantly higher degradation rate which causes a shorter maintenance cycle than that of track sections with a normal degradation path. Shock events in the degradation path may cause safety problems and, in the worst-case scenario, lead to derailment. If the degradation pattern has changed after a certain point in time, then using one linear model to study the data obviously leaves the data unfitted or leaves the data poorly explained by the model. Therefore, for those sections with an identified change point in a maintenance cycle (a shock event), two separate lines are fitted using the piecewise linear function to model the geometry degradation.

In order to predict the recovery value and the change in the degradation rate after a tamping intervention, the multivariable linear regression model was applied in the study presented in Paper 3. Multivariable linear regression, applied as a predictive analysis method, was used to explain the relationship between the explanatory variables and the recovery value after tamping. In the study presented in Paper 3, the recovery is measured in terms of the change in the standard deviation of geometry parameters under log-scale. It must be noted that the recovery value \((SDLL_{before} - SDLL_{after})\) under the log-scale corresponds to the recovery ratio \((\frac{SDLL_{before}}{SDLL_{after}})\) under the original scale. Compared to the recovery value under the original scale, the recovery ratio under the original scale, i.e. the recovery value under the log-scale, is a more natural criterion to interpret the effects of tamping, as a linear regression model with response \((SDLL_{before} - SDLL_{after})\) may produce a large recovery value, leading to a negative degradation value under the original scale. Figure 5.6 presents the relations between the recovery values, the track geometry condition before tamping and the tamping type. The results of the applied multivariable linear regression model indicate that the track geometry condition before tamping and the tamping type are statistically significant predictor variables. The results also show that the expected recovery values after complete tamping interventions are higher compared to the values expected after partial tamping (see Figure 5.6). Moreover, in the study presented in Paper 5, the number of accumulated tamping interventions was added as an explanatory variable to the recovery model. It was found that the recovery value after tamping decreases as the number of tamping interventions increases (see Paper 5 for more details). This result is in alignment with the findings of other studies about the effect of tamping on track geometry degradation (see the literature review presented in Paper 1). In addition, in the study presented in Paper 3, the change in the degradation rate was predicted by considering the degradation rate before tamping as the explanatory variable. Figure 5.7 shows the changes in the degradation rate versus the degradation rate before tamping.
Figure 5.6 Scatter plot – degradation values before tamping versus the recovery values (the values are under a log-scale)

Figure 5.7 Scatter plot – degradation rates before tamping versus the change in the degradation rate
As shown in Figure 5.7, there is an inverse linear relationship between the degradation rate before tamping and the change in the degradation rate. In the case study, it was also found that when the degradation rate before tamping is high, there is a possibility that tamping will decrease the degradation rate. It must be noted that when the degradation rate is high, there is a quick change in the track irregularities in a part of the track section and that a tamping intervention may slow that change down. In addition, as noted in several studies, the number of accumulated tamping interventions affects the degradation rate. A higher degradation rate is expected after a number of tamping interventions.

The results of the case study presented in Paper 3 show that there is a high variation in the degradation parameters of the different track sections. Therefore, in another attempt in Paper 3, an exploratory data analysis was performed to analyse the uncertainty in the degradation parameters over track sections. The least squares algorithm was used to estimate the degradation parameters, i.e. the initial degradation level, the degradation rate and the standard deviation of the error term. It was observed that the values of the parameters of the degradation model change over track sections. In order to check the pattern of the distribution of the degradation parameters, the AD and KS tests were applied. It was found that the degradation parameters are generated by a normal distribution. In consideration of the fact that the analyses in Paper 3 were performed under a log-scale, it could be inferred that the degradation parameters are generated by a lognormal distribution under the original scale. This is in alignment with the results of the other studies in the literature. In the next step, the initial degradation value, the degradation rate and the standard deviation of the error term were considered as latent discrete processes. They were considered discrete as they had different values for each track section. To determine the spatial autocorrelation of the degradation parameters, the autocorrelation function (ACF) and partial autocorrelation function (PACF) were applied.

**Figure 5.8 Sample ACF and PACF for the initial degradation value**
As can be seen in Figures 5.8, 5.9 and 5.10, all three processes are highly auto-correlated. ARMA models were used to address the spatial dependence structures within the degradation parameters of different track sections. It is noted that the time order, i.e. the one-directional dependence, is different from the spatial process because of its possible two-directional dependence. In this regard, the model proposed in Paper 3 is restricted to stationary ARMA processes with Gaussian innovations because this class of ARMA processes satisfies the time reversibility. An ARMA process with time reversibility assigns a symmetric neighbourhood relationship to its realizations (see Paper 3, Section 4.4 for more details). In order to test the stationarity of the three time series (related to the initial degradation level, the degradation rate and the standard deviation of the error term), the Augmented Dickey-Fuller (ADF) test was applied. The results of the tests confirm that all the three time series are stationary (see Paper 3, Section 4.4). In order to choose the best-fitted
model among the different ARMA models, the models were compared with respect to the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) as the model selection criteria (see Paper 3, Section 4.4 for details about the optimal AIC and BIC values for the three time series and the estimated parameters of the ARMA models). By comparing the coefficients of the ARMA models, it is observed that the highest correlation is between the sections located right beside each other. In order to test the goodness of fit of the model, the Box-Ljung (BL) test was applied. The results of the test show that the ARMA models are properly fitted. In addition, by analysing the results, it is found that the correlation structures of the estimated processes for the degradation parameters are very close to those of the true processes.

5.2. Results and discussion related to RQ 2

RQ 2: How does one predict the occurrence of isolated geometry defects at the defect-based level and section-based level?

Generally, TQIs aggregate the track geometry measurements to represent the overall condition of track sections. Therefore, TQIs may not provide complete information about severe isolated defects in track sections. Isolated defects are short irregularities in the track geometry that can cause comfort problems for passengers, unplanned maintenance activities, damage to track components, and an increase in the risk of derailment. Hence, isolated defects must be considered in track geometry degradation modelling to facilitate effective maintenance decision making. To this end, the study presented in Paper 4 was undertaken to explore the prediction of isolated track geometry defects. A defect-based model was developed to predict the changes in the amplitude of the longitudinal level defects within a maintenance cycle. In addition, a section-based model was developed using binary logistic regression to predict the occurrence of UH2 defects in a track section.

In the case study presented in Paper 4, firstly the positions prone to the occurrence of UH2 defects were identified to develop the defect-based model. To achieve this, for each track section in line section 414, the six defects with the highest values and the six defects with the lowest values in each inspection run were considered for the preliminary analysis in the study. According to the results of the preliminary analysis, the rates of the change in the amplitude of most of the defects are very small (less than 0.5 mm/year). Moreover, in the preliminary analysis it was found that the defects which had

---

4 Underhåll 2 (in English: Maintenance 2) is the upper bound for corrective maintenance.
exceeded the planning limit might turn into UH2 defects in a short period of time. Therefore, the trend of the changes in the amplitude of these defects was used to predict the occurrence of UH2 defects. Figure 5.11 displays the evolution of the amplitude of a longitudinal level defect over time. By analysing the changes in the amplitude of the isolated longitudinal level defects it was observed that the degradation within a maintenance cycle had a linear pattern. Therefore, the linear regression model was applied to model the evolution of isolated longitudinal level defects:

\[ A(t) = A_0 + \beta(t - \text{tamp}) + \epsilon \]

where \( A(t) \) is the absolute value of the amplitude of the longitudinal level defect in time \( t \), \( A_0 \) is the absolute amplitude of the longitudinal level defect after the latest tamping intervention, \( \beta \) is the degradation rate, \( \text{tamp} \) is the latest tamping time, and \( \epsilon \) is the Gaussian random error term with a mean value equal to zero.

In order to check the normality assumption for the residuals of the simple linear model, the KS test was applied. The results of the KS tests showed that all the p-values of the KS tests were larger than the significance level (0.05). Consequently, it could be inferred that the normality assumption for the residuals of the simple linear model was suitable for the performed case study. To illustrate the distribution of the degradation rates, the histogram of the degradation rates is presented in Figure 5.12.
Figure 5.12 Histogram of the degradation rates

As can be seen in this figure, the tail of the histogram indicates that there are a number of defects with a high degradation rate (higher than 2 mm/year). These defects must be monitored and analysed for planning maintenance actions as they may turn into UH2 defects in a short period of time. Obviously, the presence of switches and crossings in track sections may affect the degradation rate. Hence, a box plot was used to identify the effect of the presence of switches and crossings in track sections on the degradation rates in the case study presented in Paper 4 (see Figure 5.13).

Figure 5.13 Box plot of the degradation rates for defects located in sections with and sections without a special asset
The t-test was applied on the degradation rates for defects located in sections with and sections without a special asset. The null hypothesis of the t-test was that there was no difference between the degradation rates for defects located in sections with and sections without a special asset. The test was conducted with respect to a 5% level of significance. The p-value of the t-test was close to zero, which shows that the null hypothesis was rejected. It can be concluded that defects located in a track section with a switch and crossing have a higher degradation rate on average, see Figure 5.13. Therefore, when an isolated defect exceeds the planning limit in these sections, special consideration should be given to them.

In the study documented in Paper 3, it was found that in addition to gradual changes of the SDLL, there exist situations in which the degradation level has dramatically increased over time, i.e. situations in which a so-called shock event has occurred. To investigate the occurrence of shock events in the degradation path, the changes in the amplitude of isolated defects were analysed in the study presented in Paper 4. It was found that when there is a shock event in the degradation path of the standard deviation of the longitudinal level, there is also a change point in the degradation path of the isolated defect (see Figure 5.14). In the case study presented in Paper 4, by analysing the trend of the changes in the amplitude of the longitudinal level defects, the time and position of the shock events were identified for the studied line section. Then the information about the maintenance history from BIS\(^5\) and BESSY\(^6\) was used to identify the possible root causes of the shock events. It was found that the occurrence of shock events is generally due to exogenous factors. Two identified reasons for the occurrence of shock events are improper sleeper replacement and drum installation. If after a sleeper replacement the ballast is not compacted well, this causes an abrupt change in the degradation path. In addition, the improper installation or replacement of drums may also cause a change in the track geometry condition. In the studied line section, some of the drums are installed 1.8 m under the track surface. A malfunctioning drum may cause some damage to the track geometry condition and, as a result, an abrupt change in the degradation path. In the event of a shock event occurring in the degradation path, there will be a new degradation pattern with a higher degradation rate. Therefore, to predict the track geometry degradation with change points, the degradation path will split into two linear parts.

---

\(^5\) Trafikverket’s asset register system
\(^6\) Trafikverket’s inspection report system
In the study presented in Paper 4, the effectiveness of tamping in rectifying isolated geometry defects was also studied. To this end, the changes in the amplitude of the longitudinal level defects after tamping interventions were analysed. Figure 5.15 illustrates in detail the effect of tamping on the amplitude of isolated defects by showing the longitudinal level waveform of a given track section before and after a tamping intervention. As can be seen, there is a sudden change in the defect’s amplitude after a tamping intervention, as is indicated by the black dashed vertical line.

It was observed that for around 35% of the sections with a UH1\(^7\) defect or UH2 defect, after a tamping intervention a UH1 defect or UH2 defect occurred again at the same position. This shows that the rectification of isolated defects through tamping was not durable for the studied line section. In many cases, spot tamping using tractors and lightweight machines or tamping machines

---

\(^7\) Underhåll 1 (in English: Maintenance 1) is the lower bound for corrective maintenance.
cannot remove the root causes of defects. Many problems in the track can be considered as root causes of isolated defects, e.g. broken sleepers, track substructure problems and drainage problems.

The proposed defect-based model can be used to predict the time of the occurrence of individual UH2 defects. However, generally railway infrastructure managers prefer to plan maintenance activities based on the overall condition of track sections rather than isolated defects. Therefore, in the study documented in Paper 4, a section-based model was developed using binary logistic regression to predict the probability of the occurrence of UH2 defects in a given period of time.

![Graph showing evolution of longitudinal level irregularities for a single track section](image)

**Figure 5.15 Evolution of longitudinal level irregularities for a single track section**

Generally, the standard deviation is used to show the variation or dispersion of track geometry measurement data. Since the presence of extreme values in the geometry measurements is of importance, kurtosis can provide useful information. Higher kurtosis is the result of infrequent extreme observations, as opposed to frequent moderate deviations. Therefore, it is
expected that a track section where most of the geometry measurements are low, but which contains a UH2 defect, will have a high kurtosis. Figure 5.16 presents the relationship between the standard deviation and the kurtosis of the longitudinal level with the presence of a UH2 defect in the track section.

Figure 5.16 Relationship between the standard deviation and the kurtosis of the longitudinal level with the presence of a UH2 defect in the track section

The left panel of Figure 5.16 shows the waveforms of the longitudinal level for three track sections in the same kilometre of line section 414. The right panel of the figure shows the estimated density of the longitudinal level measurements of the track sections obtained with a kernel smoothing function. In both track sections “a” and “b” there is a UH2 defect. However, in section “a” the longitudinal levels of most of the sample points are below the planning limit, whereas in section “b” there are a number of defects which have exceeded the planning limit. As a result, section “a” has a smaller standard deviation than section “b”, but has a higher kurtosis than section “b”. This is clear from the density function of the two waveforms in that the density function of section “a” has a sharper peak and longer tails than that of section
“b”. When studying the longitudinal level waveform of section “c” in Figure 5.16, one can observe that all the longitudinal level measurements are below the planning limit. As expected, the density function for this section has a sharper peak and shorter tails than that for the other sections, which indicates that section “c” has a smaller kurtosis than the other two. In addition, the density functions of sections “a” and “b” are flatter than the density function of section "c", which indicates that section "c" has a smaller standard deviation.

In the study presented in Paper 4, to find out how the standard deviation and kurtosis of the longitudinal level are related to the occurrence of UH2 defects, binary logistic regression was applied. The results show that both the standard deviation and the kurtosis of the longitudinal level are statistically significant and have positive coefficients. This means that the higher the standard deviation and the kurtosis of the longitudinal level are, the higher is the probability of the occurrence of UH2 defects. This finding is illustrated in Figure 5.17, which shows the relationship between the standard deviation and kurtosis of the longitudinal level measurements and the presence of UH2 defects in a track section.

![Figure 5.17 Relationship between the standard deviation and kurtosis of the longitudinal level and the occurrence of UH2 defects](image)

Figure 5.17 shows that when the standard deviation or the kurtosis of the longitudinal level is low, there is no UH2 defect in the track section. Moreover,
whenever the standard deviation of the longitudinal level is higher than the
UH1 limit and the kurtosis is low (close to zero), a very small number of UH2
defects have occurred. Similarly, when the kurtosis is high and the standard
deviation is low, a small number of UH2 defects have occurred. Therefore, both
the standard deviation and the kurtosis of the longitudinal level must be
higher than some specified value for a UH2 defect to be occurring in the track
section. Therefore, considering both the standard deviation and the kurtosis
as the explanatory variables in predicting the occurrence of UH2 defects may
increase the prediction accuracy.

The section-based model proposed in Paper 4 predicts the probability that
a section which contains defects which have exceeded the planning limit will
turn into a section containing at least one UH2 defect in a given period of time.
The explanatory variables considered in the model are the following four
variables: (1) the standard deviation of the longitudinal level (SDLL), (2) the
kurtosis of the longitudinal level ($\gamma_2$), (3) the presence of defects which
exceeded the planning limit or the UH1 limit in the latest measurement $c$, and
(4) the time interval $\Delta t$ (in years). According to the results presented in Paper
4, all the variables in the model have a significant effect on the probability of
the occurrence of UH2 defects. According to the results, the coefficient of time
is positive, which means that the probability of the occurrence of a UH2 defect
in a track section is higher in a longer period of time. In addition, the
coefficient of the categorical variable $c$ is positive. This means that when a
track section has been found to have a UH1 defect in the latest measurement,
the probability of the occurrence of a UH2 defect in the time period $\Delta t$ is higher
than when the section only contains a defect which has exceeded the planning
limit. Moreover, the standard deviation and kurtosis of the longitudinal level
have a positive coefficient. This means that the higher the standard deviation
and kurtosis are, the higher is the probability of the occurrence of a UH2 defect
in the time period $\Delta t$.

The probabilities predicted using binary logistic regression can be used to
classify the outputs by using a single cut-point ($\omega$) to compare each estimated
probability with respect to $\omega$. When classification is the main goal of the
analysis, the sensitivity and specificity can be used to assess the model
performance (Hosmer Jr, Lemeshow et al. 2013). The sensitivity and
specificity are dependent on the value of the cut-point. Figure 5.18 shows the
changes in the sensitivity and specificity with respect to different cut-point
values. In particular, the sensitivity is monotonic decreasing in $\omega$, while the
specificity is monotonic increasing in $\omega$ as shown in Figure 5.18. Therefore, it
is impossible to maximize both the sensitivity and the specificity
simultaneously. In order to select the optimal cut-point, one should optimize
a suitable measure which balances the sensitivity and specificity well. For this
purpose, the sensitivity and specificity for different cut-point values are
calculated, and we select the cut-point which balances both of them, i.e. makes
the sensitivity and specificity equal, to classify the data. As can be seen in Figure 5.18, the optimal value for the cut-point should be $\omega^* = 0.23$, as this is the value at which the sensitivity and specificity curves cross each other. By considering $\omega^*$ as the optimal cut-point, the model sensitivity and specificity are 89%, which seems reasonable for this case study.

![Figure 5.18 Plot of the sensitivity and specificity for different cut-point values](image)

5.3. Results and discussion related to RQ 3

As discussed in Paper 1, long-term predictions of the track geometry condition are an essential prerequisite for evaluating different maintenance and renewal strategies with respect to the RAMS parameters and the track life cycle cost. Such predictions make it possible to determine the optimal inspection intervals, maintenance thresholds, and renewal periods. In order to obtain long-term prediction of the track geometry condition, the track geometry degradation, the occurrence of isolated defects and the post-maintenance recovery must be modelled and integrated. According to the literature review presented in Paper 1, it is recommended that one should apply random coefficient degradation models to capture the section-to-section variation in degradation patterns. An additional recommendation is that one should consider the degradation level before tamping, the tamping type and the maintenance history as explanatory variables when modelling the recovery value after tamping interventions. Paper 1 shows that most of the
researchers covered in the literature review used the standard deviation of the longitudinal level as the track quality indicator when modelling the track geometry degradation and planning maintenance actions. However, it should be noted that ignoring extreme values of isolated defects will lead to an unrealistic picture of the track condition. Therefore, both the standard deviation of the track geometry parameters and extreme values of isolated defects of the parameters must be used to represent the track geometry condition. This facilitates a better understanding of the track system degradation and provides support in making more effective and efficient maintenance decisions. In addition, by considering the complexity of the geometry aging process due to the high variation in the degradation patterns of different track sections, as well as the interaction between the degradation and maintenance processes, applying simulation techniques to analyse the performance of the track is advantageous.

In Paper 5, an integrated statistical model was proposed to investigate the effect of different inspection intervals on the performance of the track to support decision making in geometry maintenance planning. The frequency of inspection is highly important as it affects the detectability of geometry defects and associated maintenance requirements. An applicable and effective inspection interval is a key factor for determining the performance of a line section, for controlling and reducing the risk of derailment and the maintenance cost, and for keeping the punctuality of train operations at the highest possible level. In the study presented in Paper 5, the percentage of time that track sections spend in different geometry states was used as the track performance indicator. By considering the maintenance limits, three different geometry states were defined, as presented in Table 5.1.

<table>
<thead>
<tr>
<th>State</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>There is no UH2 defect or critical defect in the track section.</td>
</tr>
<tr>
<td>2</td>
<td>There is at least one UH2 defect in the track section, but there is no critical defect in that section.</td>
</tr>
<tr>
<td>3</td>
<td>There is at least one critical defect in the track section.</td>
</tr>
</tbody>
</table>

It must be noted that, as explained in Paper 5, the running behaviour of the train is influenced by the amplitude of the track irregularities. Therefore, the percentage of time that track sections spend in different geometry states has a significant effect on the level of safety and the ride comfort. If the track geometry degrades to state 2 and state 3, this will negatively affect the energy consumption, the degradation rate of other track components, and the passenger satisfaction. From an availability point of view, the number of speed restrictions that a track section suffers in a given period of time is related to the track geometry state. In fact, when the track geometry degrades to state
2 and state 3, there is a higher probability of the occurrence of temporary speed restrictions. From a safety point of view, the risk of derailment in a given track section is dependent on the number of isolated defects, the amplitude of the defects, and the length of the time during which the defects remain undetected. Therefore, if the percentage of time spent in state 3 increases, the risk of train derailment increases. The aim of the analytical model proposed in Paper 5 is to predict the percentage of time that the track will spend in the different track geometry states using various inspection intervals. A fundamental requirement for achieving this aim is modelling and integrating the track geometry degradation, the tamping recovery, and the occurrence of severe isolated defects.

In the study presented in Paper 5, a random coefficient piecewise exponential model similar to the one proposed in Paper 3 was applied to characterize the track geometry degradation and restoration. In this study, the initial degradation value and the degradation rates were considered as random variables following a lognormal distribution. This is in accordance with the findings of Paper 3. It is worth noting that the second level of the proposed model in Paper 3 can also be used to capture the correlation structure of the degradation parameters. This may provide the opportunity to consider the concept of opportunistic maintenance in the maintenance strategy. However, since considering the concept of opportunistic maintenance was out of the scope of the study presented in Paper 5, a random coefficient piecewise exponential model was used. In order to relate the covariates to the recovery value after tamping, a multivariable linear regression model was applied. In order to predict the probability of the occurrence of isolated geometry defects, the model proposed in Paper 4 was extended. In this connection, ordinal logistic regression was applied to estimate the probability of the occurrence of UH2 defects and critical defects in a given track section. In the study documented in Paper 5, it was observed that the ratio of the inspection intervals with at least one UH2 defect or critical defect to the inspection intervals without geometry defects was very small (about 5:95). Therefore, the data set was imbalanced. Since the minority class (track sections with at least one UH2 defect or critical defect) was the class of greatest interest, misclassifying rare events could cause a big error cost from a learning point of view. Therefore, the adaptive synthetic (ADASYN) sampling method was applied to address the imbalanced data problem. Since the ADASYN method was originally created to handle two-class imbalanced data, class transformation had to be used to handle the multiclass problem.

Modelling and integrating the track geometry degradation, restoration, and the occurrence of isolated defects enable us to predict the track geometry condition in each inspection interval. In the study presented in Paper 5, it was assumed that the track geometry was inspected at discrete time intervals ($\tau$) to determine its condition, and that after each inspection, one of the following
three actions were taken: preventive maintenance, corrective maintenance, and corrective maintenance with an operational restriction on the track. Figure 5.19 depicts the three mentioned actions. A description of these maintenance actions now follows.

- **Preventive maintenance (PM):** If the standard deviation of the geometry parameters exceeds the planning limit, then the track section is assessed for preventive maintenance to be performed in the first available maintenance window. By scheduling the preventive maintenance for the time periods planned for maintenance interventions, the maintenance will be performed in a way which will neither interrupt the normal traffic nor cause traffic delays.

- **Normal corrective maintenance (CM$_n$):** When a UH2 defect occurs, a corrective maintenance action is conducted on the track section. Railway companies need short-term plans for adjusting the traffic by cancelling, postponing or rerouting trains to provide time to conduct corrective maintenance actions. Such short-term plans provide a period of time during which the track can be operated until maintenance takes place.

- **Emergency corrective maintenance (CM$_e$):** When a critical defect occurs, an immediate corrective maintenance action with a speed reduction or line closure is carried out on the track section.

![Figure 5.19 Schematic description of preventive and corrective maintenance actions](image)

Owing to the complexity of the degradation process and the maintenance process, predicting the effect of employing different inspection intervals on the track geometry condition is a difficult task. Therefore, the Monte Carlo simulation technique was used in the study presented in Paper 5, to handle the variation of the various parameters within the proposed integrated model and to estimate the percentage of time spent in the different track geometry states.
In order to verify the applicability of the proposed model, a case study was performed using the track geometry data collected from line section 414. In this case study, the standard deviation of the longitudinal level was used as the degradation characteristic. The measurement data were used to estimate the model parameters. As mentioned above, the initial degradation level and the degradation rate were considered as random variables following a lognormal distribution. Figure 5.20 provides two histogram plots showing the distribution of the degradation rates and the initial degradation level.

![Histograms of the initial degradation value and degradation rate](45392254)

In order to test the lognormality of the degradation parameters, the Anderson-Darling (AD) test was applied. The p-values for the AD tests performed on the initial degradation value and degradation rate were 0.35 and 0.14, respectively. Considering the significance level of 0.05, the results of the tests show that there is no reason to reject the lognormality assumption for the initial degradation value and degradation rate.

For the construction of the tamping recovery model in the study presented in Paper 5, the following explanatory variables were included: (1) the track geometry condition before tamping interventions, (2) the tamping type (partial or complete tamping, as explained under “RQ1” in this section), and (3) the accumulated number of tamping interventions. The results presented in Paper 5 show that the recovery value after tamping decreases as the number of tamping interventions increases. In addition, it can be observed that the expected recovery values after complete tamping interventions are higher than the expected recovery values after partial tamping. These findings are visualised in the two box plots presented in Figure 5.21.
In order to estimate the probability of the occurrence of UH2 defects and critical defects, ordinal logistic regression was used in the study presented in Paper 5. In the case study presented in that paper, the response variable $Y$ took a value of 0 whenever there was no geometry defect in the track section, a value of 1 whenever there was at least one UH2 defect in the track section, and a value of 2 whenever there was at least one critical defect in the track section. The explanatory variable considered in the model to predict the occurrence of isolated defects was the standard deviation of the longitudinal level.

![Figure 5.21](a) Box plot of the recovery values for different numbers of tamping interventions, (b) Box plot of the recovery values for different tamping types

The results for fitting the model show that the standard deviation of the longitudinal level has a significant effect on the probability of the occurrence of UH2 defects and critical defects in a given track section. The relationship between the standard deviation of the longitudinal level and the probability of the occurrence of UH2 defects and critical defects is shown in Figure 5.22. As can be seen in Figure 5.22, when the standard deviation of the longitudinal level increases, the probability that a track section does not have a UH2 defect or critical defect decreases. As is obvious from Figure 5.22, increasing the standard deviation of the longitudinal level increases the probability of the occurrence of at least one critical defect in a track section. Figure 5.22 also shows that after a certain value for the standard deviation of the longitudinal level, the probability of the occurrence of a critical defect is higher than the probability of the occurrence of a UH2 defect. It must be noted that, although the probability of the occurrence of UH2 defect may decrease after a specific value of the standard deviation of the longitudinal level, the summation of the probability of the occurrence of a UH2 defect and the probability of the occurrence of a critical defect is a non-decreasing function of the standard deviation of the longitudinal level.
Finally, as described in Paper 5, the degradation model, the recovery model, and the isolated defect model were integrated, and the Monte Carlo simulation technique was used to predict the percentage of time spent in different geometry states. In total, an assessment was made of the effect of the following five different inspection intervals on the track geometry condition: 1 month, 2 months, 3 months, 4 months, and 6 months. The simulations were run for a 10-year period. The maintenance planning time for the performance of normal corrective maintenance was assumed to follow a normal distribution with a mean of $m = 5$ weeks and a standard deviation of $\sigma = 1$ week. It was assumed that tamping windows were available every 18 months for the performance of PM, and that the alert limit was 1.5 mm. For each simulation, 80,000 runs were performed to make sure that the simulation would be converged. Table 5.2 shows the percentage of time spent in the three defined geometry states for the different inspection intervals.

The results of the Monte Carlo simulation show that the inspection frequency has a significant effect on the length of time that a track section spends in different geometry states. Based on the results obtained, changing the inspection interval from one month to six months increases the percentage of time that a track section spends in state 2 and state 3 by 3.22% and 0.25%, respectively. According to Table 5.2, the shorter the inspection interval is, the shorter is the time that a track section spends in state 2. In addition, the results of the simulations indicate that decreasing the frequency of inspections
will significantly increase the percentage of time that a track section spends in state 3. When the length of the inspection interval increases, the UH2 defects will remain undetected for a longer period of time. Therefore, for longer inspection intervals there is a higher probability that a UH2 defect will turn into a critical defect.

Table 5.2 Effect of the inspection interval on the percentage of time spent in the different geometry states

<table>
<thead>
<tr>
<th>Inspection interval</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Month</td>
<td>99.26%</td>
<td>0.69%</td>
<td>0.05%</td>
</tr>
<tr>
<td>2 Months</td>
<td>98.55%</td>
<td>1.35%</td>
<td>0.10%</td>
</tr>
<tr>
<td>3 Months</td>
<td>97.85%</td>
<td>2.01%</td>
<td>0.14%</td>
</tr>
<tr>
<td>4 Months</td>
<td>97.20%</td>
<td>2.62%</td>
<td>0.18%</td>
</tr>
<tr>
<td>6 Months</td>
<td>95.79%</td>
<td>3.91%</td>
<td>0.30%</td>
</tr>
</tbody>
</table>

Prediction of the percentage of time that a track section will spend in different geometry states is very important, as it is a primary input for predicting the probability of the occurrence of temporary speed reductions and the risk of derailment. It must be noted that the proposed model can also be used to predict the number of different maintenance actions, i.e. the number of preventive maintenance actions, normal corrective maintenance actions and emergency corrective maintenance actions. However, since in the present study the preventive maintenance actions are carried out on the track at predetermined times, the number of preventive maintenance actions and the number of normal corrective maintenance actions are not sensitive to different inspection intervals. In the case of a condition-based maintenance policy being implemented, the inspection frequency may affect the number of preventive maintenance actions and normal corrective maintenance actions. In this case, by determining the different direct and indirect costs for preventive tamping actions, normal corrective tamping actions, emergency corrective tamping actions, inspection, derailment, and the loss of capacity, the life cycle cost of each inspection interval can be assessed for the purpose of choosing the most effective one.
Chapter 6

Conclusions, contributions, and future research

This chapter concludes the research, summarize the contributions and suggest future research.

6.1. Conclusions

The following conclusions can be made on the basis of the research performed for this thesis to answer the three RQs given in Chapter 1.

RQ 1: How does one model the track geometry degradation over multiple tamping cycles by considering the spatial variation in the degradation parameters?

- Exponential model with a multiplicative lognormal error term can be properly used to model track geometry degradation in a tamping cycle.
- Tamping will increase the geometry degradation rate on average. In addition, there is an inverse linear relationship between the degradation rate before tamping and the changes in degradation rate after tamping.
- The recovery value after tamping decreases as the number of tamping interventions increases.
- Using probabilistic models and multivariable regression model to predict recovery value after tamping is more accurate than simple linear regression model.
- When tamping takes place, there exists a weak correlation between the recovery of the longitudinal level and that of alignment. In
another word, tamping will not affect the vertical and lateral geometry parameters equally.

- By considering degradation parameters as latent discrete processes, there is a spatial autocorrelation in degradation parameters. It is observed that the highest correlation is between the sections located right beside each other.
- ARMA model can efficiently be used to capture the correlation structure of the parameter processes.

RQ 2: How does one predict the occurrence of isolated geometry defects at the defect-based level and the section-based level?

- Degradation of the amplitude of longitudinal level defects has a linear pattern within a maintenance cycle. However, it must be noted that for long maintenance intervals the degradation may follows an exponential as ageing track degrades faster.
- It was observed that for around 35% of the sections with a UH1 defect or UH2 defect, after a tamping intervention a UH1 defect or UH2 defect occurred again at the same position. This shows that in many cases, spot tamping using tractors and lightweight machines or tamping machines cannot remove the root causes of defects.
- It is observed that in addition to gradual degradation, there can be an abrupt change in the degradation path, i.e. shock event, in which the degradation level dramatically increases over time. The two identified reasons for the occurrence of shock events are inadequate sleeper replacement or drum installation.
- The results show that the probability of the occurrence of UH2 defects in a track section is linked to the standard deviation of geometry parameter. Therefore, it can be concluded that standard deviation as an aggregated TQI is a statistically significant predictor for the occurrence of UH2 defects.
- The kurtosis, which is a measure of the tailedness of the data, can be used efficiently to capture the information about the occurrence of UH2 defects.

RQ 3: How does one develop a model for assessing the effect of different inspection intervals on the track geometry condition?

- The results of the simulation show that the inspection frequency has a significant effect on the track performance, measured by length of time that a track section spends in different geometry states.
- Increasing the length of the inspection interval increases the percentage of time that a track section spends in a state requiring normal CM action and a state requiring emergency CM. These in
turn will increase the probability of the occurrence of temporary speed reduction, line closure, and derailment.

- When the length of the inspection interval increases, UH2 defects will be detected with a longer delay. Therefore, with longer inspection intervals there is a higher probability that a UH2 defect will turn into a critical defect.

### 6.2. Contributions

The main contributions of this research can be summarized as follows:

- Synthesizing the recent published documents into a summary of what is known and has been validated, and what is not known, to provide a comprehensive picture of the state of the art concerning different issues in this field, as a major step towards further developments.
- Evaluation of the effect of tamping on vertical and lateral geometry parameters. This provides complementary information that is useful for obtaining an accurate understanding of the tamping effect on both lateral and vertical geometry parameters.
- Proposal of a probabilistic model to predict the recovery value after tamping intervention. The proposed model facilitates the prediction of track geometry condition over consecutive tamping cycles.
- Proposal of a two-level piecewise model to predict the evolution of track geometry condition over a spatial and temporal space. The proposed model consider the effect of tamping on track geometry degradation pattern.
- Proposal of the application of ARMA model to capture the spatial dependencies of the parameters of geometry degradation model.
- Development of a defect-based model using linear regression to identify the degradation pattern of isolated longitudinal level defects.
- Development of a section-based model using binary logistic regression to predict the probability of occurrence of isolated defects associated with track sections.
- Development of an integrated model; considering geometry degradation, tamping effect, and the occurrence of isolated defects; to investigate the effect of different inspection intervals on the track geometry condition. The results of the developed model support the decision making process regarding the selection of the most adequate inspection interval.
6.3. Suggested future works

During the progress of this research, several interesting new research ideas have emerged. However, it has not been possible to peruse all of these within the research presented in this thesis. Hence, in this section, some of these ideas are presented as suggestions for further research.

- Extension of the proposed degradation model by considering the occurrence of shock events in the degradation path. Since an abrupt change can significantly increase the degradation rate, it is risky to discard data after change. This problem is closely related to outlier detection and change point detection.
- Extension of the proposed defect-based model and section-based model to predict the occurrence of isolated twist defects. The twist is particularly important as it has a considerable effect on the wheel loading and vehicle stability, which are related to derailment.
- Extension of the proposed section-based model by including more explanatory variables, e.g. soil type, traffic information, and track type into the model to improve the accuracy for prediction of UH2 defects.
- Development of a model to predict the risk of the occurrence of temporary speed reduction and derailment as a function of track geometry condition.
References


Track geometry degradation and maintenance modelling: A review

Track geometry degradation and maintenance modelling: A review

Iman Soleimanmeigouni, Alireza Ahmadi and Uday Kumar

Abstract
Increased demand for railway transportation is creating a need for higher train speeds and axle loads. These, in turn, increase the likelihood of track degradation and failures. Modelling the degradation behaviour of track geometry and development of applicable and effective maintenance strategies has become a challenging concern for railway infrastructure managers. During the last three decades, a number of track geometry degradation and maintenance modelling approaches have been developed to predict and improve the railway track geometry condition. In this paper, existing track geometry measures are identified and discussed. Available models for track geometry degradation are reviewed and classified. Tamping recovery models are also reviewed and discussed to identify the issues and challenges of different available methodologies and models. Existing track geometry maintenance models are reviewed and critical observations on each contribution are provided. The most important track maintenance scheduling models are identified and discussed. Finally, the paper provides directions for further research.

Keywords
Degradation modelling, maintenance modelling, tamping recovery, track geometry, track geometry measures, maintenance scheduling

Date received: 14 September 2015; accepted: 29 May 2016

Introduction
Railways are a vital and effective means of mass transportation. They have played a key role in modern transportation and social development. This is due to the fact that railway transportation has high capacity, high efficiency, and low pollution compared to other transportation modes. Today, railways are experiencing higher demands, which will in turn impose greater demands on railway track performance. Two types of railway tracks are slab track and ballasted. Slab tracks are usually used in tunnels. The majority of railway tracks around the world are ballasted and these are the interest of this study. Conventional ballasted tracks have lower construction cost and properly respond to different static and dynamic forces.

Track condition can be represented in two different ways: track geometry condition or track structure condition. Track geometry can be defined as three-dimensional geometry of track. The defects and irregularities in track geometry are mostly used to represent the quality of the track and to plan track maintenance. The track geometry degrades with age and usage and can affect negatively track performance and safety. When track geometry degrades to an unacceptable level, it can lead to derailment, and the consequences can be significant. This may include a high cost of railway operation, economic loss, damage to the railway asset and environment, as well as possible loss of human lives. The largest track maintenance costs relate to the rectification of poor track geometry. Hence, track geometry should be inspected at predetermined times and maintenance actions should be planned to recover the track geometry to an acceptable condition. Maintenance actions are used to control the degradation of the track and restore the damaged track sections to an operational state. Activities that may be applied in track geometry maintenance are manual intervention, tamping, and stone blowing.

In order to select applicable and effective inspection and maintenance plans, track geometry degradation needs to be predicted using appropriate models and methodologies. However, modelling and predicting the track geometry degradation is a complex task. The complexity arises from the fact that different
track sections represent different types of degradation behaviour over time. The reason for this different behaviour among track sections is the existence of heterogeneous factors that affect track geometry degradation. The influencing heterogeneous factors on track degradation might change over track length, e.g. soil type, climate condition, track slope and curvature, and track component types and material. Thus, proper approaches should be employed to deal with these challenges. In addition, there are several interactions between different track components which need to be understood. Furthermore, the effect of maintenance on track quality needs to be modelled.

In summary, a higher demand for railway transportation creates an essential requirement for higher speed and axle load. These accelerate the track ageing process and negatively affect track geometry condition. Considering the complexity discussed above, the importance of developing effective track degradation and maintenance models arises, which will support infrastructure managers to arrive at the most effective maintenance decisions. In this regard, synthesising the published results into a summary of what is known and has been validated, and what is not, is a major step towards further developments in this field.

Development in this field requires selection of an appropriate track condition index to reflect the real condition of the track. Therefore, in this study, existing track geometry indicators have been identified and discussed. Accurate prediction of the degradation behaviour between consecutive maintenance activities is also essential to estimate the time to reach to the maintenance intervention thresholds. Hence, available models for track geometry degradation in the literature have been reviewed and a classification is proposed. The factors influencing track degradation are introduced and discussed. It should be noted that whilst track geometry recovery actions will improve the track geometry condition, they cannot rejuvenate the geometry condition to an as-good-as-new state. Hence, estimation of the effectiveness of the track geometry recovery is a vital input for any track geometry maintenance planning model.

Therefore, tamping recovery models are reviewed and discussed to identify the issues and challenges of different available methodologies and models.

Degradation and recovery modelling need to be combined to make a maintenance model. This enables the prediction of the long-term behaviour of the track and to identify the most optimal inspection intervals and maintenance intervention thresholds. Hence, associated maintenance models are reviewed and critical observations on each contribution are provided.

When the inspection intervals and maintenance intervention thresholds are defined, the major task is then to decide the most appropriate time for implementation. To this end, the most important track maintenance scheduling models are identified and discussed in this study. Following the outcome of the reviews and the above identification, the study suggests directions for future research in this field.

**Track condition measures**

Determining an indicator is an essential prerequisite to represent and evaluate the condition of railway track. The track condition can be represented in two different ways: track geometry condition or track structure condition. The defects and irregularities in track geometry are mostly used to represent the quality of the track and to plan track maintenance. Track geometry measures can be divided into five classes: (1) longitudinal level, (2) alignment, (3) gauge, (4) cant, and (5) twist. Longitudinal level is the track geometry of track centreline projected onto longitudinal vertical plane. Alignment is the track geometry of track centreline projected onto longitudinal horizontal plane. Gauge is the distance between the gauge faces of two adjacent rails at a given location below the running surface. Cant (cross-level) is the difference in height of the adjacent running tables computed from the angle between the running surface and a horizontal reference plane. Twist is the algebraic difference between two cross-levels taken at a defined distance apart, usually expressed as a gradient between the two points of measurement.\(^5\,6\) Track inspection cars run over track with a specific speed, monitor track geometry, and record the mentioned track geometry measures for every assigned movement (usually 25 cm). Three indicators can be used to assess track geometry quality based on the recorded measurement data by inspection trains, i.e. mean value, standard deviation over a specific length, and extreme values of isolated defects.\(^6\) In the standard deviation indicator, which mostly used for assessing track geometry quality, the geometry measures are recorded for every 25 cm over a track section (usually 200 m) and then the standard deviation of the obtained measurement data over the track section is calculated. Different track quality indices can be developed using the three indicators. Alignment and gauge are horizontal geometry measures, whilst longitudinal level, cant, and twist are vertical measures. The track geometry irregularities can be classified into short wavelength and long wavelength irregularities. Given their nature, normally long wavelength track irregularities have a negative effect on the comfort of passengers. However, short wavelength irregularities generate more vibration on axles and wheels.\(^7\)

Railway track is a distributed system and experiences different conditions at different points along the track. The track geometry degradation is a function of time (or tonnage) and space

\[
Q = f(t, s)
\]
where Q is the track geometry quality indicator, t is time, and s represents space.

Therefore, different parts of track show different degradation behaviours and as a result require different maintenance plans. In this regard, usually the track is divided into short length sections such as 100 or 200 m sections and track geometry measures are calculated for each track section.

An artificial track quality index (TQI), which is a combination of the aforementioned measures, can be used as the decisive indicator for maintenance planners. This quality index can either be the track geometry index (TGI) or the track structure index (TSI), but these two indices are not independent variables. The TGI is the track quality from the perspective of the user and is obtained from track geometry data recorded by track geometry recording cars. It indicates only the geometry conditions of the track, such as profile, twist, gauge, and alignment, which have a direct influence on track riding comfort. The TSI represents the structural conditions of the track, including the condition of the rail, sleeper, fastening systems, subgrade, and drainage systems. This index refers directly to the actual condition of the track and can indicate the track’s potential for degradation. Another quality index called MDZ represents the track quality condition from a different point of view. The MDZ value considers horizontal and vertical deviations in the track, together with speed and lack of super elevation. The variation of acceleration can indicate the track’s potential for degradation.

Several other indicators have also been proposed by researchers. In some of the research, a number of track geometry measures have been considered separately to represent the track condition. Hamid and Gross proposed five TQIs to represent track geometry condition. In order to obtain a TQI, wavelength filtering should be done on raw track geometry data. Then, a proper statistic such as standard deviation or mean should be selected to represent the track condition. These five TQIs are standard deviation of gage in inches, 99th percentile of gage in inches, standard deviation of 20 ft wrap in inches, standard deviation of 10 ft MCO of alignment in inches per 1000, and standard deviation of unbalanced super elevation in inches. In order to identify the correlation among different track geometry measures, Faiz and Singh studied the geometry parameters used in the UK track maintenance process and applied linear regression.

A number of researchers have tried to combine track geometry measures to propose an integrated quality index. The role of rail cant in overall track geometry condition is studied by Sadeghi et al. They proposed a TGI using alignment, profile, twist, gauge, and rail cant. They proposed new indices for each of the track geometry parameters based on allowable limits. Using justified coefficients, they combined the parameters to design the TGI. The proposed index can be used for ballasted or slab tracks in different locations. In the proposed index, the ratio of parameter deficiency to the allowable limit 

\[ T_i = \frac{\sum_{x}X_i}{x} \]

is used as the indicator of condition of geometry parameters for a track section, where n is the number of recordings in the track section and x represents track geometry parameters. In the next step, by considering the appropriate coefficients, the five indices are combined to represent the track geometry quality as follows

\[ ITGI = \frac{\sum_{i=1}^{5} T_i W_i}{\sum_{i=1}^{5} W_i} \]

where ITGI is the improved TGI and W is the model coefficients that should be calculated based on the importance of the geometry parameters for different track classes. Since the allowable limits of geometry parameters are considered in the formulation the overall condition of track and condition of the track...
with respect to a certain parameter can be evaluated simultaneously. Furthermore, the proposed index can also take the effect of different track classes into consideration.

Since different track classes have different quality conditions, it is important to consider the effect of track classes in the formulation of TQIs. Sadeghi and Askarinejad extended the previous work by proposing an overall TGI by considering different track classes. In this index, the normal distribution is selected to represent the general pattern of four track geometry parameters: alignment, profile, twist, and gauge. Sadeghi used the mean and standard deviation of the normal distribution of each track geometry parameter to develop new indices for each parameter. Since in the normal distribution about 97% of data are located in $\bar{x} \pm 3 \times SD$, this interval can be used to confirm that track geometry measures are in acceptable limits. Therefore, this interval is used to define indicators for each geometry parameter. He integrated the new indices with justified coefficients to make a TGI able to represent the overall condition of the track and identify the causes of track geometry defects. 

The other integrated index is the generalised energy index (GEI) proposed by Li and Xiao. They developed a grey model to predict the GEI sequence over time. The GEI can capture various effects of different wavelength components of track irregularity to the vehicle dynamic response. Because the GEI can consider different track irregularity wavelength and speed, they argued that it is a better index than the TQI to evaluate dynamic performance. They concluded higher speed will cause a significant increase in the effect of long wavelength irregularities on a vehicle’s dynamic response. Li et al. proposed an integral maintenance index (IMI) that considers the distribution of track geometry parameters to evaluate track condition, and they developed an integral maintenance plan based on this index. The IMI combines the track geometry parameters, maintenance history, as well as lateral and vertical acceleration.

Although artificial indices have their own advantages, using it without having a physical meaning has a number of disadvantages. The index does not indicate the relationship between the state of the track and the corresponding maintenance action required to improve track quality. Moreover, when using artificial indices, some of the track geometry measures, e.g. twist or alignment, can exceed their acceptable thresholds whilst the TGI remains within its predefined limit.

Integrating the track geometry and structural defects in a hybrid model to represent the track condition is an important task. Since the visual inspections for finding the structural defects need a huge amount of time and cost, determining the correlation between geometry and structural defects can pave the way for identifying the structural defects using geometric data. This issue is studied by Sadeghi and Askarinejad who modelled track condition based on structural defects and proposed a quantitative track structural quality index. This index is defined for each track component group, i.e. rail, sleeper, fastener, and ballast. The researchers developed a track maintenance framework based on the supposition that the correlation between track structural defects and track geometry irregularities is identified by comparing them. Later, Sadeghi and Askarinejad used the neural network technique to correlate track geometry irregularities to track structural defects. The main aim was to predict structural defects without performing expensive and time-consuming visual inspections. Track geometry data were considered as inputs of the neural network model, and the structural defects in the rail, sleeper, fasteners, and ballast were quantified and considered model outputs. The researchers concluded that their model is sensitive to track quality condition, and the model accuracy is not adequate for new tracks with high quality condition.

The proposed concept for identifying the relationship between the geometrical and structural defects may create this opportunity to find the causes of shock events in track geometry degradation path.

Different track evaluation methods can result in different degradation rates. Berawi et al. compared three track quality condition measures: J Synthetic, Indian TGI, and a measure based on European Standard EN 13848-5. The latter measure considers longitudinal level, alignment, and gauge. They claimed that the TGI had the highest degradation rate of the three methods. Regression was used to model the track degradation process for these three indices. However, despite the above-mentioned studies, the majority of the research on track geometry degradation and maintenance modelling used short wavelength longitudinal level ($3-25$ m) as the crucial factor. In fact, the defects associated with short wavelength longitudinal level can very well be recovered
by tamping. The two main reasons for this selection are (1) the defects in the vertical direction grow slightly more rapidly than defects in the horizontal direction and (2) the defects in the horizontal and cross-level directions would be automatically recovered by track maintenance activities. However, ignoring other measures for evaluation of track geometry condition may lead to an inaccurate indication of real track condition and this in turn may lead to an ineffective maintenance plan. For instance, twist is a decisive factor to be considered for derailment risk. When the twist condition is not considered for identification of track condition, the result might include a severe error and lead to an ineffective decision. In addition, when just longitudinal level is considered per se, it may not fully reflect the overall recovery effectiveness. In addition, the standard deviation of geometry parameters is mostly used for representing the track quality condition. However, considering other indicators such as mean and extreme values in addition to standard deviation for designing the TQIs will result to a better understanding of track geometry condition. Moreover, although the indicators can properly be used for degradation modelling, analysing the individual measurements recorded in every 25 cm can provide more details for shock failures and partial maintenance modelling.

**Track geometry degradation models**

After identifying the proper track quality measure to use, a degradation model needs to be constructed and then the effect of different maintenance strategies on track degradation may be evaluated.

According to Tzanakakis' degradation is the 'reduction of the original quality due to various influences'. Although the degradation process in the track system is quite slow, ignoring track degradation has destructive and costly consequences. By considering and analysing the track degradation, a company can make proper decisions to optimise inspection times, estimate the residual life of the track system and components, calculate the life cycle cost, and predict a suitable time for renewal activities.

Track geometry degradation can be divided into three phases, burn-in, useful life, and wear-out. The first phase starts immediately after ballast intervention. Because of existing gaps among ballast particles, degradation is high with a decreasing rate, until the gaps are removed. This phase has unpredictable behaviour and is difficult to model. Fortunately, however, it is short compared to the other two phases and can be disregarded in degradation modelling. The second phase begins after a predetermined amount of time has passed. In this phase, geometry degradation occurs more slowly than in the first phase. An approximately linear relationship can be seen between track geometry degradation and time or load in this period. This is the most important phase for degradation modelling and a great deal of research has been conducted on it. In the third phase, the degradation rate increases with time, so entering this phase is dangerous and should be avoided.

The approaches used for track geometry degradation modelling may be classified into mechanistic and statistical models.

**Mechanistic models**

Mechanistic or physical models are based on a priori physical information. The mechanistic approach 'consists of establishing, by theory and testing, the mechanical properties of all the elements that make up the track structure and the railroad vehicles'. The aim is to predict track degradation with few geometrical data. In another definition, the mechanistic approach is described as the modelling of mechanical reactions in the track, which result in track degradation. The most important advantage of mechanistic modelling is that by using this type of model the relationship between track responses and parameters of traffic can be properly clarified. The major weakness of mechanistic modelling is its inability to cope with the innate uncertainty of the track degradation behaviour due to heterogeneous influencing factors. In fact these models cannot deal with tracks with equal operational, environmental, and maintenance conditions showing different degradation behaviour. In addition, losing only one influencing factor can lead to a result that considerably diverges from reality. Another major problem is quantifying the track and vehicle properties. Understanding the interactions among track components and properties can also be difficult in some cases. A search of the literature suggests the best known and comprehensive mechanistic track degradation model is the integrated track degradation model (ITDM) proposed by Zhang et al. They developed an integrated mechanistic track degradation model that includes the interaction of track components. Thus, the proposed model can be used to predict the behaviour of the overall track or a particular component. The model includes three sub-models for rail, sleeper, and ballast/subgrade. In the model, the impact of component degradations on other components’ deterioration is represented by a change in dynamic forces on the rail. The authors concluded that track roughness significantly affects ballast depth and subgrade stiffness. They used a system level model that considers the degradation of components and the interaction among them. Hence, these types of degradation models are more comprehensive and representative of the overall track behaviour and condition.

Although evaluation of the physical or mechanistic track geometry degradation models is not the aim of this study, a summary of the most important mechanistic models in the literature and their results is provided under here. A literature review of research
on track degradation modelling and designing track maintenance planning decision support tools is performed by Ferreira and Murray. Their study covered a broad range of track-related issues, including models considering the effect of forces on track components, decision support systems for maintenance planning, and long-term simulation of the track maintenance process.

Shenton studied the six mechanisms comprising track degradation: dynamic forces, rail shape, sleeper spacing, sleeper support, ballast settlement, and substructure. He argued that ballast settlement plays the most important role in track degradation and proposed a mechanistic track degradation model considering axle load, number of axle loads, ballast type, sleeper type and size, and subgrade condition. He found axle load to be the most important of these factors in ballast settlement and that ballast settlement is related to the fifth root of the number of axle loads. His study demonstrated that track geometry condition during its life time is strongly dependent on the initial level of the track after implementation. Also, maintenance activities such as tamping cannot affect the internal properties of track geometry quality. Tamping tries to recover the track geometry condition to the maximum quality of the track, but it cannot change a low quality track section to a high quality one. As a result, the track sections with higher quality have longer maintenance cycles.

Sato studied the degradation of Japanese railway tracks and proposed a mechanistic degradation model for growth of track irregularities. He considered the phases of track settlement mentioned above. He argued that the coefficient of track degradation is dependent on the load factor, structure factor, and state factor and developed an empirical model for describing the average growth of track irregularities in a track section. In the proposed model, the growth of track irregularity is a function of speed, structure factor, and the third root of passed tonnage.

Sato claimed that the growth of track irregularity is dependent on the aimed value for track state, maintenance work, and the coefficient of track irregularity rectification. He showed that the growth rate of the irregularity of a track section has an exponential distribution.

A mechanistic approach is applied by Chrismer and Selig to model ballast, sub-ballast, and subgrade settlement behaviours. Track settlement is composed of three parts: ballast, sub-ballast, and subgrade settlement. However, they stated that the effect of ballast on track settlement is significantly higher than that of the other two factors, especially after performing a number of tamping interventions and this was also expressed by Shenton. These researchers used an exponential function to model ballast settlement based on wheel loads, abrasion number, fouling index, and number of load cycles and applied linear and power functions to model sub-ballast and subgrade settlement, respectively.

Oberg and Andersson modelled track degradation behaviour by considering four degradation mechanisms: track settlement, component fatigue, abrasive wear, and rolling contact fatigue. They argued that different vehicles can generate different track degradation behaviour. They used the ORE model to predict track settlement and component fatigue. In addition, they applied rail surface damage models to predict abrasive wear and rolling contact fatigue degradation.

For more studies regarding the formulation of mechanistic models and finding the advantages and disadvantages of the models, readers are referred to Guler et al. and Dahlberg. Guler et al. reviewed a number of mechanistic models for track geometry degradation, e.g. the Sato track damage model, British Rail Research track degradation models, and the Technical University of Munich settlement model.

A number of important track settlement prediction models, including Japanese, US, European, South African, and Australian models are compared in the work by Dahlberg. He focused on mathematical models that consider the inelastic behaviour of track ballast and subgrade. In addition, he developed a basic computer model based on the finite element method programme LS-DYNA to model dynamic train and track interaction and predicts the long-term behaviour of track settlement. The study looked at the two phases of track settlement and argued the reasons for the second phase of settlement are the densification of ballast and the inelastic behaviour of ballast and subgrade. Furthermore, the study found that those types of track settlement models such as ORE that do not consider the track structure factor may lose their ability to distinguish the degradation progress between two different structures such as soft and stiff soils. The study also found that the Japanese model does not distinguish between different loading cycles of different magnitudes, and it assumes long-term track settlement has a linear behaviour. Meanwhile, the US model considers that the behaviour is logarithmic. The study concluded that loading smaller than a specific threshold cannot cause track settlement, loading slightly more than the threshold will cause linear track settlement, and for large loading, the track settlement behaviour is non-linear and may be dependent to the fifth power of pressure. Finally, Dahlberg suggested that plastic deformation instead of elastic deflection should be used to model track settlement.

A number of researchers have tried to combine the mechanistic and statistical approaches to use the best of both methodologies. Rhayma et al. proposed an adaptable stochastic approach to various mechanistic
models to represent track degradation behaviour and deal with the inherent variability of the track’s mechanical and geometrical parameters. They used a reliability method based on the stochastic finite element method to assess the influence of different maintenance activities on the track. They used distributions instead of fixed values for the influencing factors to model the innate uncertainty of such factors.

Sadeghi and Askarinejad developed a degradation model by combining mechanistic and statistical approaches based on regression that considered track geometry and track structural condition data. Using a degradation coefficient, they estimated the effect of initial track geometry and track structural condition, train speed, and total million gross tons (MGT) of traffic passing on the track. They observed an exponential relationship between the degradation coefficient and the parameters. The initial track quality condition was found to be the most effective parameter influencing the degradation coefficient, with the total MGT passing along the track coming second. They concluded that the degradation coefficient is more affected by parameters in turnouts, bridges, and curve bridges than by parameters in other track segment types.

Statistical models

Uncertainty is a major characteristic of the track geometry behaviour, and a mechanistic approach is unable to capture this, due to the nature of this type of modelling. In order to arrive at a more effective and accurate decision, one needs to consider the uncertainty in the degradation modelling process. This necessitates employing concepts from probability theory, stochastic processes, and methods from the theory of a statistics for degradation modelling. Using these methodologies and approaches requires sufficient track geometry data.

Over the past decade, significant improvements in track geometry measurement technologies and systems have taken place. Jovanovic et al. presented a set of condition monitoring methods that are used for railway infrastructure maintenance and renewal management around the world. Track surface inspection systems that include track surface and right of way inspection, rail surface and component inspection, and track surface measurement are described in their work. They stated that laser no-contact measuring systems have many advantages compared to measuring systems based on inertial systems and contact systems for measuring track geometry parameters. They claimed that these innovative and modern inspection systems can dramatically improve our understanding of track condition. In fact, these systems enable access to a huge amount of applicable data that provide valuable information that reflects the real track geometry condition. In addition, simultaneously computer science and computational techniques have evolved dramatically, which facilitate faster computations and the analysis process. To this end, there is an emerging trend for researchers towards the use of statistical methods for modelling the track geometry degradation.

A statistical model is constructed from a set of input and output variables. This class of modelling requires sufficient data samples. Statistical models can be used for degradation modelling to cope with a great number of descriptive factors that have an effect on track degradation. This kind of model uses a large amount of data about the output variable (track performance) and descriptive or input variables (influencing factors) to construct a relationship amongst them. Statistical methods such as correlation analysis, regression analysis, and stochastic processes may be applied to design a statistical degradation model. The main advantage of statistical modelling is that since real data are used to construct the degradation model, an accurate estimation of track degradation is derived. An important disadvantage is the lack of a mechanical background for track components and their interactions with influencing factors and this could result in some unrealistic results. Yousefikia et al. suggested that when sufficient data are available, a statistical model is preferable to a mechanistic model.

Jovanović et al. stated that to develop an accurate degradation model some basic data are needed, which are superstructure and infrastructure inventory, layout and operating, work history, and condition measurements. The presented system includes databases, assets, functionalities (e.g. condition analysis, deterioration modelling), and visualisation and reporting parts. The focus of the study is on developing a generic degradation model that can be used to predict the future condition of track. They stated that by identifying the condition parameters, essential and temporary activities, suitable curves for degradation behaviour, and the restoration model, a quite accurate long-term prediction of track condition can be achieved. The proposed approach provides an overview about track geometry degradation modelling. In addition, although different curves can be fitted on measurement data in a maintenance cycle, it is accepted that the track degradation has (follows) approximately linear behaviour in a short interval (a maintenance cycle).

An important class of statistical methods for degradation modelling are stochastic processes including the Wiener process, the Gamma process, and the Inverse Gaussian process. However, very few researchers have used stochastic processes and they have tended to focus on the Gamma process. The process \( \{ Y(t); t \geq 0 \} \), where \( t \) is time, is a Gamma process if it has independent and Gamma-distributed increments. In fact, \( \Delta Y = Y(t + u) - Y(u) \),
where \( u \) is a time point, follows Gamma \((\mu, \eta(t + u) - \eta(u))\) with probability density function

\[
f_{\Delta Y(t)}(y; \mu, \eta(.)) = \frac{\mu^{(\mu y)} \exp(-\mu y)}{(\eta(t + u) - \eta(u))} \quad (2)
\]

where \( \eta(.) \) is a monotonic increasing function with \( \eta(0) = 0 \), \( \mu \) is the scale parameter, and \( \eta(t) \) is the shape function.

Meier-Hirmer et al.\(^{41}\) modelled the changes in standard deviation of longitudinal level using the Gamma process. They assumed that the mean and variance of the degradation rates are dependent on environmental variables, such as ballast type, curves, tonnage, and ascending and descending gradients. In this respect they applied classification and the regression trees method to predict geometry degradation of track sections and classified the track section based on similar degradation behaviour. It should be noted that the Gamma process has a monotonic evolution path. Therefore, when the degradation behaviour is in the form of cumulative damage, this kind of modelling can be useful.\(^{42}\) However, when the degradation path has both negative and positive increments, the gamma process may lead to an inaccurate prediction. In such case, that is common in track geometry degradation paths, the Wiener process might be used. A stationary Wiener process is particularly well suited to modelling the evolution of a degradation mechanism characterised by a linear increase over time with random noise.

The application of the Gaussian random process to model track irregularities in vertical profile and alignment is studied by Zhu et al.\(^{43}\). They discussed the use of power spectral density analysis and cross-level statistics about track irregularities to improve track degradation modelling.

One of the important issues in track geometry degradation modelling is considering different track condition measures simultaneously. In this regard, Mercier et al.\(^{44}\) considered the longitudinal level and transversal level as track quality parameters. Since these two track quality parameters are dependent, they applied the bivariate gamma process to model track degradation and used trivariate reduction to build the bivariate gamma process.

A stochastic approach based on the Dagum distribution is developed by Vale and Lurdes\(^{45}\) who modelled track longitudinal level degradation over time. The researchers classified the track longitudinal level changes into three speed classes and different inspection intervals. Based on their analysis of the effect of right and left rails on track longitudinal level degradation, they argued that the degradation rate for both of rails is similar.

Obviously, employment of the stochastic processes may lead to a more robust maintenance plan. However, it should be considered that application of stochastic processes for modelling track geometry degradation has some drawbacks. When the variance term in the model is significantly bigger than mean term of the model the stochastic process may lead to a very inaccurate prediction. Moreover, in industrial application, the main flow of degradation models based on stochastic processes is the complexity in computation.

Path analysis is another powerful tool that has been widely applied in track geometry degradation modelling and identifying the effect of different factors on track geometry degradation. Factors with a significant impact on track degradation can be categorised into three main classes: (1) factors related to track condition, (2) factors related to vehicle, (3) factors related to environment. The main factors related to track geometry condition are construction materials, gradient, curvature, cant, age, rail type, sleeper type, rail length, weld type, ballast quality, and ballast thickness. The most influential factors related to the vehicle are speed and tonnage, but vehicle type and axle load are also important. Common environmental factors in degradation studies are weather, water, landslide, flooding, falling rock, and snow. Another classification is proposed by Famurewa et al.\(^{46}\) who categorised the factors influencing track geometry degradation into five classes: maintenance and age, traffic characteristics, material, design parameters, and environmental condition. Maintenance history is one of the most important factors in track condition and has a significant influence on track degradation. Maintenance activities related to ballast, such as manual and mechanical tamping, ballast profiling and stabilisation, and mechanical ballast cleaning, have an impact on vertical geometry measures. Small component maintenance, small component renewal, sleeper renewal, and rail renewal affect the horizontal geometry.\(^{36}\) Jovanovic\' et al.\(^{39}\) mentioned that rail corrugation has a significant effect on track geometry failure by increasing the dynamic forces and so could be the root cause of some of these types of failures. Hence, this failure mode should be monitored and planned for maintenance.

Multivariate statistical analysis is an advanced tool for determining the relationships among two or more variables of interest. This technique is applied in track geometry degradation modelling to identify the impact of heterogeneous factors along the track on geometry degradation. Lyngby\(^{9}\) suggested a methodology for evaluating track degradation in terms of track geometry irregularities and proposed a multivariate regression model to demonstrate the relationship between the track degradation measure variable and other influencing variables. He applied two exponential functions to deal with the initial and second phases of degradation. Since different sections of track are not identical, the track was split into

homogeneous sections with similar influencing variables. He concluded that: (1) axle load has a non-linear relationship with degradation, (2) degradation after tamping is dependent on the number of previous tampings, (3) soil consisting of clay material will settle sooner than other types of soil, (4) light rail tracks degrade faster than heavy rail tracks; and (5) harsh rainfall increases degradation rate. However, it should be considered that small values for the estimated coefficients may not reflect the magnitude of influencing factors. More analysis is required to find the effect of influencing factors on track degradation rate. In addition, in order to dividing track into similar track sections for identifying the influencing factors on degradation, it should be noted that the track sections with similar influencing variables may show different degradation behaviour.

The usual method for track segmentation is to divide the track into fixed short length track sections. In fact, since different track sections show different degradation behaviours we need to model the degradation of each track section separately. Another approach is segmentation of the track based on similar traffic, track structure, environmental condition, and maintenance history characteristics. Guler et al. developed a degradation model by using multivariate statistical analysis for different track geometry parameters. First, they divided the track into homogeneous sections based on gradient, curvature, cant, speed, age, rail type, and rail length. They used a multiple linear regression model to model the degradation rate of track geometry parameters in terms of the independent variables. They examined the effect of traffic load, speed, curvature, gradient, cant, sleeper type, rail type, rail length, falling rock, landslide, snow, and flood. They concluded landslide and snow do not affect track geometry degradation, but rail type and rail length do have an effect. The model found a high correlation between cant and curvature. It also showed that cumulative load affects degradation, but negatively. In their model, when the cumulative load is increased, the track geometry degradation is reduced, but this is contradictory to other studies.

Westgeest et al. applied a regression method to model track degradation and maintenance. They used a combination of track geometry parameters to create the key performance indicator (KPI) as the track quality indicator. The KPI is considered the response variable in the proposed regression models. They studied the effect on the KPI of different types of subsoil, sleeper, tonnage, and engineering structures and considered two types of tamping, manual and mechanical. The results showed that the proposed degradation model can properly address changes in the KPI over time, but it is not efficient in terms of track behaviour prediction. They concluded that the track segments have different degradation rates depending on a number of factors, e.g. closeness to switches, sleeper types, and subsoil types. The results of the study support the fact that section-to-section variation should be considered for track geometry degradation modelling.

Hamid and Gross considered the effect of 11 physical factors in the process of degradation modelling for five selected TQIs. These factors are classified into three groups, i.e. track structure, traffic, and maintenance (basic and discretionary). A stepwise autoregressive (AR) model is used to model track degradations for unmaintained tracks. The track is segmented into many unequal length homogeneous sections. Statistical analysis is carried out to determine the contribution of the mentioned factors on each of the TQIs. He concluded that more than 80% of the variation in the TQIs can be explained by the proposed models.

He et al. applied a statistical exponential model for each track geometry defect to model the relationship between the degradation rate and influencing factors such as monthly traffic (MGT), total number of cars per month, total number of trains per month, and number of inspections since the last observed critical geometry defect. They found that various track geometry defects have different degradation rates. Most are sensitive to traffic volume and only a few do not reflect an increase in traffic volume.

Bing selected five TQIs using correlation analysis and applied a statistical degradation model based on step-wise linear regression using these TQIs. He found a strong dependency between the track quality of a segment at the present time and its quality at a previous time step. However, the model was not accurate in terms of track degradation prediction. In the next phase, a mechanistic approach was used for surface track degradation modelling. The present surface TQI, annual tonnage, axle load mix, train speed, ballast type, and track modulus were considered as high significant factors. The regression model was used to represent the relationship between TQI and initial quality and influencing factors.

The path analysis methods can increase our knowledge about the parameter selection for modelling track geometry degradation. However, due to high level of uncertainty in geometry degradation behaviour, performing path analysis on other track lines in different places may give a different result. An interesting area for further research is using the results of the path analysis in a model with random coefficients to increase the accuracy of track geometry degradation models.

Different types of data-driven approaches are also widely applied by researchers in track degradation modelling. These approaches may be categorised into machine learning methods, fuzzy methods, filtering, and data-driven statistical methods. The machine learning models include support vector machines, artificial neural networks (ANNs), and Bayesian networks.
An ANN is a computational model that can be used to predict the approximate behaviour of track geometry with time by considering a large number of inputs. Guler\(^7\) used this method to model the degradation of different track geometry parameters. The track was divided into homogenous segments and then he estimated the degradation rates between each two consecutive maintenance interventions using linear regression. The average degradation rates were the output variables of the ANN model. The model considered traffic load, velocity, curvature, gradient, cross-level, sleeper type, rail type, rail length, falling rock, landslide, snow, and flood as influencing factors. By considering R-square values, he concluded an ANN model can provide a good estimate of the degradation rate.

A number of studies have tried to make short-term predictions for track geometry degradation. These studies may have application for prognosis analysis. Xu et al.\(^49\) proposed an approach based on historical changes in track irregularity to predict the short-term track degradation. They estimated the non-linear behaviour of track irregularity during a maintenance cycle using a number of short-range linear regression models. When regular inspections are performed, a family of regression models can be constructed. A non-linear model can be estimated using this family, based on integral theory. The proposed model is able to predict track irregularities two months in advance. In another study, Xu et al.\(^50\) proposed a track measures data mining model to predict railway track degradation for a short time period. They considered twist and alignment when they performed a validity test and observed the prediction errors for both had a normal distribution with mean close to zero and a standard deviation smaller than one. They concluded that the model can accurately address track condition, two or three months in advance. Using waveform data, Liu et al.\(^51\) proposed a short range prediction model to estimate any track irregularity index over a short track section length (25 m) and on a day-by-day basis. The proposed model uses linear regression to predict track geometry condition for every day. By new inspection, the states of the model are updated and this cycle is optimised on a rolling basis. They concluded that the total process of track surface change over track sections is non-linear and different track sections have different non-linear process. However, accepting this conclusion needs more studies and discussions. Kawaguchi et al.\(^52\) developed two degradation models to predict track alignment irregularities. First, they developed a degradation model based on analysis of lateral track deformation to estimate mean time to maintenance of track alignment irregularities and evaluate various maintenance plans. Second, they designed another degradation model based on the exponential smoothing method to accurately predict the track alignment irregularities a maximum of one year in advance. Except for a few cases, the proposed model was able to accurately predict track alignment irregularities. They found that the growth of alignment irregularities has a larger value in the neighbourhood of normal joints. The first proposed degradation model can be used to compare different maintenance strategies from an economic point of view and the second can be used to construct an annual track scheduling plan.

Bai et al.\(^53\) proposed an approach called the tree-augmented naïve Bayes-TQI to predict railway track irregularities for a short-term horizon. They assumed that the difference in the TQI between two inspection times represents the effect of heterogeneous factors on track irregularities. The historical data related to four previous inspection intervals were used to predict the next inspection irregularities. The proposed model proved more accurate than models in similar previous work. A machine learning model based using a multi-stage framework is developed by Xu et al.\(^54\) who predicted changes in track irregularity over time. They defined different stages of track irregularity changes based on maintenance thresholds and linear regression was used to predict track degradation in each stage. Jovanovic\(^27\) developed a generic degradation model that is suitable for modelling degradation of different parameters. To develop a generic degradation model, he argued that the condition parameters that represent the condition of track components and essential and temporary activities affecting these components should be determined. He observed different degradation patterns based on various intervals between essential or temporary activities. Various curve types, such as linear, exponential, and quadratic, may be used to explain the degradation patterns.

Time series models can be used to model track geometry degradation model. An AR model can predict the next track geometry condition based on recent track geometry condition data

\[
Q(t + 1) = a_1Q(t) + a_2Q(t - 1) + \cdots + w(t)
\]  

(3)
where $Q$ represents track geometry condition and $w$ is a random normal variable

$$w_{(t)} \sim N(0, \sigma^2) \quad (4)$$

The number of terms that should be used in AR model for accurate prediction of track geometry degradation can be computed based on correlation analysis.

Autoregressive moving average model (ARMA) is another type of time series models. The formulation of ARMA model is as follows

$$Q(t + 1) = a_1 Q(t) + a_2 Q(t - 1) + \ldots + w_{(t)} + \beta w_{(t-1)} \quad (5)$$

Chaolong et al.\textsuperscript{56} applied time series analysis to predict track irregularity using standard deviation time series data. They identified different patterns and specifications of track irregularity using the clustering approach. In order to predict the changing trends of track irregularity, they used the linear recursive model and the linear ARMA. However, it should be noted that the proposed model is applied for a short length track section.

The grey model can be used for prediction with very few measurement data. However, the grey model does not follow a global trend; rather, it tries to follow the original degradation pattern. Famurewa et al.\textsuperscript{46} evaluated track geometry condition using control charts and compared the accuracy of linear, exponential, and grey models in the estimation and prediction of track geometry degradation. Time is the descriptive variable in the proposed model. The comparison demonstrated that the grey model has a lower mean percentage error than the linear model and an approximately equal error value compared to the exponential model. Since grey model adapts its parameters to new conditions as new data become available, its application may be advantageous over simple linear and exponential regression. However, it should be noted that, although the grey model can be used for prediction with only four measurements with a good fit, the obtained estimation may not be reliable. In fact, very few number of measurements cannot determine the degradation path. Liu et al.\textsuperscript{57} developed a model by integrating a grey model with a Markov chain to predict track quality condition. When they conducted a case study, they observed that the proposed model was better for track quality prediction than traditional grey models. Since the grey model mainly deal with the data that has a quite smooth trend and has no merit in capturing uncertainty, using the Markov model for prediction of residual values will increase the model accuracy. In a comparative study, Quiroga and Schnieder\textsuperscript{58} examined the efficiency of the double exponential smoothing method, a generic degradation model, and an AR model for track degradation prediction. They considered a non-linear model for modelling track restoration after tamping. By comparing the degradation models, they observed the degradation rate dramatically changed after tamping intervention. Moreover, the three models lose their efficiency in track degradation prediction after performing a number of tamping procedures. After considering these issues, they developed a hybrid discrete-continuous framework based on a grey box model. Tamping intervention is considered as a jump action in the model. After comparing these four models, they concluded that the proposed hybrid model is more efficient in terms of track degradation behaviour prediction. Chaolong et al.\textsuperscript{59} developed a modified grey model to analyse track irregularity time series data for a fixed measuring point and obtain a medium–long-term prediction of track cross levelling. By validating the model, they found that the proposed model can accurately predict the long-term track irregularity changes. In addition, they compared the stochastic linear AR model, the Kalman filtering model, and an ANN model with respect to short-term track cross levelling prediction of a track section. They observed that the ANN model was more accurate than the two other models.

In order to cope with the probable non-linearity behaviour of track geometry degradation between two consecutive maintenance actions, some researchers have used multistage linear regression models. The multistage degradation models are applied to deal with the periodic differences in the degradation process. Guo et al.\textsuperscript{60} modelled the relation between track alignment irregularity and gross passing tonnage for a short track section and between two consecutive maintenances. Three degradation phases are considered in their study and a multiphase linear model is constructed based on maintenance thresholds. Chang et al.\textsuperscript{61} proposed a multistage linear model to predict changes in track irregularity. By assuming three properties for track geometry degradation, i.e. cyclic, multiage, and exponential changes, they were able to model the track degradation between two consecutive maintenance interventions. Based on multistage and exponential changes in track irregularities, they modelled different stages of TQI changes using a number of linear models. The different stages of track irregularity changes were based on the TQI distribution. Finally, after conducting a case study, they concluded that the proposed model gives more accurate results than linear models. However, it should be noted that accepting the cyclic feature of track geometry degradation needs more discussion. In fact, the track may show different behaviour in different maintenance cycles. In another study, a multistage linear model is used by Guo and Han\textsuperscript{62} to cope with different phases of degradation between two consecutive maintenance interventions and the exponential growth of track irregularity.
Although applying the multistage linear models increases the accuracy of degradation modelling, a number of issues restrict their application for track geometry degradation modelling. The main issues are determining the number of degradation stages, requiring a large amount of data, and complex computation for model parameters estimation. The basic reason for using multistage linear models is to deal with different rates in the geometry degradation path. Although based on the theoretical background supported by bathtub curve different phases of degradation should be considered in the degradation modelling, in practice the burn-in and wear-out phases are too short or will not happen. In fact, geometry degradation mostly shows a more or less linear behaviour in a maintenance cycle. Another assumption that mainly is considered in these papers is cyclic behaviour that means track geometry degradation in different maintenance cycles has a same behaviour. However, it is accepted as a true that the maintenance intervention will affect the track geometry degradation model parameters.

An important field of study is the estimation of the probability that a track section needs an unplanned maintenance. Andrade and Teixeira applied logistic regression to estimate this probability. They considered that the response probability is dependent on two groups of variables. The first group includes the standard deviation of longitudinal levelling defects and the standard deviation of horizontal alignment defects which are common indicators for planned maintenance. The second group consists of two binary variables that demonstrate whether a track section has at least a switch and whether a track section is located on a bridge. They found that the two mentioned indicators for planned maintenance are statistically significant predictors of unplanned maintenance needs. However, they observed that the standard deviation of horizontal alignment has a more significant effect on the probability of unplanned maintenance needs. In addition, the presence of bridges and switches in track sections may have different effects on unplanned maintenance needs.

The section-to-section variability should be considered in the geometry degradation modelling to describe the differences in degradation processes of sections from same track line. Incorporating random coefficients into degradation models enables description of section-to-section variability. Andrade and Teixeira assumed that the track is comprised of four groups: switches, bridges, stations, and plain track and considered a linear model for track longitudinal level degradation. They performed statistical correlation analysis for each group section and fitted the log-normal distribution to the track’s longitudinal level degradation parameters. The researchers detected a significant correlation between the initial longitudinal level standard deviation and the degradation rate for opposite directions with a profound impact on unavailability parameters. They argued that the correlation between adjacent sections shows the need for clustering the tamping interventions. Furthermore, track sections that include switches and bridges have a meaningfully shorter tamping intervention cycle. This approach is beneficial in the way that it considered the section-to-section variation in degradation rates. In fact, it is considered that the parameters of linear regression model are random variables that follow a certain distribution.

In order to evaluate a track geometry degradation model and deal with the uncertainty of its parameters, Andrade and Teixeira used a Bayesian approach. They considered the track longitudinal level deviation to have a linear relationship with passing tonnage and assumed the initial longitudinal level and degradation rate would have a bivariate log-normal prior distribution. The researchers divided the track into four groups based on infrastructure features. They calculated the posterior distribution of the parameters for different stages, i.e. design, after first inspection, between first inspection and first tamping, and between the second and remaining tamping interventions. They argued that the parameter uncertainties are significant in the design stage. Therefore, degradation prediction and maintenance planning must be done at a logical time after the design stage. The advantage of the proposed model is considering the uncertainty of the parameters of model in simulation of the degradation process.

A hierarchical Bayesian approach combined with a conditional AR term is also put forwarded by Andrade and Teixeira who modelled track geometry degradation. They assumed that the standard deviation of longitudinal level has a normal prior distribution. They also assumed that the mean of this normal distribution for different track sections is dependent on the initial quality before or after track renewal, the degradation rate before or after track renewal, and the disturbance effect of tamping. They considered a normal distribution for the disturbance effect of tamping and used an inverse gamma distribution for variances in the hierarchical Bayesian model. Finally, they performed a Markov Chain-Monte Carlo simulation to deal with difficult calculations of the proposed model. The proposed model deals with the spatial dependency of the track geometry degradation parameters of neighbour track sections. By assuming that there is a spatial dependency between consecutive track sections, the proposed model can reduce the uncertainty of track geometry degradation parameters over track length. In addition, the other advantage of the model is considering the tamping and renewal effect on degradation parameters.

The models with random coefficients can capture the uncertainty of degradation parameters over track
length. By considering more influencing factors on geometry degradation parameters, better fitted distributions with lower variance will be obtained for modelling the geometry degradation parameters. In order to avoid complex computations of degradation models with random coefficients, the random parameters and associated distributions must be selected as simple as possible. The output of the path analysis that is described before can be used to find the candidate factors for designing the models with random coefficients. In fact, by identifying the influencing factors on the track geometry degradation, the degradation model parameters can be considered as the random variables that are a function of the mentioned factors.

Another class of methods for modelling track geometry degradation between two consecutive maintenance interventions are Markov models. However, the Markov model has limitations that restrict its application for track degradation modelling. First, in a Markov chain, the transition probabilities amongst different states should be constant over time, but before this assumption is accepted in the track degradation field, more studies are required. Second, in a Markov process, the current state is only dependent on the immediate previous state. Finally, the size of the Markov model grows dramatically with an increase in the number of components.4,19

Application of a stochastic Markov model to evaluate track degradation between two consecutive maintenance interventions is addressed in the work by Bai et al.67 The states of the Markov model were defined by classifying the TQI values into four classes. The researchers considered the inspection interval times between two consecutive maintenance actions as Markov model time steps. They constructed a Markov transition probability matrix by considering various heterogeneous factors. They argued that the existence of these heterogeneous factors caused two maintenance units with the same mileage to show different degradation behaviour.

Yousefikia et al.24 modelled tram track degradation using a Markov model to obtain the optimal maintenance strategy. In their view, since tram tracks include numerous tight curves, the effect of curves on track degradation in terms of real wear should be considered through degradation modelling. A two-stage model to optimise track maintenance planning was developed by Shafahi and Hakhamaneshi.68 The track was divided into six classes based on traffic and topography, and its condition was evaluated based on a combined track record index. First, the researchers used Markov model to model track degradation behaviour, making use of data obtained from an Iranian railway to construct the transition probability matrix. Second, they applied dynamic programming optimisation to obtain optimal decisions for maintenance activities.

Before applying the Markov model for track geometry degradation modelling between two consecutive maintenance actions a number of issues should be considered. Since the track geometry measures are continuous and the Markov chain is a discrete model, the track geometry measures should be discretised that can lead to an inaccurate estimation of track degradation. Discretising the track geometry measures to a few numbers of states will lead to simple estimation of transition matrix; however, it will result in an inaccurate estimation of track degradation. On the other hand, defining large number of states for discretising will increase the size of transition matrix and associated calculations. Moreover, in Markov chain it is considered that the sojourn times in different states follow exponential distribution. This assumption limits the application of Markov chain for track geometry degradation modelling. In addition, due to section-to-section variation of geometry degradation different transition matrices should be defined for different track units which contain many track sections with similar characteristics. However, due to high uncertainty of geometry degradation over track length, finding the track maintenance units with similar transition matrix is a difficult task. A summary of the reviewed papers for track geometry degradation modelling is provided in Table 1.

**Track geometry maintenance models**

By combining recovery and degradation models the long-term behaviour of the track geometry may be predicted and different maintenance strategies may be evaluated with respect to different objectives. As a result, it is possible to determine optimal inspection intervals and maintenance intervention thresholds. Thresholds are defined as alert points to prevent the track from entering hazardous zones by performing maintenance intervention.

Acceptable thresholds for different types of fault and various speed levels can be found in European Standard EN 13848-5. The main threshold levels are the immediate action limit (IAL), the intervention limit (IL), and the alert limit (AL). When degradation reaches the IAL, speed constraints should be enforced or immediate modifications should be made to the track. Once degradation passes the IL, the track requires corrective maintenance. If the AL is passed, the condition of the track geometry should be regularly analysed and planned maintenance should be performed.69 When a track section exceeds a predetermined degradation level, more inspections are required. A degraded track section is more vulnerable to defects; thus, if a certain degradation threshold is passed, immediate maintenance should be performed and modifications made.69 Applying dynamic threshold levels that are dependent on the age of the track structure and the number of previous interventions...
Table 1. Summary of the reviewed papers for track geometry degradation modelling.

<table>
<thead>
<tr>
<th>Model</th>
<th>Author(s)</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curve fitting (linear or non-linear)</td>
<td>Jovanovic\textsuperscript{27} and Quiroga and Schnieder\textsuperscript{58}</td>
<td>To simply model track geometry degradation</td>
</tr>
<tr>
<td>Grey model</td>
<td>Li and Xiao,\textsuperscript{17} Famurewa et al.\textsuperscript{46} Liu et al.,\textsuperscript{57} Quiroga and Schnieder,\textsuperscript{58} Chaolong et al.\textsuperscript{59}</td>
<td>To get the benefit from updating the model parameters using new measurement data and to deal with small data</td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>Guler\textsuperscript{7} and Chaolong et al.\textsuperscript{59}</td>
<td>To predict degradation by considering a large number of influencing factors</td>
</tr>
<tr>
<td>Time series analysis</td>
<td>Chaolong et al.\textsuperscript{56}</td>
<td>To estimate the next degradation level using recent measurement data.</td>
</tr>
<tr>
<td>Stochastic process</td>
<td>Meier-Hirmer et al.\textsuperscript{41} Zhu et al.\textsuperscript{43} Mercier et al.\textsuperscript{44} and Vale and Lurdes\textsuperscript{45}</td>
<td>To capture the uncertainty of track geometry degradation over time (time continuous)</td>
</tr>
<tr>
<td>Markov chain</td>
<td>Yousefikia et al.\textsuperscript{24} Bai et al.\textsuperscript{57} and Shafahi and Hakhamaneshi\textsuperscript{58}</td>
<td>To capture the uncertainty of track geometry degradation over time (time discrete)</td>
</tr>
<tr>
<td>Multistage linear models</td>
<td>Gou et al.\textsuperscript{60} Chang et al.\textsuperscript{61} and Guo and Han\textsuperscript{62}</td>
<td>To deal with the non-linear (mainly exponential) degradation path in a maintenance cycle</td>
</tr>
<tr>
<td>Data mining</td>
<td>Xu et al.\textsuperscript{49,50} Liu et al.\textsuperscript{51} Bai et al.\textsuperscript{53} Xu et al.\textsuperscript{54} and Berggren\textsuperscript{55}</td>
<td>To short-term prediction of track geometry condition</td>
</tr>
<tr>
<td>Path analysis</td>
<td>Lyngby,\textsuperscript{9} Hamid and Gross,\textsuperscript{13} Bing,\textsuperscript{14} Guler et al.\textsuperscript{36} He et al.\textsuperscript{48} Westgeest et al.\textsuperscript{47}</td>
<td>To identify the influencing factors on track geometry degradation</td>
</tr>
<tr>
<td>Models with random coefficient</td>
<td>Andrade and Teixeira\textsuperscript{54-66}</td>
<td>To capture the section-to-section variation in geometry degradation</td>
</tr>
</tbody>
</table>

will improve the track maintenance planning and increase the track lifetime.\textsuperscript{70}

In this paper, we define track geometry maintenance as all technical activities that are applied to retain and restore the track geometry condition to the acceptable thresholds, to meet safety and quality standards at lowest possible cost. Activities that may be applied in ballast maintenance are manual intervention, tamping, and stone blowing.\textsuperscript{3,4} According to Andrews,\textsuperscript{19} ‘A tamping machine packs the ballast under the sleepers. It does this by measuring the track geometry, calculating the required adjustments, lifting the track and inserting vibrating tamping arms either side of a sleeper’. Stone blowing is executed with a special track vehicle that measures the current track geometry parameters and compares them with desirable values. It then tries to remove the gaps and improve track geometry condition by injecting ballast under the sleeper. The most important benefit of stone blowing over tamping is that in the stone-blowing process the position of existing ballast will not change and the ballast particles will not break up.\textsuperscript{19} Selection of a proper alternative between tamping and stone blowing depends on the availability of tools and equipment and the track maintenance history.\textsuperscript{71}

Tamping can repair vertical and horizontal track geometry defects, but it has a destructive effect on ballast. The tamper arms can break the ballast stones.\textsuperscript{19,71} After a specific length of time, tamping frequencies will increase dramatically, causing the maintenance cost to also increase. In this situation, ballast renewal could be used to restore track geometry to an ideal condition.\textsuperscript{27} As the age of ballast increases and with recurrent tamping interventions, the track geometry degradation accelerates, and a reduction in tamping efficiency might be observed. Whilst tamping will improve the track geometry condition, it cannot rejuvenate the geometry condition to an as-good-as-new state.

Tamping is an imperfect maintenance action and its effect on track geometry condition is twofold: a jump reduction and a change in degradation rate after tamping. Most of the researchers have only considered the jump reduction as the important factor for estimation of tamping effectiveness. However, analysing the effect of tamping on degradation rate is also of crucial importance. Therefore, the necessity of modelling tamping effectiveness may be raised. These types of models are called restoration or recovery models. By combining the modelling of track geometry degradation between two consecutive maintenance interventions and also maintenance restoration, track geometry maintenance models may be designed. In order to evaluate tamping effectiveness, Arasteh khoy et al.\textsuperscript{25} compared the level of vertical track geometry before and after tamping and clustered the data associated with vertical track geometry into three classes: bad, good, and excellent. They observed a significant variability in the tamping efficiency within different track sections and found that
the tamping efficiencies within sections were lying on a neighbourhood of a bad cluster. They also studied the effect of ballast age on tamping efficiency and found no significant relationship between tamping efficiency and ballast age. However, this result is contradictory to other research and more studies are needed to accept it as a general rule.

Arasteh Khouy et al. developed a model to optimise the tamping ILs with the objective of minimising total maintenance cost. These researchers applied exponential regression trend analysis to estimate constant degradation rates of different 200 m track sections. They took the short wave longitudinal level as the track quality measure and two specific twist defects to represent the safety faults, with a Weibull distribution representing their occurrence probability. The maintenance cost objective function included inspection cost, corrective tamping cost, accident risk cost, and capacity lost cost. Tamping effectiveness was computed by assessing how much the longitudinal level was reduced after tamping. They found that performing tamping interventions at higher limits was not efficient in terms of either energy consumption or ride comfort. In addition, performing tamping interventions at higher intervention thresholds might lead to significant speed reduction.

The recovery after tamping can be modelled using deterministic or probabilistic approaches. The selection of the approach should be made according to the level of uncertainty in the recovery values after tamping. In the deterministic approach the recovery value after tamping is directly modelled as a function of the influencing variables, such as the track geometry condition before tamping, the speed, and the maintenance history. In this approach, the model parameters are considered unknown but deterministic (constant), and the uncertainty will be captured through a confidence interval. The deterministic approach is normally used when the model uncertainty is low.

In the probabilistic approach, the uncertainty is captured through modelling the recovery value after tamping by selection of a proper distribution, while the parameters of the distribution are a function of the influencing variables. In this approach, for every set of input variables, a unique distribution will be obtained for the recovery value after tamping. The probabilistic approach is normally used when the model uncertainty is high.

A number of studies have estimated tamping recovery using a deterministic approach. They applied linear regression and considered that the tamping effectiveness is dependent on the track quality condition at the moment before tamping. This approach is used to develop track maintenance models and to optimise scheduling.

Meier-Hirmer et al. proposed an optimisation model for minimising track maintenance cost. They applied the Gamma process to model track geometry degradation and longitudinal level is selected as the track geometry condition indicator. They assumed that track inspections are perfect during the track life cycle. Linear regression is used to model maintenance intervention recovery based on the track degradation level before performing maintenance.

Vale et al. selected a linear model to represent the recovery of track quality by tamping intervention to track quality condition before tamping. This recovery value was combined with the degradation model to show the track quality condition in each section and time period.

A number of the other researchers have used linear regression to develop tamping restoration models which are used for scheduling track maintenance activities, e.g. Gustavsson, Famurewa et al., Oyama and Miwa, Li et al., and Miwa. The main advantage of these models is its simplicity. Therefore, in the cases that recovery values of tamping for a track line have a low level of uncertainty, this approach can be applied.

However, it should be noted that mostly there is a high variation in the recovery values after tamping even for similar track geometry condition before tamping. Since deterministic approach cannot capture this variation, some other researchers have assumed that the recovery value after tamping is a random variable with a given distribution. Audley and Andrews evaluated the influence of tamping on track geometry. Their aim was to find a distribution of time for degradation to a specific level. In addition, they fitted a distribution to model the probability of achieving a specific vertical level, after a maintenance intervention. When they fitted a two-parameter Weibull distribution on the generated data for different data sets, they found a significant trend between the Weibull parameters and the number of tampings. In addition, they observed that speed, maintenance history, and level of track quality had a profound impact on time to achieving the specific track quality state. In fact, in a high speed level and after a high number of tampings, the track was found to degrade quite a bit faster. They also fitted a two-parameter log-normal distribution on the probability of achieving a specific vertical level, considering different speeds and maintenance history. Increasing the speed and the number of tamping interventions decreased the probability of achieving a specific track quality. In another approach, Quiroga and Schnieder simulated railway track geometry degradation considering tamping interventions by using Monte Carlo simulation. They assumed that the longitudinal level after performing a number of tampings is a log-normally distributed random variable and the degradation development between two tamping interventions can be modelled by an exponential function. They observed that by increasing tamping intervention, the variances of initial longitudinal level after tamping and the degradation growth rate increased. They observed that the longitudinal...
level showed an accelerated growth in initial tampings, but this behaviour increased with more tamping interventions. This reduction showed an approximately constant behaviour after the fifth tamping. Finally, they evaluated the effect of speed on geometry degradation and observed that more tamping intervention is needed for greater operational speeds. The results of the research by Audley and Andrews\(^{71}\) and Quiroga and Schnieder\(^{78}\) demonstrate the importance of considering the effect of number of previous tampings and train speed on tamping effectiveness. However, they did not analyse the effect of tamping on other track geometry measures, e.g. alignment and cant. An interesting area for further research is to analyse the effect of tamping on different track geometry measures and performing correlation analysis to find out how tamping affects the different geometry measures. In addition, one should consider the fact that a tamping intervention may be performed on the whole length of a track section or on just a fraction of a section. Those tamping interventions that were performed on a fraction of a track section can be called partial tamping. Therefore, researchers should distinguish between partial and complete tamping interventions when estimating the recovery values after tamping. In this regard, a significant reduction in the variation of the estimated recovery values after tamping for a specific track geometry condition can be obtained. Another interesting area of study is analysing the effect of tamping intervention on degradation rate. Table 2 summarised the reviewed papers about modelling recovery of tamping.

After quantifying the tamping effectiveness, the next step is to combine the degradation and recovery models to obtain the optimal maintenance strategy. There are two methods for track geometry maintenance modelling, and these are the age- and degradation-based approaches. Simply stated, age-based modelling considers regular times for performing maintenance; whilst degradation-based strategies perform a maintenance activity when degradation exceeds intervention thresholds. Shafiee et al.\(^{79}\) developed an age-usage rail maintenance strategy that minimises the average rail infrastructure maintenance cost. The proposed methodology integrated these two approaches to simultaneously optimise usage and age thresholds. They assumed that all maintenance activities return the track to as-good-as-new condition, so the interval between two maintenance interventions is a renewal cycle. Their model incorporated the associated cost of maintenance activities and penalty costs due to traffic disruption and passenger dissatisfaction. They considered a heavy haul line in Sweden as a case study and compared the proposed methodology to age- and usage-based maintenance. They observed a reduction in the maintenance cost with the proposed methodology and an increase in traffic movement did not affect the optimal usage threshold.

In this review, we focus on degradation-based maintenance modelling. The available approaches for maintenance modelling are classified into Markov models, Petri net (PN) models, and simulation-based models. In these approaches, different degradation stages are defined based on maintenance and safety thresholds.

Quiroga et al.\(^{80}\) developed a simulation procedure by considering stochastic degradation and restoration models as well as a heuristic model for tamping scheduling optimisation. They considered three zones for track geometry quality, i.e. intervention, warning, and objective zones. Two different maintenance strategies in terms of constant and dynamic intervention thresholds are compared from required efforts for maintenance strategy and track geometry quality aspects. The monetary cost and time distributions for tamping interventions are considered as features of required efforts for the maintenance strategy and the longitudinal level is considered as the track geometry quality index.

Simpson et al.\(^{81}\) also proposed a track maintenance modelling approach to simulate and optimise the related cost of different maintenance strategies. The proposed model can evaluate the effect of traffic changes on maintenance and operation cost. The main output of the model is finding optimal maintenance strategies based on optimal thresholds for tamping and sleeper replacement thresholds. The researchers used the ITDM to predict track degradation condition, and they calculated train operation

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Applied model</th>
<th>Influencing variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linear regression</td>
<td>Probabilistic model</td>
</tr>
<tr>
<td>Gustavsson,(^2) Famurewa et al.,(^{70}) Meier-Hirmer et al.,(^{71}) Vale et al.,(^{74}) Oyama,(^{75}) Li et al.,(^{76}) and Miwa(^{77})</td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Audley and Andrews(^{71}) and Quiroga and Schnieder(^{78})</td>
<td>✅</td>
<td>✅</td>
</tr>
</tbody>
</table>
costs based on train delays. They also considered opportunistic maintenance in neighbouring segments. The strength of the model is the consideration of the rail, sleeper, and ballast maintenance activities simultaneously. In addition, the model can evaluate the effect of different track and traffic conditions for given maintenance scenarios.

Arasteh Khoy et al.\textsuperscript{11} proposed a model to optimise track geometry inspection intervals with respect to total ballast maintenance cost. Their study considered the Q-value as the track geometry measure and B-fault and C-fault limits were used to perform preventive and corrective maintenance. They used a two-parameter Weibull distribution to predict twist problems, B-faults, and C-faults. They observed that extending inspection intervals from two months to four months decreased the total maintenance costs. They argued that the reason for this is the low degradation rate for critical faults in the track. In addition, they found that more inspections are needed for sections with a high degradation rate. However, it should be considered that they assumed the track sections are identical and maintenance effectiveness is perfect. Relaxing these assumptions will make the model more complicated but will lead to a more realistic result.

A well-known approach for track geometry maintenance modelling is the Markov model. A Markov model consists of a number of different states that are linked by the possible transition paths. Different track geometry conditions can be considered as possible states for the Markov model and the degradation rates can manage the transition from one state to another one. The main output of the Markov models for track geometry maintenance modelling is the average probabilities of the track section being in different states. These results can be used to evaluate different maintenance strategies based on the long-term condition of track.

In order to consider the degradation, inspection, and maintenance of track, Prescott and Andrews\textsuperscript{82} applied a Markov model. The main aim of their study was to evaluate the effect on track condition by changing the intervention threshold. They considered the effect of maintenance history on the degradation process, dividing the degradation process into a number of phases with a constant degradation rate. The result of the proposed Markov model represented the expected time of the track being in different states, namely good condition, maintenance requested, speed restriction, and line closure.

Similarly, Prescott and Andrews\textsuperscript{83} modelled track asset maintenance management using a Markov model. They combined a degradation model with maintenance strategies to evaluate different concepts, such as cost and risk of different management approaches. They evaluated different maintenance strategies by varying a number of parameters, e.g. maintenance threshold, inspection interval time, renewal period, and mean time to perform normal priority maintenance. The model outputs included the mean time the track spent in different degradation level states and the probability of performing a specific maintenance intervention on the track within the planning horizon. The two mentioned proposed models are considered for an individual track section. In fact, it is considered that planning the maintenance for a track section is independent of the other sections in the line. However, in practice the maintenance activities are performed based on the condition of the whole track line.

Meier-Hirmer et al.\textsuperscript{84} proposed a maintenance model based on the gamma degradation process which is also considered fixed inspection intervals and delays of interventions. Intervention thresholds and inspection intervals were the model's decision variables. They evaluated the efficiency of tamping by computing the differences between longitudinal levels before and after tamping. They used a regression model to relate the gain from intervention to the degradation before tamping and observed that the intervention gain follows a normal distribution. In their work, the track states at inspection times follow a Markov chain. The transition rates for the Markov renewal technique are a function of deterioration before tamping, tamping effectiveness, and deterioration after tamping.

In another work, Podofillini et al.\textsuperscript{85} also proposed a methodology to optimise application of an ultrasonic inspection tool to minimise the related cost and risk of rail maintenance strategies. The main aim of the methodology was to evaluate the risk and cost related to a rail inspection strategy using a Markov model and minimise these values using a multiobjective optimisation model. Rail derailment was the risk criterion in the optimisation model. The results of the proposed methodology were optimised inspection intervals and optimised waiting time to perform maintenance. Corriere and Di Vincenzo\textsuperscript{86} used a Markov model and the dynamic programming optimisation method to model track maintenance planning based on the rail quality index (Italian IQB). The Italian IQB includes a number of defectiveness indices, i.e. longitudinal level, alignment, transversal level, and wedging.

It should be noted that using a Markov chain for track maintenance modelling has a number of limitations; it is restricted to just an exponential distribution for the estimation of the sojourn times. The transitions between track states should occur with a constant rate.\textsuperscript{19} In order to cope with the limitations of Markov chain approach in track geometry maintenance modelling, a number of researchers have recently used PNs. A PN is a graphical tool that includes nodes, arcs, and transitions. PNs are used as a description of distributed systems.

In order to model track geometry behaviour by considering the effect of tamping, Andrews\textsuperscript{19} used a PN. The ‘time to degrade to different states’ was
modelled for homogenous track sections, using Weibull distributions, for different track features (region, type of rail, type of sleeper, speed classes, and annual cumulative tonnage) and phase of life features (number and sequence of performed interventions). These distributions were used for transitions in the proposed PN model. The inspection intervals and intervention thresholds were decision variables. Andrews\(^4\)\(^5\) concluded that degradation between interventions has a linear behaviour and observed that tamping efficiency is not perfect and cannot recover the ballast to as-good-as-new condition. In practice, the tamping will not perform for a single section that exceeds the degradation threshold. In fact, the tamping intervention will be planned based on the condition of a group of track section. Therefore, more complicated models that can consider opportunistic maintenance should be applied to obtain the realistic result.

In order to extend the previous work, a PN is also applied by Andrews et al.\(^4\) to predict track degradation behaviour by considering different track maintenance strategies. They argued that the degradation rate is different between different intervention intervals. They used a Weibull distribution to determine the time required for the track to degrade to a particular quality level. Their model considered four decision parameters to evaluate various maintenance strategies: inspection time, renewal time, routine repair time, and intervention threshold. The effects of these parameters were evaluated with respect to number of interventions, time in different track conditions such as poor and good states, and the required number of speed restrictions and line closures. They found no significant relationship between inspection time and number of interventions. In addition, extending renewal times had no meaningful influence on the time spent in good condition. By involving associated costs in inspection and maintenance activities, along with penalty costs due to speed restrictions and line shutdowns, a life cycle cost model can be designed based on the proposed approach. Although the proposed model can properly predict the track geometry condition of a track section over a long-term period, however, the model is unable to predict the degradation at track line level. Moreover, the proposed model requires a large enough data set to estimate the transition rates. Especially, in order to taking into account the effect of number of preformed maintenance interventions on transition rates a long trend of measurement data for a track section is required. Finally, the proposed model considered the longitudinal level for planning maintenance activities. Since other track geometry measure may compete to reach the intervention threshold, considering alignment, cant, and twist in the model will increase the model flexibility.

The track geometry maintenance modelling for a track line is studied by Prescott and Andrews\(^3\) that applied the PN method to address this issue for track degradation, inspection, and maintenance modelling. One important issue is to consider tamping and stone blowing simultaneously in maintenance modelling. They considered that a limited number of tamping and stone-blowing machines should be assigned to different track sections. The model considered the practical limitations of performing tamping after stone blowing. The researchers used a two-parameter Weibull distribution for the degradation time distribution and considered four limits: opportunistic maintenance permitted, maintenance needed, speed reduction needed, and line closure needed thresholds. Based on these limits, they were able to define four states of degradation. Transitions among these states occurred by sampling from the degradation time distributions. The main difference between their work and previous studies is considering maintenance decision-making module in their proposed PN model. By doing so, the opportunistic maintenance can be considered and the single section model can be extended to the network level model. The decision for grouping the maintenance actions was made based on track section degradation states, availability of machines, and section locations. Considering total maintenance cost and line availability as the criteria for opportunistic maintenance decision making will increase model flexibility.

A hierarchical framework based on the PN method is applied in the work by Rama and Andrews\(^8\)\(^9\) to model track infrastructure systems and their different sub-systems. The proposed PN model has three main parts: track degradation, track inspection, and maintenance intervention. The degradation times between different states were obtained using given distributions. Overall, the proposed framework was able to evaluate the concept of opportunistic maintenance. In their work, the sleeper and rail are considered to construct the model. However, the same concept can be integrated with the other PN models mentioned in this section to construct a holistic railway infrastructure maintenance model. In fact, a PN model that can consider different track components as well as track sections along the track line is highly desirable for infrastructure managers. In order to achieve this goal, the effect of the state of each individual track sections and track component (rail, sleeper, etc.) on the state of whole track line should be identified. In addition, the effect of performing opportunistic maintenance on the whole track line should be determined. However, considering all this issues together will dramatically increase the model complexity.

Recently, a framework based on hierarchical coloured PNs has been proposed to model track risk and maintenance. Shang and Berenguer\(^88\) considered the gauge spread as the track geometry defect in their study and a multistage model is used to predict the gauge degradation over time. They assumed that gauge degradation follows an exponential distribution in each level. The proposed framework consisted of
three phases, namely component, system, and operational phases. Different inspection policies may be evaluated through the proposed framework in terms of availability, risk, and cost aspects.

Another application of PNs can be found in the work by Quiroga and Schnieder. They proposed a framework based on PNs to evaluate the influence of different organisational maintenance strategies, such as new assignments of human and technical resources, on track availability. Their model incorporated four reliability states: faultless, detected partial faults, undetected partial faults, and faulty. The model considered that failure occurred in two ways, first, through degradation and second, through random shocks. The results demonstrated that when maintenance is optimally assigned, a significant improvement in track availability can be obtained. In Table 3 the recent papers that used Markov-based and PN-based approaches for track geometry maintenance modelling are classified and the main characteristics of each approach are provided.

A number of researchers have also developed decision support approaches and frameworks for track maintenance. Based on decision rules, Guler developed a decision support framework to evaluate the condition of track components and propose optimal times for maintenance and renewal operations. The rules were generated based on expert opinions, a track component degradation model, and a track component intervention threshold. The maintenance and renewal activities for ballast, sleeper, rail, and track bed layers are considered in the study. Therefore, the benefits of the opportunistic maintenance are considered in the framework. In addition, since other parameters such as track structure and components properties are considered along with degradation models, more practical decisions for track maintenance activity can be obtained. Since the applied rules are based on expert opinions, the implementation of the proposed framework requires a set of experienced experts with high knowledge about track maintenance. The proposed framework has the capability to upgrade itself by considering up to date degradation models.

Developing a framework is the subject of work by Lovett et al., who look for planning and scheduling track maintenance activities. First, they used a Weibull curve to represent the track degradation model. Second, they chose case-based reasoning (CBR) for maintenance work selection. In addition, the benefit/cost ratio was used to evaluate selected maintenance activities. Risk reduction related to derailment and other defective track failures was considered benefits in this model. Third and finally, they suggested a scheduling optimisation model based on the knapsack approach. Time and location of maintenance activities were decision variables, with budget and time thresholds for performing maintenances as the constraints. The proposed framework is not limited for track geometry maintenance and different track failure modes can be considered. The applied degradation model in the framework is basic and as-good-as-new assumption is considered for maintenance activities. The proposed approach based on CBR can be applied for selection of the proper maintenance activity among tamping, stone blowing, manual intervention, and ballast renewal for rectifying track geometry defects.

By combining an ANN and a fuzzy inference system, a decision support system for tamping intervention planning is provided by Dell’Orco et al. The intervention thresholds were considered linguistic variables in the fuzzy logic. The track degradation was modelled using an adaptive neural network. To consider opportunistic maintenance, they clustered tamping interventions using the fuzzy C-mean method in both time and space.

To sum up, with respect to the characteristics of the PN, there is an increasing trend to use the PN for track geometry maintenance modelling. Since shock events usually happen in track geometry degradation path, they should be considered for planning corrective or even preventive maintenance activities. Developing a module in the PN for considering the shock events will increase the model efficiency.

As a conclusion in order to design a track maintenance model and enhance the efficiency of available approaches a number of issues should be considered. The different modes of track geometry defects should be considered in the model. Accurate and robust degradation and recovery models for each of the track geometry measures should be employed for long-term prediction of track condition. In addition, the effect of shock failures on track geometry degradation levels

| Table 3. Summary of recent papers that used Markov- and Petri nets-based approaches for track maintenance. |
|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| **Markov model**                                | **Petri nets**                                   | **Characteristics**                               |
| Prescott and Andrews                            | Prescott and Andrews, Andrews et al., Andrews, Rama and Andrews, and Shang and Berenguer | Constant transition rates are applied. It is used for the individual track section level. Different distributions can be used for transitions among states. The approach is used for both individual track section and line levels. |
should be evaluated. The opportunistic maintenance for grouping the major maintenance and renewal works of track components must be considered. The both of preventive and corrective maintenance activities should be considered in the framework. The maintenance intervention thresholds and inspection intervals should be optimised with respect to life cycle cost as well as reliability, availability, maintainability, and safety parameters. The effect of operational condition of track (e.g. track type and traffics) on different maintenance strategies should be evaluated.

Scheduling geometry recovery operations

After obtaining the optimal maintenance strategy by determining the optimal inspection intervals and intervention thresholds, appropriate times should be assigned for conducting maintenance activities. This needs proper scheduling of maintenance activities along a railway track. The main goal of scheduling the track maintenance activities is to minimise the needed time for maintenance intervention and overall maintenance cost whilst preserving the railway track in desirable reliability and availability limits, subject to the practical constraint about available maintenance resources. Mostly, the problem is modelled using mixed integer programming. Since the proposed mathematical programming models need massive computation, different approaches such as heuristic methods are also applied in the literature. The main objective function of the available models is based on cost minimisation for track geometry scheduling optimisation. The maximisation of track quality condition improvement in terms of reducing track surface irregularities is considered as the model’s objective function in the work by Oyama and Miwa. An all integer linear programming model is deployed by Oyama and Miwa who investigate the area of scheduling tamping interventions by considering track degradation and restoration. The logistic distribution was fitted on real data to represent the growth of track irregularities. The researcher applied the exponential smoothing method on real data to construct the track degradation for each track section and used regression to compare the track irregularities before and after tamping intervention and to find a mathematical model for track irregularity restoration. In order to select potential track sections for improving track quality condition by applying tamping, the researcher applied the maintenance unit selection model. Finally, he tested the model with real data obtained from a track line in Japan and found the model to be efficient in terms of reducing track irregularities. Based on the problem variables, the optimisation model can be deterministic or probabilistic. Employment of probabilistic mathematical optimisation model can capture associated uncertainty in scheduling problems and may lead to a more robust solution. In addition, in this approach the units are selected such that the total expected geometry recovery obtained due to tamping is maximised. Although such model has its own benefits, of maximising geometry recovery of track units, however, it does not consider availability of track line and associated maintenance costs. Application of opportunistic tamping can support a more effective grouping of track section and to arrive at a more cost-effective solution at higher availability performance.

Minimising the number of preventive tamping interventions is another alternative for the objective function of a scheduling model, and this is employed in the work by Vale et al. They developed a mathematical programming model based on mixed 0–1 integer programming. The optimisation model considered four issues as constraints: degradation rate of TQI over time, track layout, track quality limits, and effect of condition of track quality in the maintenance moment of the track quality recovery. These researchers used linear regression to estimate the degradation rate. Another approach for geometry maintenance scheduling is to minimise the number of tampings over all track sections and time horizons and this is the objective function of the study by Vale and Ribeiro. Three constraints are addressed in the work by Vale and Ribeiro who developed a mixed 0–1 non-linear programming model by considering stochastic degradation. These constraints are as follows: (1) track quality recovery after tamping is dependent on initial quality of track, (2) track quality should be within predetermined thresholds, and (3) the start and finish of tamping operations must be located on a straight line. The degradation rate of track geometry measure was estimated using a three-parameter Dagum distribution and maintenance interventions over time were assumed to follow the Generalised Pareto distribution. Considering stochastic degradation process will improve the model accuracy and will lead to more realistic maintenance scheduling plan.

Minimisation of the maintenance cost whilst keeping track surface irregularities at a safe level is the objective function of the work by Miwa. He developed a mathematical programming model to optimise the track irregularity maintenance scheduling. The exponential smoothing method was used in this study to model track degradation and linear regression was applied to represent the track condition restoration after tamping. Developing a decision support system framework by considering degradation and restoration models for tamping scheduling optimisation is the subject of the work by Miwa et al. They fitted the logistic distribution on track irregularity data to express the track condition, with the parameters of the distribution related to type of alignment, rail, and sleeper; depth of ballast; and maintenance history using the exponential
smoothing method. The restoration model was developed by comparing the track irregularities before and after tamping intervention and by considering the track's structural characteristics. They developed an integer programming optimisation model to schedule tamping interventions aiming to minimise the average track surface irregularities. The proposed degradation model considers the section-to-section variation in geometry degradation and assumed that tamping will be conducted on all the sections in a unit. Relaxing this assumption by considering opportunistic maintenance may dramatically decrease the total maintenance cost.

Comparing different objective functions of scheduling models is in part the subject of the work by Quiroga and Schneider who proposed a heuristic method to optimise the tamping scheduling problem. The three objective functions are as follows: (1) minimisation of total reduction track geometry deviation, (2) expected time to failure, and (3) expected deviation of longitudinal level at next campaign. They first proposed a track geometry degradation model that consisted of two parts: the achieved value after a tamping action that followed a normal distribution and the degradation between two consecutive tamping actions that evolve based on an exponential function. The researchers used a stochastic noise variable to deal with small variations in longitudinal levels due to measurement errors. The heuristic algorithm specified an upper bound for the objective function and then found the optimal feasible solution. They validated the model using data from a track line in France and applied the last objective function in their model. They observed the optimal solution of the proposed model was very close to the upper bound. In addition, the proposed model significantly reduced the calculation time.

One of the important issues in track geometry maintenance scheduling is considering renewal activities parallel to optimising tamping activities. Zhao et al. pointed out this issue and developed a model to optimise ballast tamping and renewal activities with respect to renewal, tamping, and penalty costs, simultaneously. Their non-linear model included a combination of initial and linear degradation to model TQI deterioration from traffic load (tonnage). In addition, they assumed that ballast condition restoration would decrease after each tamping intervention with a specific ratio. The work examined three maintenance policies: fixed intervention level, constant interval of tamping, and optimal non-constant intervals of tamping. By comparing the three mentioned policies, the researchers discovered that the latter policy has more benefits in terms of Life Cycle Cost (LCC) and ballast service life than the other two policies. It should be noted that the proposed model consider that track sections are identical.

Santos and Teixeira stress the importance of determining the optimal track length for performing maintenance with a tamping machine. To do so, the simulated annealing method is employed to optimise the maintenance intervention scheduling for a long period. Their work featured a basic degradation mode, with different constant track degradation rates assigned to 10 km track sections using a normal distribution. The longitudinal level was the measure of track condition and the model's cost function included operational and transportation costs related to tamping.

Aiming to find the time to next inspection that will guarantee a high safety level for the railway track, Mercier et al. address the application of a bivariate gamma degradation process in maintenance scheduling. They observed that when two degradation indicators are considered at the same time, there is a significant reduction in time to next maintenance intervention. In fact, by considering the dependency between the two degradation indicators, the researchers were able to increase the required number of inspections to achieve the same safety level as if they had considered these indicators separately. They concluded the safety level achieved by considering the dependency of degradation indicators was significantly higher than the level achieved by considering the indicators individually or separately.

As a best practice in railway engineering management, a track line is divided into large numbers of track sections. Obviously, planning the maintenance activities for individual track section is not practical in real life.

To solve this problem, a number of researchers combined several sections as a unit and planned the tamping for each track units. Although the obtained scheduling plan by this approach may be practical, it may not be optimal in the terms of cost and availability. In order to deal with these challenges, opportunistic maintenance can be modelled in terms of time, space, and components. In this policy maintenance decisions of a track section are based on the geometry condition of other track sections in a given neighbourhood. Since track is an integrated system such that its function is dependent on all of its components and their interaction, scheduling different track component maintenance simultaneously can create large benefits in terms of cost and availability.

The importance of allocation of track possession times is pointed out by Famurewa et al. The applied objective function in their model is minimising direct tamping intervention costs and indirect costs related to possession times whilst keeping track geometry parameters at acceptable levels. Their study looked at both preventive and corrective tamping interventions. These researchers used an exponential function to represent the degradation rate of the longitudinal level for the sections. An empirical regression method was selected to model tamping effectiveness. The track possession time included travelling, setup, tamping,
and takedown times. They observed the degradation rate of different sections to be non-homogeneous and the distribution of track section degradation to be strongly skewed to the right. Therefore, they argued that track sections cannot be considered as identical units. In addition, increasing the number of preventive tampings reduced the required corrective tamping interventions. The proposed model can be used to obtain the optimum number of preventive tamping interventions that minimise track possession time and tamping intervention cost (preventive and corrective). They considered that tamping interventions can be grouped for the track sections that passed the maintenance threshold and are located close together.

The subject of opportunistic maintenance in adjacent segments and for different track components is investigated in the work by Caetano and Teixeira. They evaluated the trade-off between the economy of scale benefits and loss of life due to opportunistic maintenance. A mixed integer linear programming model that integrated track components’ degradation is developed to optimise track renewal intervention scheduling with respect to the track’s life cycle cost. The longitudinal level and alignment were the track geometry parameters used and a linear regression model was applied to model their degradation. The LCC model consisted of renewal intervention costs, track maintenance intervention costs, economy of scale benefits, and residual value of track components at the end of the planning horizon. They evaluated the effect of the available budget on track life cycle cost and found budget limitations could restrict the performance of an optimal renewal plan and increase track LCC. In their case study, they observed that the ballast maintenance cost was higher than that for the rail and sleepers; moreover, the opportunistic maintenance of rail and sleepers was commonly based on the ballast renewal schedule. The proposed model revealed that considering different track component degradation models and opportunistic maintenance in track renewal modelling significantly can reduce the track life cycle cost. The model outputs are optimised renewal intervals and the identification of track components and section groups for opportunistic maintenance. It should be considered that the proposed model is for planning the renewal activities. This approach can be extended to consider the opportunistic maintenance and renewal activities simultaneously.

The importance of opportunistic maintenance is highlighted in the work by Zhang et al. who developed a mathematical programming model to optimise track maintenance scheduling. They assumed that the degradation time is a Weibull random variable. The model sought to minimise four costs: cost due to unsafe transportation, cost of loss of useful life, cost of maintenance, and cost of travel. Since the proposed model is NP hard, these researchers selected an enhanced genetic algorithm to find the optimal solution. The proposed model considers the track sections that are waiting for maintenance and optimally assign them to different maintenance teams in different times.

Opportunistic maintenance in terms of different components is in part the subject of work by Gustavsson where a mixed integer linear programming model was developed to optimise tamping scheduling. The track line was considered a multicomponent system and cost dependencies among the track sections were included in the study. The model extended a previous one developed in Vale et al. The objective of the mathematical optimisation model was to minimise the cost of tamping interventions in track sections and the related cost of maintenance at any time step. The study assumed that the condition of track segments at every time step was only dependent on its condition in the last time step and integrated the linear and exponential degradation models to represent the track irregularities. A degradation model considered the increasing degradation rate due to degraded track components, whereas a linear model represented the tamping recovery of track segments. In this study, Gustavsson employed a disaggregated constraint in their model that created a stronger lower bound for the branch-and-bound algorithm to deal with the issue. To demonstrate its cost-saving benefits, the proposed model was compared with the greedy policy and the age policy in different conditions, i.e. constant and increasing degradation and different setup costs. The proposed model proved better than the other two policies in terms of the required cost for performing the scheduled tamping. The main focus of Gustavsson’s work is on computational aspect of the optimisation model.

The importance of considering the tactical horizon for track geometry maintenance scheduling is pointed out in the work by Li et al. where a mixed integer programming model is developed to schedule tamping activities. The objective function of the proposed model was to minimise the total tamping cost. The proposed model considered the merits of economies of scale from the point of view of both time and space. They made the following assumptions: track degradation has a linear behaviour; the degradation rate is constant over time; and based on the ORE model, tamping recovery is linearly dependent on the initial track quality at the beginning of tamping. The track degradation thresholds for different speeds were considered constraints in their model. The tamping machine costs consisted of warming up, tamping, ramping down, and driving costs. Due to opportunistic maintenance, they found a significant reduction of nearly 50% in total tamping cost using the proposed model. The model provided more flexibility for changing tamping times at the tactical level. They concluded using tactical planning instead of yearly planning would result in lower maintenance costs.
In addition, a longer tactical planning horizon would reduce the total tamping cost. They suggested that by applying the total tamping cost instead of the number of tampings as the model’s objective function created more practical solutions with lower maintenance cost.

The subject of considered routine tasks and projects simultaneously is addressed by Budai et al. who modelled the track preventive maintenance scheduling problem for medium-term planning in a finite time horizon. The model objective was to minimise the possession cost and track maintenance cost. Opportunistic maintenance was developed in the study by clustering different routine tasks and projects to reduce the track possession time. Based on test results, researchers argued that more available time for performing maintenance activities and more possibilities for grouping routine tasks and unique projects will reduce the track maintenance cost. The proposed approach did not consider the degradation and recovery models. In fact, the application of the proposed approach is scheduling and clustering a specific set of maintenance activities for a given railway track condition.

A decision support framework is developed by He et al. to rectify track geometry defects. The proposed framework combined three phases, namely a statistical degradation model to predict the degradation behaviour of different track geometry defects, a model to evaluate the derailment risk, and an mixed integer linear programming model to optimise defect rectification activities aiming to minimise the total expected maintenance cost and derailment risk probability. It is considered that the degradation rates of track geometry defects have an exponential relationship with monthly traffic, monthly total number of cars and trains, and number of inspections. The corrective or unplanned maintenance were considered track rectification activities in their study, with different track geometry defects classified as yellow or red tags. They remarked that two different approaches can be used to optimise rectification activities. Local track masters tend to use models aiming to minimise the derailment risk, but a cost-based model is used to keep the total track maintenance cost at an acceptable level. One of the advantages of the proposed approach is that it considers different track geometry defects simultaneously. In addition, considering different objective functions for track maintenance scheduling is one of the important issues that are addressed in this research. It should be noted that the proposed approach does not consider recovery model for long-term prediction of track geometry condition.

A major interest of railway managers is developing an integrated model to optimise the planning of track components’ maintenance. This area is the subject of the work by Zhao et al. who developed a mixed integer programming model to optimise the scheduling plans of rail, sleeper, and ballast renewal simultaneously. This model’s objective was to maximise the cost benefit of combining different renewal activities in both time and space aspects. Because of the difficulty of the proposed model, researchers used a genetic algorithm to obtain its solution. The model did not use track component degradation models; rather, component renewals were planned based on time and tonnage. Since the model schedule the renewal activities, the track component will back to as-good-as-new condition after the intervention. Relaxing the assumption of fixed life time of track components by using degradation and recovery models will improve the decision efficiency. Table 4 summarises the characteristics of the reviewed papers for track maintenance scheduling.

To make a conclusion from the reviewed papers, the most important issues for railway track maintenance scheduling are presented here. The main inputs for scheduling models are degradation and recovery models. Applying a proper degradation model that can capture time and section-to-section variation of track geometry degradation and using the probabilistic approaches for modelling the recovery of tamping will result in a more accurate estimation of track condition for planning the maintenance activities. In order to corporate these degradation and recovery models in scheduling problem, the probabilistic optimisation model should be employed. The other important point that must be considered for track maintenance scheduling is considering the occurrence of random shocks. Since random shocks are a common event in track geometry degradation, they should be involved in the models to obtain a practical scheduling plan. Since the track geometry condition can be represented by different measures, considering the longitudinal level, alignment, cant, twist, and gauge problems at the same time will lead to a more efficient schedule plan. As mentioned before, the most important issue in track maintenance scheduling is opportunistic maintenance in terms of time, space, and track components. Without considering the opportunistic maintenance in terms of space, the obtained scheduling plans are proper for track section not for track line and will lead to an impractical plan.

It should be noted that a basic assumption in railway management is that operation and maintenance tasks shall both be scheduled together in a globally efficient way. Although the purpose of the infrastructure is to enable the movement of train traffic, train operations may not always have precedence over maintenance. Instead they are both needed and are dependent on each other – without maintenance train operation will in the end be impossible, and without operation the maintenance is of no interest. Thus, a healthy balance must exist, which should be reflected in the planning and scheduling processes and tools. Therefore, planning the maintenance activities considering train traffic is another challenge to arrive at the most effective solution. In this regard, Lidén
Table 4. Summary of reviewed papers for track maintenance scheduling.

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Objective function</th>
<th>Opportunistic maintenance</th>
<th>Model variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not considered</td>
<td>Considered for fixed units</td>
<td>Track quality index</td>
</tr>
<tr>
<td>Oyama and Miwa(^7) &amp; Maximising the total expected improvement by the planned maintenance &amp; ✓</td>
<td>Surface irregularities</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Vale et al.(^7) &amp; Minimising the number of preventive tamping interventions &amp; ✓</td>
<td>Longitudinal level</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Vale and Ribeiro(^9) &amp; Minimise the number of tampings &amp; ✓</td>
<td>Longitudinal level</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Miwa(^7) &amp; Minimising the mean value of the standard deviation of surface irregularities &amp; ✓</td>
<td>Surface irregularities</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Miwa et al.(^9) &amp; Minimising the mean value of the weighted surface irregularities &amp; ✓</td>
<td>Surface irregularities</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Quiroga and Schneider(^9) &amp; Expected longitudinal level at next time interval for tamping &amp; ✓</td>
<td>Longitudinal level</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Famurewa et al.(^7) &amp; Direct and indirect tamping intervention costs &amp; ✓</td>
<td>Longitudinal level</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Zhao et al.(^9) &amp; Total cost per unit traffic load per unit track &amp; ✓</td>
<td>Longitudinal level</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Mercier et al.(^4) &amp; Not considered &amp; ✓</td>
<td>Longitudinal and transversal levels</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Caetano and Teixeira(^7) &amp; Minimising the life cycle cost &amp; ✓</td>
<td>Longitudinal level and alignment</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Zhang et al.(^9) &amp; Minimising the costs of safety, useful life loss, maintenance, and travel &amp; ✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gustavsson(^2) &amp; Minimising the total maintenance cost &amp; ✓</td>
<td>Longitudinal level</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Li et al.(^7) &amp; Minimising the total tamping cost &amp; ✓</td>
<td>Longitudinal level and alignment</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Zhao et al.(^1) &amp; Maximising the cost benefit obtained by opportunistic maintenance &amp; ✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
wrote a report aiming to explain and classify railway infrastructure maintenance work from a planning and scheduling point of view, in which train traffic and operational restrictions are considered. He concentrated on planning the defined maintenance activities, whilst the study also focussed on track maintenance modelling based on the analysis of track degradation behaviour. Combining these studies will enhance our understanding of about track maintenance modelling. In another research, Soh et al.102 also reviewed the scheduling techniques for railway preventive maintenance that aims to minimise the overall maintenance cost and optimise makespan of preventive maintenance.

**Discussion**

Railway track degradation and maintenance modelling has established an important research area. The increasing number of publications in this area confirms this trend. To begin with, to represent the track quality condition various researchers applied different TQIs. Mainly, the three indicators, i.e. mean value, standard deviation over a specific length, and extreme values of track geometry parameters are used to define TQIs. Based on the reviewed papers an integrated TQI that can properly represent the track quality condition should: (1) represent the overall condition of track geometry and individual condition of track geometry parameters simultaneously, (2) consider the different weights for different track geometry parameters by using justified coefficient in the formulation, (3) consider the track class and the maintenance history of the line, (4) use different indicators of track geometry parameters (i.e. mean value, standard deviation, extreme values), and (5) represent both geometrical and structural defects. Another approach for representing the track geometry condition is using the individual indicators of track geometry parameters. The benefit of this approach is that since each of the track geometry parameters has a unique degradation path, the concept of competing failure modes can be applied for degradation modelling.

In order to model track geometry degradation in a maintenance cycle, many approaches are used in the literature, namely linear and exponential regression, multistage linear model, grey model, time series analysis, path analysis, ANN, data mining, models with random coefficient, Markov model, and stochastic process. Among the proposed models, linear and exponential regression models are widely used in the literature due to the simplicity and ability to represent the underlying degradation path. However, it should be considered that the regression model cannot be updated by new data and are not dependent to the previous observations. In addition, these models need a large amount of measurement data to have an acceptable accuracy. In this regard, grey model and time series analysis are used in the literature to overcome the mentioned issues. It is widely pointed out that different track sections degrade differently from each other. This kind of difference in degradation is called section-to-section variation, due to the variability in track structure, traffic, environment, maintenance history, etc. In order to find the effect of different factors on track geometry degradation, the path analysis is used in the literature. By using the path analysis the potential factors that cause different degradation behaviours of different track sections can be identified. In order to capture the section-to-section variation, the random coefficient models are used in the literature. In models with random coefficients, the section-to-section variation is captured by considering a random effect representing the individual section characteristics. In addition to section-to-section variation, the time-varying dynamic of track geometry degradation should be considered in degradation modelling. In order to capture the uncertainty of degradation path over the time, the discrete and continuous stochastic models are used in the literature. However, still there is need for more advanced stochastic processes to accurately model track geometry degradation over time. In addition, only a few researchers considered the shock events in the degradation modelling. In addition to normal degradation, railway track is also exposed to random shocks that may increase the degradation level. Random shocks in degradation models are assumed to be negative, causing the increase in degradation level and even the failure of the system. In general, failure is a result of the interaction and competition of performance degradation and random shocks. To sum up, an interesting field for future research is proposing a data-driven-based model with random effects and by considering random shocks for track geometry degradation modelling.

In order to model recovery of tamping, two main approaches are used in the literature, i.e. linear regression and probabilistic model. It is observed in the literature that the recovery values after tamping even for same track geometry condition before tamping are variable. The linear regression is a deterministic model and cannot capture the variation in the recovery values after tamping. Hence, in the case of high variations in recovery values after tamping, using a probabilistic model has advantage over linear regression. Another issue that should be considered for tamping recovery modelling is identifying the partial and complete tamping. Since partial and complete tamping interventions have different effect on track geometry condition, they should be analysed in different ways. A challenge for further study is clustering the partial and complete tamping and applying a probabilistic model with random coefficients to model tamping recovery. In addition, almost all the recovery models in the literature defined the recovery value of tamping as the difference between standard deviation
of a geometry parameter before and after tamping. In order to extend the previous works and to analyse the variation in recovery values after tamping, the individual measurements recorded in every 25 cm and the associated distributions before and after tamping intervention must be analysed.

In order to predict the long-term condition of track geometry and to model track maintenance, different approaches are used in the literature. The most important applied approaches are Markov model and PNs. The main objective of these approaches is to find the optimal inspection interval and maintenance intervention thresholds for planning track geometry maintenance with respect to direct and indirect maintenance costs. Initially, researchers used Markov model for track geometry maintenance modelling. However, as mentioned before Markov model has some limitations such as constant transition rate that restrict its application for track geometry maintenance modelling. In this regard, in the recent years there is an increasing trend for using the PNs. By applying PNs different distributions can be used to control the transitions among different states of the model. The other important point that should be pointed out is that in the past most of the track geometry maintenance models were applied for an individual track section. However, since track line is composed of many track section, the recent studies tried to find the optimal maintenance strategy for the whole track line. A challenge for further study is considering probabilistic models with random coefficient for degradation and recovery modelling in maintenance models. In addition, considering the random shocks in the model that urge performing the corrective maintenance activities is another challenge. Another interesting area for future research is considering different track geometry measures and also different track failure modes, e.g. sleeper and rail defects to propose a maintenance model based on competing failure modes.

The last phase of track maintenance modelling is scheduling to find the right times for performing maintenance activities. Various approaches with different objective functions are proposed in the literature for track maintenance scheduling. The following tasks should be done to obtain a proper track maintenance scheduling: defining an objective function with respect to RAMS parameters and track life cycle cost; applying robust degradation and recovery models; considering practical constraints of maintenance activities; and involving the concept of opportunistic maintenance in terms of time, space, and track components. Considering the opportunistic maintenance and applying multiobjective functions for track maintenance scheduling is another area for future study.

Conclusions and directions for future research

The aim of this study is to identify the existing gaps and challenges for track geometry degradation and maintenance modelling and propose some directions to address these challenges in order to develop new methodologies for robust and reliable track geometry maintenance.

The challenges and opportunities for railway track degradation and maintenance modelling may be categorised into four areas: Finding a proper track geometry indicator, predicting track geometry behaviour accurately, modelling track geometry recovery after maintenance and modelling track maintenance strategies, as well as maintenance scheduling.

Concerning track geometry measures, it is observed that several organisations have developed their own indices to represent the condition of track geometry. However, the majority of the researchers have used longitudinal level as the critical factor. It should be noted that ignoring other measures such as alignment and twist may lead to an unrealistic understanding of the track condition. An interesting area for further research is to develop a composite index in which defects associated with geometry and structure are considered together. This facilitates a better understanding of the track system degradation and supports a more effective and efficient maintenance decision.

Railway track is a distributed system, and different heterogeneous factors affect track degradation behaviour along the track lengths. Therefore, the track must be divided into a number of shorter track sections to enable degradation and maintenance analysis as each position along the track may need different types of maintenance actions. Constant equal length sections are usually considered for planning track maintenance activities, but a more efficient method is to cluster sections based on their structural, environmental, and operational characteristics, to make practically optimal segmentation and for effective use of service windows.

Obviously, uncertainty is the major characteristic of track geometry behaviour and mechanistic approaches are unable to capture this. It is observed that there is an emerging trend to use statistical methods for modelling
track geometry degradation. One reason is that the significant improvements in track geometry measurement technologies and systems which have taken place over the past decade have enabled access to huge amounts of data. Second, computational techniques have evolved tremendously facilitating faster computations in the analysis process. However, the results obtained through statistical approaches need to be supported by physical models developed using mechanistic approaches. An interesting area for further research is combining mechanistic and statistical modelling, which is a hybrid approach, to deal with the innate uncertainties of track geometry behaviour.

According to the literature, a majority of track geometry recovery studies are based on the ‘as-good-as-new’ assumption. Hence, they have assumed that tamping can return the track quality to its original condition. Obviously, there is a need for a robust and effective model that properly estimates tamping effectiveness. One of the most important prerequisites for modelling tamping recovery is to have a reliable database that includes sufficient information about all the performed maintenance activities, e.g. the time and location of all maintenance interventions along the track as the most key data. However, such a complete database is not available. Therefore, different approaches should be selected to identify the time and locations of the performed tamping on track sections in the databases. In fact, there is a need to develop a methodology for analysing the massive data available, so-called big data, and to define some criteria about the jump reduction or change in degradation rate after tamping. This enables identification of the time and location of tamping interventions along the track.

Tamping effectiveness is dependent on various factors, e.g. the number of previous tampings, initial track quality condition, and the level of quality condition just before tamping, etc. In order to propose a proper tamping recovery model, the influencing factors and their contribution to tamping effectiveness should be identified. Although very few researchers have dealt with this issue, it is still an important area for further study and research. In addition, a number of studies have considered the longitudinal level to assess the effectiveness of tamping. However, it is also essential to address the other measures such as alignment and cant for evaluation of tamping effectiveness.

In the recent years, there has been an obvious trend towards using simulation-based methods for modelling of track geometry maintenance strategies. This will change the practices of current maintenance modelling and analysis to arrive at a more effective and efficient track geometry maintenance decisions.

Furthermore, as track geometry failure involves both degradation and shock failures, which are two dominant and competing causes of failure, proper track geometry maintenance should address these two failure modes jointly. Only a few researchers have addressed this issue in the literature. In order to fill this gap, a competing failure mode approach is also an interesting subject to study in order to develop a more efficient track geometry maintenance model.

In practice, track geometry maintenance planning needs to consider many other technical and organisational factors such as demand on track time and availability of maintenance resources. Therefore, introducing an interval limit or a linguistic variable for the track degradation threshold (instead of a crisp value) constitutes another attractive field of study for maintenance planning.

One of the most important issues in track maintenance scheduling is opportunistic maintenance planning. Just a few researchers have considered opportunistic maintenance approaches in terms of time, space, and components simultaneously in their work.

The other important issue in track maintenance scheduling optimisation is the planning horizon for a maintenance model. As discussed previously, considering tactical planning would bring more benefits regarding opportunistic maintenance compared to annual planning. Therefore, in future studies a tactical planning horizon should be considered for maintenance scheduling optimisation models. Based on the literature review, another important point for maintenance scheduling is to consider the proper stochastic degradation and restoration models in the track geometry maintenance scheduling.

Acknowledgement
The authors wish to thank Dr Nat Jack for his time and comments to improve the paper.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research received financial support from Trafikverket and the Luleå Railway Research Center.

References


Evaluation of the effect of tamping on the track geometry condition: A case study

Evaluation of the effect of tamping on the track geometry condition: A case study

Iman Soleimanmeigouni1, Alireza Ahmadi1, Iman Arasteh Khoy1 and Christophe Letot2

Abstract
Tamping is one of the major activities undertaken by railway maintenance managers to recover the track geometry condition. Modelling the effectiveness of tamping along with track geometry degradation is essential for long-term prediction of track geometry behaviour. The aim of this study is to analyse the effect of tamping on the different track geometry measurements, i.e. longitudinal level, alignment and cant, based on inspection car records from a part of the Main Western Line in Sweden. To model recovery after tamping, a probabilistic approach is applied. The track geometry condition before tamping was considered as the dominant factor for modelling the model parameters. Correlation analysis was performed to measure the linear relation between the recoveries of the different geometry measures. The results show a moderate correlation between the recovery of the longitudinal level and that of the cant, and a weak correlation between the recovery of the longitudinal level and that of the alignment. Linear regression and Wiener process were also applied to model track geometry degradation and to obtain degradation rates. The effect of tamping on degradation rate was analysed. It was observed that degradation rate increased after tamping.

Keywords
Correlation analysis, geometry measures, probabilistic model, railway maintenance, tamping, tamping effectiveness, track degradation, track geometry, Weiner process

Introduction
The railway track degrades through aging and usage, thus losing its functionality over time and resulting in failure. Effective maintenance can be employed to compensate for the shortcomings of railway track functionality and reliability. The main aim of maintenance development is to facilitate easy and cost-effective maintenance strategies so that the desired functionality and reliability of the track can be guaranteed. Models are needed to make optimal decisions with regard to reliability and maintenance during the design, development and operation stages.

Track geometry maintenance activities comprise a great part of the railway maintenance cost. Activities that may be applied in track geometry maintenance are manual intervention, tamping and stone-blowing.1 Tamping is one of the major activities undertaken by infrastructure maintenance managers to remedy track geometry failures. The tamping machine arms squeeze the ballast under the sleepers to provide a proper geometrical condition in the vertical and horizontal directions.

In order to evaluate the effect of tamping on the track geometry condition, different measures, i.e. longitudinal level, alignment, cant and twist, must be considered in the analysis. Although tamping will improve the track geometry condition, it cannot restore the geometry condition to as good as the new state. In fact, tamping is an imperfect maintenance action and will cause two changes in the track geometry condition, i.e. jump reduction in the track geometry measurements and a change in the degradation rate after tamping (see Figure 1). In the literature, it is mostly the jump reduction that is considered in the analyses of the effect of tamping on the track geometry condition. However, analysing the effect of tamping on the degradation rate is also of crucial importance.

1Division of Operation and Maintenance, Luleå University of Technology, Luleå, Sweden
2Machine Design and Production Engineering Lab, Research Institute for the Science and Management of Risks, University of Mons, Mons, Belgium

Corresponding author:
Iman Soleimanmeigouni, Division of Operation and Maintenance, Luleå University of Technology, SE-971 87 Luleå, Sweden.
Email: iman.meigouni@ltu.se
Modelling the recovery value after tamping along with the track geometry degradation is essential for long-term prediction of track geometry behaviour.

The recovery value after tamping can be modelled using deterministic or probabilistic approaches. The selection of the approach should be made according to the level of uncertainty in the recovery values after tamping. In the deterministic approach, the recovery value after tamping is directly modelled as a function of the influencing variables, such as the track geometry condition before tamping, speed, maintenance history, etc. In this approach, the model parameters are considered to be unknown but deterministic (constant), and the uncertainty is captured through a confidence interval. The deterministic approach is normally used when the model uncertainty is low.

In the probabilistic approach, the uncertainty is captured through modelling the recovery value after tamping by the selection of a proper distribution, while the parameters of the distribution are a function of the influencing variables. In this approach, for every set of input variables, a unique distribution is obtained for the recovery value after tamping. The probabilistic approach is normally used when the model uncertainty is high.

A common deterministic approach for modelling the recovery after tamping is the use of linear regression. In this approach, it is assumed that the recovery after tamping is linearly dependent on the track geometry condition before tamping. If the recovery values after tamping follow a normal distribution, a common linear regression model can be applied for modelling the recovery after tamping. Vale et al. selected a linear model to represent the relation between the recovery value after tamping and the track quality condition before tamping. This recovery value was combined with the degradation model to show the track quality condition for a long time period. An optimization model for minimizing the track maintenance cost was developed by Meier-Hirmer et al., who used linear regression to model the recovery after tamping. In addition, several researchers have also used linear regression to model the recovery after tamping for scheduling track maintenance activities.

Several other researchers have applied the probabilistic approach for modelling the recovery after tamping. The effect of tamping on the track geometry condition was studied by Audley and Andrews. They found that a two-parameter Weibull distribution could be applied to estimate the time for degradation to a specific level. They also concluded that there was a significant dependency between the Weibull parameters and the number of previous tamping interventions. In addition, they observed that the speed, the maintenance history and the level of track quality had a profound impact on the time required to achieve a specific level of the track geometry condition. With high-speed traffic and after a high number of tamping interventions, the track was found to degrade.
In this study, Quiroga and Schnieder\textsuperscript{10} used the Monte Carlo simulation to simulate track geometry degradation using tamping interventions. They assumed that the recovery value after tamping was a log normally distributed random variable and applied an exponential function to model the track geometry degradation in a maintenance cycle. They observed that by increasing the number of tamping interventions the variances of the recovery values after tamping and the degradation rate increased.

Soleimanmeigouni et al.\textsuperscript{11} developed a long-term prediction model combining degradation, shock event and tamping recovery models. It simulates the Wiener process to model track geometry degradation, simulates shock event times using an exponential distribution, and uses a probabilistic approach to model recovery after tamping.

Arasteh Khouy et al.\textsuperscript{12} also evaluated the effectiveness of tamping and compared the longitudinal levels before and after tamping. By applying the UIC tamping intervention graph,\textsuperscript{13} the recovery values of the longitudinal level are clustered into ‘bad’, ‘good’ and ‘excellent’ categories. They observed a significant variability in the tamping efficiency along the track length. In another study, Arasteh Khouy et al.\textsuperscript{14} investigated the optimal tamping intervention limits to minimize the total maintenance cost. They found that tamping at higher intervention limits was not cost effective in terms of capacity loss due to regulatory speed reduction, ride comfort and energy consumption. For more studies regarding the track geometry degradation and restoration models as well as the advantages and disadvantages of the models, readers are referred to Soleimanmeigouni et al.\textsuperscript{15}

In this study, a probabilistic model has been proposed to estimate the recovery value after tamping. It should be noted that tamping may be conducted on a track section either in partial or complete types. Since complete and partial tamping interventions have different effects on the track geometry condition, this issue is considered in the evaluation of tamping effectiveness in this study. Furthermore, the effect of tamping on different track geometry measurements, i.e. the longitudinal level, alignment and cant, has been analysed. In order to find out how tamping affects the different geometry measures, a correlation analysis has been conducted. In addition, the effect of tamping on the geometry degradation rate has been studied. In this regard, a linear regression model and the Wiener process have been used to model track geometry degradation.

The rest of this paper is organized as follows. The information of the line used as the case study is provided in following section. Then, the data collection and cleaning processes are described. The tamping practices in Line 414 within the Main Western Line in Sweden are then discussed and the classification of tamping interventions is demonstrated. A model is proposed to estimate the recovery value after tamping. ‘Correlation analysis of recovery values of geometry measures’ section discusses the results of the correlation analysis. ‘Effect of tamping on the track geometry degradation rate’ section is dedicated to the study of the effect of tamping on the degradation rate. Finally, the conclusions are provided.

**Line information**

In this study, the line section 414 between Järna and Katrineholm Central Station was used as a case study. This line section is a part of the Main Western Line in Sweden (Västra Stambanan). It is a double-track, electrified and remotely blocked line, used by both passenger and freight trains. The maximum speed of trains on the Main Western Line is around 200 km/h. Line 414 is 82 km long, and the line consists of UIC 60 and SJ 50 rails, M1 ballast material, Pandrol e-clip fasteners and concrete sleepers. The data on the line used in the case study were collected from 2007 to 2015 using Optram. Optram is the system that has been used since 2007 by Banverket (the former Swedish Rail Administration) and Trafikverket (the Swedish Transport Administration) for studying the measurements performed on track and overhead lines. This system visualizes and graphically represents the track geometry measurements. Optram provides functionality for analysis and displays data trends.\textsuperscript{16}

**Data collection and cleaning**

Due to the effect of different environmental and structural factors on the track geometry condition, various geometry degradation behaviours can be observed along the track length. In order to acquire a better understanding of the track geometry condition over the track length, line section 414 was divided into 200 m track sections in this study, and track geometry measurement data and track maintenance history data were used.

In order to remove corrupt and inaccurate records from the data set, data cleaning was performed. After an initial analysis of the track geometry measurement data and through consultation with railway track maintenance experts, it was found that there were some issues to be considered in the data cleaning process. First, there is a slight difference between the measurements obtained by different types of measurement wagons. In fact, using the measurement data gathered by different measurement wagons might have resulted in significant errors in the estimation of the track geometry degradation rate and the recovery values after tamping. Therefore, to avoid this kind
of error, only the measurement data gathered by measurement wagons STRIX and IMV200 were used in our study. The second point concerning data cleaning is the importance of the measurement wagon speeds. Based on the opinion of railway experts, a measurement wagon running at a speed lower than 40 km/h may not accurately measure the changes in the track geometry parameters. After analysing the data, we found that this particular issue can be easily observed in the track geometry measurement data. Thus, the measurement data obtained by measurement wagons running at low speeds (below 40 km/h) were removed from the data set. It should be noted that the inspection interval lengths used for the measurement wagons STRIX and IMV200 are not equal. In addition, in some cases, the time period between two inspections is significantly long. In these cases, the recovery values after tamping cannot be represented properly by the measurement data obtained with a long inspection interval. In order to obtain an accurate estimation of the recovery after tamping, it is necessary to use the measurement data in a short interval after the tamping interventions. Therefore, in this study, we did not consider these data for modelling the recovery values after tamping. Moreover, it must be noted that the degradation behaviour of the track sections containing points and level crossings is different from that of normal track sections. In addition, different types of maintenance activities are performed on the former track sections to remedy bad track geometry conditions. Accordingly, in this study, the track sections with points and level crossings were excluded from the analysis.

**Tamping practices in the Main Western Line**

Nowadays, tamping machines are equipped with special devices to perform tamping as well as levelling and lining actions simultaneously. Therefore, in addition to the vertical geometry defects, the horizontal track geometry defects are maintained using tamping interventions. Tamping tines pack the ballast material under the sleeper to provide a stable sleeper bed. The tamping tines penetrate the ballast and compact it under the sleeper through a squeezing movement. Tamping machines can be classified into different categories based on their speed, which is mainly influenced by the number of sleepers that can be tamped simultaneously. In another classification, tamping machines can be divided into continuous action tamping machines and intermittent action tamping machines. Intermittent action tamping machines tamp a group of sleepers, moving along for a couple of meters and then stopping to perform tamping on the next group of sleepers. Continuous action tamping machines are equipped with a tamping trolley under the machine itself, so that the trolley can move while the tamping machine is at a constant speed. There are also special tamping machines for switches and crossings. For point failures, tamping can also be performed by tractors and lightweight tamping machines. The machines used in Line 414 are single and double sleeper tampers for straight lines and a special tamping machine with flexible units for turnouts.

Immediately after the tamping intervention, some of the ballast particles do not lie in a consolidated and stable position. The stabilization process can be done in two different ways. First, the ballast is settled over time under traffic. But temporary speed restrictions must be imposed during the period of the initial settlement. Second, track stabilization machines are used to settle the ballast bed to avoid speed restrictions and irregular settlements. In track Line 414 for point failures, there is no demand for stabilization. Tamping on Line 414 is preferably done with mechanized stabilization. However, depending on the availability of machines and planning time, sometimes the track is stabilized by the traffic load. In this case, the traffic speed is reduced immediately after traffic until the track is stabilized after a few hundred thousand gross tons.

In the Line 414, no specific tamping interval is defined. In fact, the infrastructure managers follow a corrective maintenance strategy and try to perform tamping based on the track condition which can be checked by Optram. However, based on the data in Line 414, it can be observed that the time between two major tamping interventions is around five years. However, in addition to major tamping interventions, a number of track sections are tamped at different times.

**Complete and partial tamping classification**

It is a commonly accepted truth that the recovery after tamping is strongly dependent on the track geometry condition before the tamping intervention. In this study, we defined the recovery value after tamping as the difference between the standard deviation of the track geometry measurements before the tamping intervention and the corresponding standard deviation after the intervention. In Figure 2, the recovery values after tamping of the longitudinal level are plotted against the standard deviation of the longitudinal level (SDL) before tamping for the Line 414.

As is obvious in Figure 2, there is a high variation in the recovery values after tamping. In this connection, one should consider the fact that a tamping intervention may be performed on the whole length of a track section or on just a fraction of a section. Therefore, in this paper, those tamping interventions that were performed on a fraction of a track section are called partial tamping. Partial tamping can be categorized into four different types (see Figure 3). In the first type of partial tamping, a fraction of a track section is tamped, for example, because of a
point failure in that fraction. (When a point failure occurs, a fraction of a track section or, in a few cases, the whole length of a track section is tamped.) In the second type of partial tamping, the tamping intervention starts somewhere in one fraction of a section and then continues for a number of other sections. In the third type, the tamping is terminated in a fraction of a track section. The fourth type is a combination of the second and third ones. The main reason for partial tamping is point failures.

Figure 2. Recovery values after tamping versus the condition before tamping of the SDL for Line 414. SDL: standard deviation of the longitudinal level.

Figure 3. Schematic view of different partial tamping types.
tamping plotted against the track geometry condition before the tamping intervention. As shown in Figure 4, the expected recovery values after tamping for partial tamping are lower than those for complete tamping. The effectiveness of partial tamping should be analysed based on three main factors, i.e. the type of partial tamping used, the length of the tamped fraction and the location of the track geometry failure. It has been observed that the first type of partial tamping is more effective than the other types. This is due to the fact that this type of tamping is performed mainly to remedy point failures.

Another important issue concerns the track sections with a bad geometry condition. A very small number of track sections have an SDL bigger than 2. Since there were not enough data in that range, those sections were considered as bad-condition track sections and removed from the process of modelling the recovery after tamping.

The proposed model to estimate the recovery value after tamping

In this study, it was assumed that the recovery values after tamping of different track geometry measurements, i.e. longitudinal level, alignment and cant, were random variables with specific distributions. Therefore, a proper distribution that could be fitted on the recovery values was to be selected. After selecting the distribution, the parameters of the distribution should be considered as a function of influencing factors such as the track geometry condition before tamping, the speed and the maintenance history. However, in this study, the track geometry condition before tamping was considered as the dominant factor for modelling the distribution parameters. Since the entire track sections were used by trains running at the same speed, the factor of speed was not considered in the proposed model. According to the data collected during the eight years, there are a small number of track sections with more than one complete tamping. Therefore, the number of previous tamping interventions was not considered in the study. In order to obtain a complete understanding of the recovery after tamping, we needed to ascertain the effect of tamping on different geometry measures. In this regard, the recovery values after tamping for SDL, standard deviation of cant (SDC) and standard deviation of alignment (SDA) were investigated as follows.

Recovery values after tamping for the longitudinal level

In order to find the best-fit distribution on the recovery values after tamping, the Anderson–Darling and Kolmogorov–Smirnov tests were used as the goodness-of-fit criteria. The null hypothesis...
was the best choice for representing the data. Therefore, the three-parameter Weibull distribution has the smallest Kolmogorov–Smirnov values. In addition, this distribution has a P-value greater than 0.05 (the significant level). In addition, the Anderson–Darling test, for most of the three-parameter distributions, there is no established method for calculating the P-value. Therefore, both Anderson–Darling and Kolmogorov–Smirnov tests are used in this study in order to select the best-fit distribution on data and check the results. As can be seen in Table 1, the three-parameter Weibull distribution has the smallest Anderson–Darling value and has a P-value greater than 0.05 (the significant level). In addition, this distribution and the two-parameter Weibull distribution have the smallest Kolmogorov–Smirnov values. Therefore, the three-parameter Weibull distribution was the best choice for representing the data. For the fitted distribution, the estimated location, shape and scale parameters are $-\gamma = 0.029$, $\beta = 2.566$ and $\alpha = 0.603$, respectively. It must be noted that the location parameter in this case is very close to zero. However, based on the results of the Anderson–Darling test, the three-parameter Weibull distribution is used to ensure accuracy of the data considering the sensitivity of the model to the location factor. By doing so, the same model can be used to estimate the recovery values after partial tamping that contain negative values. In this study, we considered that the recovery after tamping followed the three-parameter Weibull distribution and the parameters of the distribution were a function of the track geometry condition before the tamping intervention.

Figure 5 shows that when the SDL before tamping increases, the recovery value after tamping and the related variation increase. Accordingly, it can reasonably be assumed that by increasing the SDL before tamping, the mean and variance of the three-parameter Weibull distribution should be increased. The three-parameter Weibull distribution for the recovery value of the longitudinal level has the following probability density function:

$$f(R_{SDL}) = \frac{\beta}{\alpha} \left( \frac{R_{SDL} - \gamma}{\alpha} \right)^{\beta - 1} e^{-\left( \frac{R_{SDL} - \gamma}{\alpha} \right)^{\beta}}$$  

(1)

where $R_{SDL}$ is the recovery value for the SDL, $\beta$ is the shape parameter, $\alpha$ is the scale parameter and $\gamma$ is the location parameter. The location parameter will shift the distribution. It should be noted that both the shape and the scale parameters are definite positive values. The mean and variance of the distribution can be represented as follows:

$$Mean(R_{SDL}) = \gamma + \alpha \Gamma \left( 1 + \frac{1}{\beta} \right)$$  

(2)

$$Var(R_{SDL}) = \alpha^2 \left\{ \Gamma \left( 1 + \frac{2}{\beta} \right) - \left( \Gamma \left( 1 + \frac{1}{\beta} \right) \right)^2 \right\}$$  

(3)

where $\Gamma$ represents the gamma function. Since the mean and variance of the recovery of the longitudinal level are directly proportional to the scale and location parameters and are inversely proportional to the shape parameter, it is assumed that the scale, shape and location parameters are a function of the SDL before tamping:

$$\alpha = a \theta_{SDL} + b$$  

(4)

$$\beta = c \theta_{SDL} + d$$  

(5)

$$\gamma = e \theta_{SDL} + f$$  

(6)

where $\theta_{SDL}$ is the SDL before tamping and $a, b, c, d, e$ and $f$ are the model coefficients. In order to estimate the values of these coefficients, the method of maximum likelihood estimation (MLE) was applied. The estimated parameters were $\hat{a} = 0.896$, $\hat{b} = 0.091$, $\hat{c} = 0$, $\hat{d} = 0.088$, $\hat{e} = -0.28$, $\hat{f} = -0.22$. As can be seen, since $\hat{c} = 0$, we can state that the shape parameter is not dependent on the SDL before tamping and has a constant value. Figure 5 shows a comparison between the simulated recovery values after tamping for the SDL using the three-parameter Weibull model and the real recovery values after tamping.

As is evident, the simulated recovery values after tamping obtained by the proposed model are very close to the real recovery values after tamping for the SDL.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Anderson–Darling</th>
<th>P-value</th>
<th>Kolmogorov–Simonov</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.414</td>
<td>0.332</td>
<td>0.079</td>
<td>0.139</td>
</tr>
<tr>
<td>3-Parameter Gamma</td>
<td>0.308</td>
<td>*</td>
<td>0.078</td>
<td>0.148</td>
</tr>
<tr>
<td>2-Parameter Weibull</td>
<td>0.245</td>
<td>&gt;0.250</td>
<td>0.073</td>
<td>0.198</td>
</tr>
<tr>
<td>2-Parameter Gamma</td>
<td>1.413</td>
<td>&lt;0.005</td>
<td>0.100</td>
<td>0.026</td>
</tr>
<tr>
<td>Logistic</td>
<td>0.689</td>
<td>0.043</td>
<td>0.088</td>
<td>0.069</td>
</tr>
<tr>
<td>3-Parameter Weibull</td>
<td>0.177</td>
<td>&gt;0.500</td>
<td>0.073</td>
<td>0.197</td>
</tr>
</tbody>
</table>

SDL: standard deviation of the longitudinal level.

* A closed form expression for p-value does not exist.

Table 1. Results for the fitted distribution to recovery values after tamping for SDL.
Recovery values after tamping for the cant

The same approach was applied for modelling the recovery of the cant. As can be seen in Table 2, the three-parameter lognormal distribution has the smallest Anderson–Darling and Kolmogorov–Smirnov values. For the fitted distribution, the estimated location, scale and threshold parameters are $\mu = -0.750$, $\sigma = 0.211$ and $\tau = -0.326$, respectively.

The three-parameter lognormal distribution for the recovery value of the cant has the following probability density function:

$$f(R_{SDC}) = \frac{1}{(R_{SDC} - \tau)\sigma\sqrt{2\pi}} \times \exp\left\{ -\frac{(\ln(R_{SDC} - \tau) - \mu)^2}{2\sigma^2} \right\}$$

where $R_{SDC}$ is the recovery value of the cant, $\mu$ is the location parameter, $\sigma$ is the scale parameter and $\tau$ is the threshold parameter. The mean and variance of the distribution are as follows:

$$\text{Mean}(R_{SDC}) = \tau + \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

$$\text{Var}(R_{SDC}) = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$$

Figure 5 shows that the mean and variance of the recovery value of the cant are directly proportional to the distribution parameters. Therefore, it is assumed that the location, shape and threshold parameters are a linear function of the SDC before the tamping intervention as shown in the following equations:

$$\mu = a'R_{SDC} + b'$$

**Table 2.** Results for the fitted distribution to recovery values after tamping for SDC.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Anderson–Darling</th>
<th>P-value</th>
<th>Kolmogorov–Smirnov</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>1.471</td>
<td>$&lt;0.005$</td>
<td>0.070</td>
<td>0.256</td>
</tr>
<tr>
<td>3-Parameter Lognormal</td>
<td>0.313</td>
<td>*</td>
<td>0.037</td>
<td>0.936</td>
</tr>
<tr>
<td>3-Parameter Weibull</td>
<td>0.964</td>
<td>0.012</td>
<td>0.056</td>
<td>0.538</td>
</tr>
<tr>
<td>3-Parameter Gamma</td>
<td>0.366</td>
<td>*</td>
<td>0.038</td>
<td>0.919</td>
</tr>
</tbody>
</table>

SDC: standard deviation of cant.

* A closed form expression for p-value does not exist.

**Recovery values after tamping for the cant**

The same approach was applied for modelling the recovery of the cant. As can be seen in Table 2, the three-parameter lognormal distribution has the smallest Anderson–Darling and Kolmogorov–Smirnov values. For the fitted distribution, the estimated location, scale and threshold parameters are $\mu = -0.750$, $\sigma = 0.211$ and $\tau = -0.326$, respectively.

Figure 6 shows that the mean and variance of the recovery value of the cant are directly proportional to the distribution parameters. Therefore, it is assumed that the location, shape and threshold parameters are a linear function of the SDC before the tamping intervention as shown in the following equations:

$$\mu = a'o_{SDC} + b'$$

SDC: standard deviation of cant.

* A closed form expression for p-value does not exist.
\[ \sigma = c' \theta_{SDC} + d \]  
\[ \tau = e' \theta_{SDC} + f' \]

where \( \theta_{SDC} \) is the standard deviation of the cant before tamping and \( a', b', c', d', e' \) and \( f' \) are the model coefficients. The unknown parameters, i.e. \( a', b', c', d', e' \) and \( f' \), were estimated using MLE and the estimated values are 0.063, 2.224, 0.015, 0, -0.0003 and -9.36, respectively. A comparison between the simulation results for the recovery values after tamping of the cant using the proposed approach and the real recovery values of the cant is shown in Figure 6. It is obvious that the variation in recovery values corresponding to bad geometry condition before tamping is higher than recovery values corresponding to good geometry condition before tamping. In addition, it can be seen that the mean value of the model is increased linearly by increasing the SDL before tamping.

**Recovery values after tamping for the alignment**

As Figure 7 shows, one can observe that in a number of track sections, tamping has a negligible positive effect and sometimes even has a negative effect on the alignment. Therefore, the pair-wise comparison test was applied to find whether tamping has a significant effect on the alignment or not. The results of the test show that the standard deviation of the alignment is smaller after tamping than before tamping. The 95% confidence interval for the recovery value of the alignment is (0.07, 0.10). Hence, it can be concluded that, although tamping will affect the alignment, the recovery of the alignment is small compared to that of the longitudinal level and the cant. The results of the goodness-of-fit test show that the three-parameter lognormal distribution has the smallest Anderson–Darling and Kolmogorov–Smirnov values and is properly fitted on the recovery values after tamping of the alignment (see Table 3). For the fitted distribution, the estimated location, scale and threshold parameters are \( \mu = -0.506, \sigma = 0.176 \) and \( \tau = -0.525 \), respectively.

The three-parameter lognormal distribution for the recovery values of the alignment has the following probability density function:

\[ f(R_{SDA}) = \frac{1}{(R_{SDA} - \tau)\sigma\sqrt{2\pi}} \times \exp \left\{ -\frac{(\ln(R_{SDA} - \tau) - \mu)^2}{2\sigma^2} \right\} \]

where \( R_{SDA} \) is the recovery value of the alignment. As in the case of the cant, it is assumed that the distribution parameters are a linear function of the SDA.
before the tamping intervention as shown in the following equations:

\[
\mu = a''\theta_{SDA} + b'' \\
\sigma = c''\theta_{SDA} + d'' \\
\tau = e''\theta_{SDA} + f''
\]

where \(\theta_{SDC}\) is the standard deviation of the cant before tamping and \(a'', b'', c'', d'', e'' \) and \(f''\) are the model coefficients. These model coefficients were estimated using MLE and the estimated values are 0.222, 1.115, 0.035, 0, −0.457 and −3.152, respectively. Figure 7 shows a comparison between the real recovery values for the alignment and the simulated recovery values using the proposed approach.

**Correlation analysis of recovery values of geometry measures**

Correlation analysis was used to measure the linear relation between the recoveries of the longitudinal level, alignment and cant, and the results are presented in this section. In this regard, the strength and direction of the linear relation between the recoveries of the track geometry measurements were evaluated by applying the Pearson correlation coefficient. The range for this value is from −1, which represents a perfect negative linear relationship, to +1, which represents a perfect positive linear relationship.

The results of the correlation analysis are presented in Table 4.

It can be seen that there is a moderate correlation between the recovery of the longitudinal level and that...
of the cant, a weak correlation between the recovery of the longitudinal level and that of the alignment, and finally, a very weak correlation between the recovery of the cant and that of the alignment. Accordingly, it can be inferred that tamping will not affect different track geometry measurements in a similar manner. For example, a tamping intervention may have a significant influence on the longitudinal level but a negligible effect on the alignment. In fact, although tamping machines perform levelling, lining and tamping at the same time, the effect of the tamping intervention on different geometry measurements is not equal.

It should be noted that alignment is a horizontal geometry measurement, while longitudinal level and cant are vertical measurements. This fact may support the results of the correlation analysis that show a higher correlation between the longitudinal level and cant when compared with the correlation between the longitudinal level and alignment. In addition, another reason for the weak correlation between recovery after tamping of the longitudinal level and alignment is the difference between the standard deviation of the longitudinal level and SDA before the tamping intervention. In fact, as the track geometry condition before tamping has a significant influence on recovery value after tamping, the difference between the standard deviation of the longitudinal level and SDA before tamping for a track section may cause the difference between the effect of tamping on the longitudinal level and alignment.

**Effect of tamping on the track geometry degradation rate**

In order to obtain the degradation rates, it is essential to employ an appropriate approach for modelling the track geometry degradation. The two main approaches for modelling the track geometry degradation are the stochastic and the deterministic approaches. In this study, both the approaches were used to obtain the degradation rates and to evaluate the effect of tamping on the degradation rate. The stochastic method applied was the Wiener process, while the deterministic method used was linear regression modelling.

The Wiener process has been widely applied for degradation modelling in different fields, e.g. bearings, laser generators and milling machines.\(^{18}\) In addition, this process has been used to model the performance deterioration of pavements.\(^{19}\) A stationary Wiener process is particularly well suited to model the evolution of a degradation mechanism characterized by a linear increase over time with random noise. Such a process has continuous sample paths and has independent, stationary and normally distributed increments.\(^{20}\) A Wiener process-based model is characterized by a drift parameter representing the expected degradation rate and by a Brownian motion term to consider the non-monotonic behaviour. Letot et al.\(^{21}\) used the Wiener process to model track geometry degradation. It should be noted that, in contrast to the gamma process, the Wiener process has both negative and positive increments. In this study, it was assumed that the negative increments could be due to noises or unregistered maintenance activity (not only tamping).

According to Kahle and Lehmann,\(^{22}\) the formula for the Wiener process can be expressed as follows:

$$Z(t) = z_0 + \mu t + \sigma W(t)$$  \hspace{1cm} (17)

where \(Z(t)\) is the degradation measure described by the model, \(z_0\) is the initial degradation, \(\mu\) is the drift coefficient, \(\sigma\) is the diffusion coefficient and \(W(t)\) is the standard Brownian motion representing the stochastic dynamics of the degradation process.\(^{20}\) The Wiener process is associated with the concept of Brownian motion. Hence, it can be represented by the following formulation:

$$Z(t) = z_0 + N(\mu t, \sigma \sqrt{t})$$  \hspace{1cm} (18)

The expected value of the Wiener process with the drift coefficient \(\mu\) is \(E(Z(t)) = \mu t\) and the variance is \(\text{Var}(Z(t)) = \sigma^2 t\).

The drift coefficient and diffusion coefficient were estimated for all track sections of Line 414 and the maintenance cycle. The estimated drift coefficients represent the degradation rates. The obtained degradation rates before and after the tamping interventions were compared by applying pair-wise comparison. The null hypothesis of the t-test was that there was no difference between the drift coefficient before and after tamping. The test was conducted with respect to a 5% level of significance. The means of the drift coefficients before and after tamping for the different sections were 0.085 and 0.11, respectively. The \(P\)-value of the t-test is almost zero, which shows that the null hypothesis is rejected. Therefore, it can be inferred that tamping will increase the degradation rate. The 95% confidence interval for the increment of the degradation rate is (0.013, 0.034).

The other approach used in this study for modelling the track geometry degradation is linear regression. The regression slope in a maintenance cycle is considered as the degradation rate. The geometry degradation rates were estimated for all of the track sections and maintenance cycles. The rates before and

<table>
<thead>
<tr>
<th>Variables</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery of longitudinal level and cant</td>
<td>0.558</td>
</tr>
<tr>
<td>Recovery of longitudinal level and alignment</td>
<td>0.267</td>
</tr>
<tr>
<td>Recovery of cant and alignment</td>
<td>0.100</td>
</tr>
</tbody>
</table>

**Table 4. The results of the correlation analysis.**
geometry measurements should be evaluated to understand the effect of tamping interventions on different track geometry behaviour using the nature of the process. Since complete and partial tamping interventions have a different effect on the track geometry condition, they should be clustered and analysed separately. In this case study, it has been observed that complete tamping has a significantly higher effect on the track geometry condition than partial tamping. Through the separation of complete and partial tamping, a significant reduction in the variation of the recovery values after tamping can be obtained. However, in this case study, quite a high level of uncertainty can still be observed in the recovery values after tamping. In this regard, in order to deal with this uncertainty, a probabilistic approach has been used. The three-parameter Weibull distribution was applied for modelling the effect of tamping on the SDL. In addition, the three-parameter lognormal distribution was used for modelling the effect of tamping on the standard deviations of the alignment and cant. The track geometry condition before tamping is considered as the dominant factor for modelling the recovery after tamping. It has been observed that the mean and variance of the distributions applied for recovery after tamping modelling are directly proportional to the track geometry condition before tamping. However, considering more influencing factors, such as the number of previous tamping interventions and the train speed, can increase accuracy of the model. It has been observed that the proposed approach can estimate properly the effect of tamping on different track geometry measurements. In fact, the proposed model can deal adequately with the uncertainty concerning the effect of tamping on different track geometry measurements that arises from various factors influencing the track geometry condition. It seems that the longitudinal level is the dominant factor for planning the tamping intervention in this case study. One possible reason is that the degradation rate of the longitudinal level is slightly higher than that of the cant and alignment. The results of the analysis in this case study show that tamping has the most significant effect on the longitudinal level. It has been observed that tamping has a significantly lower effect on the alignment than it has on the longitudinal level. The results of the correlation analysis show that there is a moderate correlation between the recovery after tamping of the longitudinal level and that of the cant. However, there is a weak correlation between the recovery after tamping of the longitudinal level and that of the alignment, and a very weak correlation between the recovery after tamping of the cant and that of the alignment. Therefore, one can infer that tamping will not affect different track geometry measurements equally. Thus, in order to obtain an accurate understanding of the effect of tamping on the track geometry condition, different track geometry measurements should be involved in the analysis. Finally, the track geometry degradation was modelled using linear regression and the Wiener process by

**Conclusions**

The effect of tamping interventions on different track geometry measurements should be evaluated to obtain a proper prediction of the long-term track geometry behaviour. Since complete and partial tamping interventions have a different effect on the track geometry condition, they should be clustered and analysed separately. In this case study, it has been observed that complete tamping has a significantly higher effect on the track geometry condition than partial tamping. Through the separation of complete and partial tamping, a significant reduction in the variation of the recovery values after tamping can be obtained. However, in this case study, quite a high level of uncertainty can still be observed in the recovery values after tamping. In this regard, in order to deal with this uncertainty, a probabilistic approach has been used. The three-parameter Weibull distribution was applied for modelling the effect of tamping on the SDL. In addition, the three-parameter lognormal distribution was used for modelling the effect of tamping on the standard deviations of the alignment and cant. The track geometry condition before tamping is considered as the dominant factor for modelling the recovery after tamping. It has been observed that the mean and variance of the distributions applied for recovery after tamping modelling are directly proportional to the track geometry condition before tamping. However, considering more influencing factors, such as the number of previous tamping interventions and the train speed, can increase accuracy of the model. It has been observed that the proposed approach can estimate properly the effect of tamping on different track geometry measurements. In fact, the proposed model can deal adequately with the uncertainty concerning the effect of tamping on different track geometry measurements that arises from various factors influencing the track geometry condition. It seems that the longitudinal level is the dominant factor for planning the tamping intervention in this case study. One possible reason is that the degradation rate of the longitudinal level is slightly higher than that of the cant and alignment. The results of the analysis in this case study show that tamping has the most significant effect on the longitudinal level. It has been observed that tamping has a significantly lower effect on the alignment than it has on the longitudinal level. The results of the correlation analysis show that there is a moderate correlation between the recovery after tamping of the longitudinal level and that of the cant. However, there is a weak correlation between the recovery after tamping of the longitudinal level and that of the alignment, and a very weak correlation between the recovery after tamping of the cant and that of the alignment. Therefore, one can infer that tamping will not affect different track geometry measurements equally. Thus, in order to obtain an accurate understanding of the effect of tamping on the track geometry condition, different track geometry measurements should be involved in the analysis. Finally, the track geometry degradation was modelled using linear regression and the Wiener process by

![Figure 8. Track geometry degradation modelling using a linear regression model and the Wiener process.](image)
considering the longitudinal level as the main measure for representing the track geometry condition. It was observed that tamping will increase the slope of the regression line and the drift coefficient of the Wiener process. Therefore, it can be inferred that tamping has a negative impact on the degradation rate. The proposed approach in this study can be extended for modelling recovery after partial tamping interventions by making small modifications.

Acknowledgements

The authors wish to thank Trafikverket and Luleå Railway Research Center (JVTC) for the technical support provided during this project. Special thanks is extended to Dr. Arne Nissen from Trafikverket for his valuable support and discussion.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study is part of the project ‘RAMS and LCC for Railway’, which is sponsored by both Trafikverket and Luleå Railway Research Center (JVTC).

References

Modelling the evolution of ballasted railway track geometry by a two-level piecewise model

Modelling the evolution of ballasted railway track geometry by a two-level piecewise model

Iman Soleimanmeigounia, Xun Xiaoα, Alireza Ahmadiα, Min Xieb, Arne Nissenc and Uday Kumarα

αDivision of Operation and Maintenance, Luleå University of Technology, Luleå, Sweden; βDepartment of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong; cTrafikverket, Luleå, Sweden

ABSTRACT

Accurate prediction and efficient simulation of the evolution of track geometry condition is a prerequisite for planning effective railway track maintenance. In this regard, the degradation and tamping effect should be equipped with proper and efficient probabilistic models. The possible correlation induced by the spatial structure also needs to be taken into account when modelling the track geometry degradation. To address these issues, a two-level piecewise linear model is proposed to model the degradation path. At the first level, the degradation characteristic of each track section is modelled by a piecewise linear model with known break points at the tamping times. At the second level, Autoregressive Moving Average models are used to capture the spatial dependences between the parameters of the regression lines indexed by their locations. To illustrate the model, a comprehensive case study is presented using data from the Main Western Line in Sweden.

1. Introduction

Track geometry degrades with age and tonnage, and can affect negatively track safety, availability and quality of ride. To ensure the highest possible track performance, maintenance should be planned to control degradation and to restore the track geometry to an operational state. To determine track condition, an inspection vehicle runs along the track to measure its vertical and lateral deviation at short (usually 25 cm) intervals. The measurements are processed to produce numerical characteristics representing the track geometry condition over a track section of a specific length (usually 200 m). The characteristics include, e.g. the mean, standard deviation and extreme values within specific wavelengths. The standard deviations of the longitudinal level and alignment may represent the condition of track geometry in vertical and lateral direction, respectively; and in term of distributed irregularities. In addition to standard deviation of geometry parameters, isolated defects also might be considered in the evaluation of track geometry condition. The isolated defects are short irregularities in track geometry (e.g. 1–3 m) that can dramatically increase the dynamic forces between wheel and rail. However, for most railway industries, the standard deviation of the short wavelength (3–25 m) longitudinal level defects is the decisive factor in planning maintenance activities (UIC, 2008). Accordingly, most researchers use this characteristic to represent the condition of the track geometry (Soleimanmeigouni, Ahmadi, & Kumar, 2016).

The track geometry degradation characteristics should be kept within specific limits. The European Standard EN 13848-5 (2008) defines three limits for maintenance. (1) Intermediate Action Limit (IAL): this is the safety limit. If the IAL is exceeded, there is a risk of derailment; the risk can be reduced by speed reduction, line closure, or corrective maintenance. Different indicators are mentioned in EN 14363 (2016) for evaluation of safety against derailment such as the ratio of guiding force and vertical wheel force on the outer wheel Y/Q. The readers are referred to EN 14363 (2016) for more information about the factors influencing the safety against derailment of vehicles. (2) Intervention Limit (IL): this is the corrective maintenance limit. If the IL is exceeded, corrective maintenance should be conducted before the IAL is reached. (3) Alert Limit (AL): this is the preventive maintenance limit. If the AL is exceeded, the track geometry condition should be analysed and considered in regularly planned maintenance operations. Tamping is the main maintenance action used to remedy the track geometry condition and keep it within the required limits. Nowadays, tamping machines are equipped with special devices that provide this opportunity to do tamping, levelling and lining actions, simultaneously. Therefore, in addition to vertical geometry defects, the horizontal track geometry defects will be maintained by tamping action. Tamping machine lifts the track to a specific level and the tamping tines pack the ballast under the sleepers. The tamping tines penetrate the ballast and compact the ballast under the sleeper with a squeezing movement to provide a stable sleeper bed.

To model the evolution of the geometry condition of a track line, two main challenges must be addressed: (1) modelling the temporal evolution of track geometry condition by considering...
the effect of maintenance; (2) modelling the spatial variation in degradation parameters, along with the possible spatial dependences. A number of important track geometry degradation models are reviewed by Sadeghi and Askarinejad (2010). However, the degradation patterns within one maintenance cycle are relatively simple. For the standard deviation of the longitudinal level, a simple linear model or an exponential model can be used, with time or accumulated traffic load as the explanatory variable (Caetano & Teixeira, 2014; Esveld, 2001; Lichtberger, 2005).

Nevertheless, the degradation path always has an obvious break point after a maintenance activity. Tamping usually causes two changes in the track geometry condition, i.e. a sudden reduction in the current track geometry level and a change in the degradation rate (Quiroga & Schneider, 2012). In other words, if it is assumed that the degradation path is linear within one maintenance cycle, the degradation paths before and after tamping will follow two lines with different slopes and intercepts. Although tamping will improve the track geometry condition, it cannot restore it to an as-good-as-new state. Generally speaking, the existing literature shows tamping tends to decrease the current degradation value significantly but, at the same time, to increase the degradation rate (Audley & Andrews, 2013; Soleimanmeigouni, Ahmadi, Arasteh Khouy, & Letot, 2016). The recovery value of tamping is defined as the difference between the degradation characteristics before and after tamping. In general, this value is modelled considering the track geometry condition right before tamping as the dominant factor (Miwa, 2002; Vale & Ribeiro, 2014). Nonetheless, there is a variation in recovery values even for sections with almost the same track geometry conditions before tamping. It is necessary to consider the effect of different tamping actions and the heterogeneities between track sections.

Analysing the effect of tamping on track geometry condition has been the main concern of a number of studies. Arasteh Khouy, Schunnesson, Juntti, Nissen, and Larsson-Kraik (2013) evaluated the effect of tamping and compared the longitudinal levels before and after tamping. By applying International Union of Railways (UIC) tamping intervention graph (UIC, 2008), they clustered the recovery values of the longitudinal level into ‘bad’, ‘good’ and ‘excellent’ categories and observed a significant variability in the effect of tamping along the track length. To evaluate the influence of tamping on track geometry, Audley and Andrews (2013) used the Weibull distribution to model the times for a degradation characteristic to reach a specific level and a log-normal distribution to determine the probability that the track would not achieve specific track geometry conditions after tamping. They found a significant trend between the degradation rate and the number of tampings. They also observed that the train speed, track maintenance history and level of track quality before tamping all have a profound impact on the time required to achieve a specific track quality.

Soleimanmeigouni et al. (2016) evaluated the effect of tamping on different track geometry measures. They found that tamping, i.e. complete or partial tamping, has a significant effect on the recovery values. Tamping intervention may be performed on the whole length of a track section or on just a fraction of a section. If a tamping intervention was performed on a fraction of a track section is called partial tamping, otherwise it is called complete tamping. In addition, they observed that, on average, tamping tends to increase the degradation rate.

Different models have been proposed to combine the degradation process within the maintenance cycle and the maintenance effect on track geometry condition to address the long-term temporal evolution of the track geometry condition. A survey of these studies can be found in a review paper by Soleimanmeigouni et al. (2016). Caetano and Teixeira (2016) applied a linear function of the accumulated axle load to model the degradation of two track geometry characteristics, i.e. longitudinal level and alignment. They introduced a stochastic variable obtained from an analysis of the accuracy of estimation to consider uncertainty in predicting track geometry condition. To model the recovery of tamping, they applied a linear regression model, considering the track geometry condition before tamping as the explanatory variable. They also considered the age of the track sections as a covariate.

Quiroga and Schneider (2012) applied an exponential function to model track geometry degradation in a maintenance cycle. They considered the degradation level after tamping and the degradation rate to be log-normally distributed random variables dependent on the number of tampings performed on a track section. They observed that the mean and variance of degradation rates increased with the number of performed tamping interventions. Famurewa, Juntti, Nissen, and Kumar (2016) applied an exponential function to model track geometry degradation in a maintenance cycle. They assumed the track geometry condition after the first tamping to be a linear function of track geometry condition before tamping and the track geometry condition after any further tamping to be dependent on the quality achieved in the last intervention, the cumulative number of tamping interventions and a quality loss factor.

In addition, several researchers have used stochastic processes to model track geometry degradation. Meier-Hirmer, Riboulet, Sourget, and Roussignol (2009) developed a maintenance model based on Gamma process with fixed inspection intervals and delays of interventions. They used a regression model with a Gaussian random error term to relate the recovery value after tamping intervention to the degradation level before tamping. Mercier, Meier-Hirmer, and Roussignol (2012) applied the bivariate gamma process to model track geometry degradation by considering the longitudinal level and transversal level as track geometry characteristics. A stochastic approach based on the Dagum distribution is developed by Vale and Lurdes (2013) who modelled track longitudinal level degradation over time.

Yousefikia, Moridpour, Setunge, and Mazloumi (2014) modelled tram track degradation using a Markov model to obtain the optimal maintenance strategy. In their view, since tram tracks include numerous tight curves, the effect of curves on track degradation in terms of real wear should be considered through degradation modelling. Prescott and Andrews (2013a) used a Markov model to predict the long-term track geometry condition. To cope with the limitations of using a Markov model in track geometry degradation modelling, some researchers have recently used Petri nets (Andrews, 2013; Prescott & Andrews, 2013b). Andrews, Prescott, and De Rozières (2014) applied a Petri net to predict the track geometry condition given different asset management strategies. They used the proposed model to evaluate the life cycle cost of the various strategies to find the most effective one.
Considering the spatial issues, the degradation patterns vary from section to section. Hence, it is natural to employ probabilistic models to quantify the uncertainties of their patterns. It is also shown that there is no significant effect on the degradation rate by the types of traffics or track constructions (Esved, 2001). Andrade and Teixeira (2011) considered a number of infrastructure features and applied a multivariate linear regression model to explain the section-to-section variation in initial degradation values and degradation rates. Most of the regression coefficients were not statistically significant and the model could not explain the section-to-section variations with respect to those considered explanatory variables. It must be noted that the neighbouring track sections often tend to exhibit a more similar degradation pattern than track sections with a significant distance between them. The reason of this behaviour may be that the neighbouring track sections have similar structural, environmental, and operational conditions.

Recently, researchers have developed spatial models for the track geometry. Andrade and Teixeira (2015) developed a Hierarchical Bayesian model to predict the degradation characteristics, i.e. the standard deviations of longitudinal level and alignment. They stated that track geometry degradation in a maintenance cycle has a linear behaviour. They also considered the spatial dependency of model parameters in consecutive track sections and applied Conditional Autoregressive (CAR) model to capture the possible spatial dependence. Andrade and Teixeira (2016) used the same approach and considered both planned and unplanned maintenance to evaluate different alert limit strategies.

This paper proposes a two-level piecewise linear model to characterize track geometry degradation and restoration with possible spatial dependences. It should be noted that the model is proposed for ballasted railway tracks. At the first level, the track geometry degradation of each section is modelled using a piecewise linear model with break points at the maintenance times. Then, a multivariable linear regression model is used to link different covariates on the maintenance actions and track section conditions with the responses, i.e. the recovery values and changes in degradation rates after a tamping action for a track section. At the second level, Autoregressive Moving Average (ARMA) models are used to capture the spatial dependences in the parameters of the regression lines indexed by their locations. The ARMA model is widely used to deal with auto-correlated data; the main advantage is its simplicity. Since the hierarchical model tends to be complex and difficult for the statistical estimation, a simple ad hoc procedure is performed to estimate all the unknown parameters.

The rest of this paper is organised as follows. Section 2 presents the two-level piecewise linear model. Section 3 explains the estimation procedure for the unknown parameters in the proposed model. Section 4 contains a case study on the degradation measurement data and maintenance history data for a track line in the Swedish railway network. Finally, Section 5 gives the main conclusions and suggests directions for further research.

2. Track geometry degradation model

The evolution of the track geometry degradation of a track line with length $L$ over a time horizon $[0, T]$ can be regarded as either a deterministic or a stochastic function. This function is defined over the two-dimensional spatiotemporal space $\mathbb{D} = [0, L] \times [0, T]$ and increases monotonically or stochastically. Here, time could be the chronological time or any monotonic transformation of the time, e.g. the accumulated axle load (Million Gross Tons, MGT) of a track section. The track geometry degradation data exhibit many uncertainties and the observed degradation path of a specific section usually presents an increasing trend but may decrease between two or three consecutive measurements. In this paper, the degradation path of the track geometry characteristic is assumed to be a stochastically increasing function $D(s, t)$ mapping the domain $\mathbb{D}$ to the range $\mathbb{E}$, i.e. a subset of the real line $\mathbb{R}$. Since most of the track geometry degradation characteristics are strictly positive (e.g. the standard deviation of the longitudinal level investigated in this paper), $\mathbb{E} = (0, +\infty)$ is widely used.

The original data can be transformed into log-scale by simply taking the logarithm. Then, the degradation model under the log-scale as $Y(s, t) = \ln D(s, t)$ is obtained. As the log transformation is monotone, it does not affect the time when a degradation characteristic hits the maintenance threshold. After the transformation, the range of $Y(s, t)$ will be $\mathbb{R}$. The log-scale will provide us many flexibilities and benefits in statistical modelling and analysis. In the rest of this paper, all issues are investigated under the log-scale, unless stated otherwise.

In practice, the measurements of the degradation characteristics are taken over a discretised time schedule and the values of the degradation characteristics are calculated for individual track sections with the same length $l$. Therefore, the degradation characteristics measured over an interval, e.g. the first section $[0, l]$, will be the same. Both the spatial space $[0, L]$ and the temporal space $[0, T]$ are discretised into $S = \{1, 2, \ldots, S\}$ and $T = \{t_0, t_1, \ldots, t_n\}$, where $S = L/l$ and $0 = t_0 < t_1 < \cdots < t_n = T$. Since $S$ becomes an index set, the degradation model over the discretised domain $S \times T$ can be rewritten as $Y_s(t)$, where $s \in S$ and $t \in T$.

2.1. Degradation model within a maintenance cycle

It is widely pointed out that different track sections have different degradation parameters (Andrade & Teixeira, 2015). The degradation characteristic of the $i^{th}$ section is assumed to follow a linear model with Gaussian error term as follows:

$$Y_i(t) = \alpha_i + e^\gamma t + \epsilon_i, t \in T,$$

where $\alpha_i$ is the initial degradation value, $e^\gamma$ is the degradation rate, and $\epsilon_i \sim N(0, \sigma_i^2)$. The Gaussian error term $\epsilon_i$ is included in the model to consider the uncertainty in the evolution of track geometry condition. The slope of this linear model is parameterised as $e^\gamma > 0$, where $\gamma_i \in \mathbb{R}$ to ensure that $Y_i(t)$ is stochastically increasing. Although the model is defined over the discretised time domain $T$, it is obviously trivial to generalise it to the continuous time domain $[0, T]$.

In fact, this model describes the evolution of a single degradation characteristic in one maintenance cycle, i.e. the time between two consecutive maintenance activities. Note that this linear model with an additive normal error term under the log-scale is equivalent to an exponential model with a multiplicative lognormal error term in the original scale. In a traditional setting,
it could be assumed that the evolution of a degradation characteristic follows a linear model or an exponential model with an additive normal error term (Quiroga & Schnieder, 2012). This traditional setting means it is possible to get a negative degradation value at any measurement time. Even though the probabilities are usually extremely small, it still makes the additive normal error term in the original scale an unsatisfactory assumption.

2.2. Degradation modelling in multiple maintenance cycles

When the measurements of the degradation characteristic identify that some track sections are reaching the maintenance threshold, tamping will be directly scheduled to restore the track geometry condition. It is assumed that there is no delay in maintenance once a track section hits the threshold. Suppose that for the $s^{th}$ section, there are $k_s$ maintenance actions right after the known times $t_{m_{1,s}}, \ldots, t_{m_{n,s}} \in T$, where $0 \leq k_s < n$, $m_{1,s} \geq 3$ and $m_{j+1,s} - m_{j,s} \geq 3$. The $m_{j,s}$ represents the measurement number at $j^{th}$ maintenance action for the $s^{th}$ track section. The constraints $m_{1,s} \geq 3$ and $m_{j+1,s} - m_{j,s} \geq 3$ ensure there are at least three measurements in each linear segment, and the regression coefficients are estimable.

The degradation path $Y(t)$ over the whole time horizon $T$ for the $s^{th}$ section is assumed to be a piecewise linear model as follows:

\[
Y_s(t) = \sum_{j=1}^{k_s} \left( \alpha_{s} - \sum_{j=1}^{k_j} r_{j,s} I(t < t_{m_{j,s}}) + e^{r_{j,s} (t_{m_{j,s}} - t)} \right) Y_{s,j} + \epsilon_s,
\]

where $r_{j,s}$ is the recovery effect on the degradation level of the $j^{th}$ tamping, $\omega_{j,s}$ is the effect on the degradation rate of the $j^{th}$ tamping, $Y_{s,j}$ is the natural logarithm of degradation rate, and $I(\cdot)$ is the indicator function. Here, $\tilde{Y}_s(t)$ represents the deterministic linear part of $Y_s(t)$. The piecewise linear function is also well-defined over the continuous time interval $[0, T]$. If the delay of the maintenance action is deterministic and denoted by $\Delta \in (t_{m_{1,s}}, t_{m_{n+1,s}})$, it is simple to incorporate this delay effect into Equation (2) by replacing $I(t > t_{m_{j,s}})$ with $I(t > t_{m_{j,s}} + \Delta)$

This model defines a piecewise linear function over time $t$. It contains $k_s$ change points and $k_s + 1$ pieces of linear segments. If $k_s$ is zero, it reduces to the simple model presented in the last subsection. The model can be written in a separate form as follows:

\[
Y_s(t) = \begin{cases} 
\alpha_s + \beta_s t + \epsilon_s, & 0 \leq t < t_{m_{1,s}}, \\
\alpha_s - r_{1,s} + e^{r_{1,s} (t_{m_{1,s}} - t)} + \epsilon_s, & t_{m_{1,s}} < t < t_{m_{2,s}}, \\
\ldots \\
\alpha_s - \sum_{j=1}^{k_s} r_{j,s} + e^{r_{j,s} (t_{m_{j,s}} - t)} + \epsilon_s, & t_{m_{j,s}} < t \leq T.
\end{cases}
\]

where $\beta_s > 0$ is the degradation rate. For the simplicity of the further discussions the log-rate function $\Gamma_s(t)$ is defined as:

\[
\Gamma_s(t) = \gamma_s + \sum_{j=1}^{k_s} \left[ \omega_{j,s} I(t > t_{m_{j,s}}) \right],
\]

which is a piecewise constant function. By ignoring the error term $\epsilon_s$, $\Gamma_s(t)$ is the logarithm of the first-order derivative of the deterministic linear part of $Y_s(t)$ and summarises all degradation rates over $T$.

For the long-term prediction of the track geometry condition, it is necessary to model $r_{j,s}$ and $\omega_{j,s}$ properly to take the future maintenance activities into account. Recovery value after tamping can be assumed to be a deterministic function, depending on certain covariates of the track sections (Vale, Ribeiro, & Caicada, 2011). For example, for a specific section, the degradation value before tamping is the critical covariate used to predict degradation level after tamping (Famurewa, Xin, Rantatalo, & Kumar, 2015). In addition, the degradation rate after tamping is highly related to the number of accumulated tamping interventions (Quiroga & Schnieder, 2012). However, there are still a great amount of uncertainties in both $r_{j,s}$ and $\omega_{j,s}$, which cannot be explained by these two covariates. For example, the tamping activity could be complete tamping for a whole section or partial tamping for only a fraction of a section. Complete tamping is expected to improve the degradation level more than partial tamping. As pointed out by Audley and Andrews (2013), the number of accumulated tamping and train speed also affects the tamping effectiveness.

For the $s^{th}$ section, the degradation value and the degradation rate right before the $j^{th}$ tamping are given as $Y_s(t_{m_{j,s}})$ and $\Gamma_s(t_{m_{j,s}})$. It is further assumed that the rest of the possible covariates are $\{x_{v,s,j}, v = 1, \ldots, U\}$ for the $j^{th}$ tamping of the $s^{th}$ section. Therefore, two linear regression models are suggested to link $r_{j,s}$ and $\omega_{j,s}$ with all the covariates as:

\[
r_{j,s} = a_1 + b_{10} Y_{s,j} + \sum_{u=1}^{u} b_{1u} x_{v,s,j} + \delta_1
\]

\[
\omega_{j,s} = a_2 + b_{20} \Gamma_{s,j} + \sum_{u=1}^{u} b_{2u} x_{v,s,j} + \delta_2
\]

where $a_1, a_2, b_{1v}, b_{2v}, v = 0, \ldots, u$ are regression coefficients, and $\delta_1 \sim N(0, \Gamma_{1}^2)$, $\delta_2 \sim N(0, \Gamma_{2}^2)$ are Gaussian random error terms. In addition, $b_{1v}$ and $b_{2v}$ are listed separately to emphasize their importance. More specifically, Model (5) and Model (6) are referred to as the recovery model and the change in rate model, respectively. Note that neither $r_{j,s}$ nor $\omega_{j,s}$ is observed, and both can be linked to $\tilde{Y}_s(t)$ and $\Gamma_s(t)$ as follows:

\[
r_{j,s} = \tilde{Y}_s(t_{m_{j,s}}) - \tilde{Y}_s(t_{m_{j,s}} + \Delta)
\]

\[
\omega_{j,s} = \Gamma_s(t_{m_{j,s}} + \Delta) - \Gamma_s(t_{m_{j,s}})
\]

where $\tilde{Y}_s(t)$ and $\Gamma_s(t)$ means the right limits of the functions at time $t$ defined over the continuous time domain $[0, T]$.

2.3. Spatial dependence in $\gamma$, $\sigma_s$, and $\sigma_v$

To predict the geometry condition of a specific track line, given sufficient historical data, the model presented in the previous subsection will suffice. To simulate the long-term evolution of track geometry for any track line, however, the parameters of the degradation model need to be simulated for all track sections.

To address the possible spatial dependency in track degradation, it is assumed that the degradation parameters are
independently generated by stationary ARMA processes. In particular, the initial value \( \alpha_s \), the log-rate \( \gamma_s \), and the log standard deviation \( \rho_s \). In \( \sigma \) follow three independent stationary ARMA processes over \( S \). It must be considered that \( \alpha_s \), \( \gamma_s \) and \( \rho_s \) are latent discrete stochastic processes. They are considered as discrete as they have different realised values for each track section. The ARMA \((p,q)\) model, wherein \( p \) is the order of autoregressive terms and \( q \) is the order of moving average terms, is given as follows:

\[
\begin{align*}
\alpha_s &= c_1 + \sum_{i=1}^p \varphi_{1,i} \alpha_{s-i} + \sum_{i=1}^q \theta_{1,i} e_{s+1-i} + e_{s,1}, \\
\gamma_s &= c_2 + \sum_{i=1}^p \varphi_{2,i} \gamma_{s-i} + \sum_{i=1}^q \theta_{2,i} e_{s+i} + e_{s,2}, \\
\rho_s &= c_3 + \sum_{i=1}^p \varphi_{3,i} \rho_{s-i} + \sum_{i=1}^q \theta_{3,i} e_{s+i} + e_{s,3},
\end{align*}
\]

where \( e_{s,1}, e_{s,2}, e_{s,3} \) are constant terms, \( \varphi_{1,i}, \varphi_{2,i}, \varphi_{3,i} \) are auto-regressive parameters, \( \theta_{1,i}, \theta_{2,i}, \theta_{3,i} \) are moving average parameters, and \( e_{s,1}, e_{s,2}, e_{s,3} \) are independent Gaussian innovations.

Herein, it is assumed that the stationary condition hold for the whole track. Although this assumption looks restrictive, it actually reflects the effect and consistency of the operation and management of the rail transportation department. At the very least, the stationary condition should hold locally. If the observed data show the underlying degradation parameters are instationary over the spatial space, the autoregressive moving average model, including any exogenous covariates (ARMAX), may be used to explain the instationary part by examining the relevant covariates. In fact, the ARMAX model can always be used if we want to incorporate exogenous covariates into the model. Certainly, \( \bar{Y}(t_{m,s}) \) or \( \Gamma_s(t_{m,s}) \) used in the maintenance effect models should not be included as covariates, as they are endogenous.

As stated in Gaetan, Guyon, and Bleakley (2010), there is no theoretical reason to limit the use of ARMA models in spatial statistics. Some might argue that the time order, i.e. the one-directional dependence, as explicitly presented in Model (9), is different from the spatial process because of its possible two-directional dependence. Note that we restrict Model (9) to spatial ARMA processes with Gaussian innovations because only this class of ARMA processes satisfies the time reversibility (Azrak & Mélard, 2006). With this property, the reverse process \( \{ \tilde{\alpha}_s, \tilde{\gamma}_s = \alpha_{s+i} \} \) follows the same ARMA process in Model (9) as the original process \( \{ \alpha_s \} \). The same argument holds for \( \{ \gamma_s \} \) and \( \{ \rho_s \} \). This shows that an ARMA process with time reversibility assigns a symmetric neighbourhood relationship to its realisations.

Furthermore, the CAR model originally proposed by Besag (1974) has been used by Andrade and Teixeira (2015) to capture spatial dependence within degradation parameters. The basic idea of CAR model is that the probability of a value in a location is conditional on the values of neighbouring locations. For the one-dimensional discrete spatial space, i.e. the regular lattice system over a line, the CAR model is equivalent to another class of the spatial AR model – the Simultaneous Autoregressive (SAR) model (see Gaetan et al., 2010). Clearly, the track line system defined in our paper is a one-dimensional regular lattice system.

Finally, an example presented in Gaetan et al. (2010) suggests an AR model can be transformed into a SAR model. In fact, it can be shown that any finite realisation of a stationary AR process over \( S \) can be characterised as the realisation of a CAR process over \( S \). The proof is presented in the Appendix 1.

2.4. Link between original scale and log scale

Although the focus in this paper is on the log scale, it is sometimes necessary to show the track geometry degradation process using the original scale. As discussed in Section 2.1, under the original scale, our model becomes an exponential model. Let \( D_s(t) = e^{\gamma_s(t)}, D_s(0) = e^{\gamma_s}, \gamma_s = e^{\gamma_s}, \) and \( \beta_s = e^{\beta_s}, \) the track geometry degradation within a maintenance cycle for the \( s \)th section can be modelled using an exponential model as follows:

\[
D_s(t) = D_{s,0} e^{\beta_s t} e_s, \ t \in T
\]

where \( D_s(t) \) is the degradation value at time \( t \in T \), \( D_{s,0} > 0 \) is the initial degradation level, \( \beta_s > 0 \) is the degradation rate and \( e_s \) is the random error term following a lognormal distribution with the location parameter zero and the scale parameter \( \sigma_s \).

Furthermore, the degradation model under multiple tamping interventions can be modelled as follows:

\[
D_s(t) = \frac{D_{s,0}}{\Pi_{j=1}^k e^{\gamma_s(t_{m-s})}} \exp \left( \beta_s \sum_{j=1}^k e^{\gamma_s(t_{m-s})} \right) e_s = \tilde{D}_s(t) e_s
\]

The above are simple transformations of the model under two different scales. Now, the change ratio \( R \) of a maintenance action taken at time \( t_{m,s} \) for the \( s \)th section is defined as follows:

\[
R_{s} = e^{\gamma_s} = \frac{\tilde{D}_s(t_{m,s})}{\tilde{D}_s(t_{m,s} +)}.
\]

It is easy to check that \( R \in (0, +\infty) \). A bigger recovery ratio indicates a more effective maintenance action; hence, there is a smaller degradation value right after this maintenance action. The recovery value under the log-scale corresponds to the recovery ratio under the original scale. Compared to the recovery value under the original scale, \( \tilde{D}_s(t_{m,s}) - \tilde{D}_s(t_{m,s} +) \), the recovery ratio under the original scale, i.e. the recovery value under the log scale, is a more natural criterion to interpret the effects of tamping, as a linear regression model with the response \( \tilde{D}_s(t_{m,s}) - \tilde{D}_s(t_{m,s} +) \) may produce a large recovery value, leading to a negative degradation value under the original scale.

3. Statistical estimation of the proposed model

In practice, only the measurement schedule \( T \), the observed degradation characteristic \( Y_s(t), t \in T \) and the maintenance times \( \{ t_{m,s} \} \in T \) are available. All the unknown parameters, i.e. the first level parameters in the piecewise linear degradation model, the second level parameters in three ARMA models, and
the parameters in the two maintenance effect models, should be estimated.

Estimating the parameters simultaneously by maximising the likelihood or using the Markov Chain Monte Carlo method is complicated and time-consuming. Therefore, in this section, a heuristic procedure to estimate all the parameters in several steps is presented. At the expense of some statistical estimation efficiency, this procedure is fast and easily understood and implemented by practitioners. The model estimation takes the following steps:

(1) Use the least square algorithm fitting all the \( k + 1 \) segments divided by \( \{ t_{m_1}, t_{m_2}, \ldots, t_{m_k} \} \) separately. The estimates are denoted by the intercept \( \hat{\alpha}_0, \ldots, \hat{\alpha}_k \), the slope \( \hat{\beta}_0, \ldots, \hat{\beta}_k \), and all the residuals \( \hat{\epsilon}_t \) correspond to the best fitted ARMA model by specifying certain model parameters and introducing heteroscedasticity into the model. This result does not preserve the unbiasedness of the estimator. Since both \( \hat{\gamma} \) and \( \hat{\sigma}_t \) are maximum likelihood estimators, they follow normal distributions with the true rate as the mean asymptotically. Therefore, the estimates for the change of rate model might be biased but remain consistent.

(2) For the initial degradation value and the recovery value, respectively, \( \hat{\alpha}_i = \hat{\alpha}_{0,i} \) and \( \hat{\gamma}_{i,j} = Y_s(t_{m_1}) - Y_s(t_{m_1+1}) + \hat{\beta}_{i,j} t_{m_1+1} - \hat{\beta}_{i-1,j} t_{m_1} \). For the degradation rate, \( \hat{\gamma}_i = \ln \hat{\beta}_{0,i} \) and \( \hat{\gamma}_i = \ln \hat{\beta}_{0,i} - \ln \hat{\rho}_{i,j} = 1, \ldots, k \). For the ARMA model, the ARMA model can be used to fit all the regression coefficients, \( a_0, a_1, b_0, b_1, \ldots, u \), and the standard deviation of the error terms \( \tau^2 \) and \( \sigma^2 \).

(3) Replace \( r_{j,i}, \omega_{j,i} \), and \( \Gamma_{j,i} \) by \( \hat{r}_{j,i}, \hat{\omega}_{j,i} \), and \( \hat{\gamma}_{j,i} \) in Equations (5) and (6), the least squares algorithm can be used to fit all the regression coefficients, \( a_0, a_1, b_0, b_1, \ldots, u \), and the standard deviation of the error terms \( \tau^2 \) and \( \sigma^2 \).

(4) Fit three ARMA models to the process \( \{ Y_s(t_0) \} \) and \( \{ \hat{r}_i \} \) as the estimated parameters for the true process \( \{ a_0 \} \) and \( \{ \hat{r}_i \} \) using the maximum likelihood approach within the function auto.arima provided by the package forecast in R. For more information about package forecast in R, the readers are referred to Hyndman and Khandakar (2007).

(5) The procedure concludes.

The procedures mainly use the least squares algorithm. In addition, R is chosen as the major tool to carry out the computation as it provides the most powerful statistical modelling toolbox for both researchers and practitioners (R Core Team, 2016). It can fit a linear model extremely efficiently via the function lm provided by the package stat. The function auto.arima can automatically select the optimal orders \( p \) and \( q \) corresponding to the best fitted ARMA model by specifying certain model selection criteria, i.e. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Before fitting the ARMA model, it is necessary to confirm that the process is stationary; the pre-analysis is presented in the case study. Matlab or SAS can also be used to carry out all the above calculations without difficulty. For the sake of the theoretical justification of the above procedure, we include the following remarks to show our estimates are at least consistent. The reader can check the references at the end of the paper to find the related theory. The estimated parameters can be used to predict the future degradation path of the current track line or to simulate the degradation behaviour of a new one. For prediction, we can directly use the estimated first level parameters and parameters in the maintenance effects model. For the simulation, we should use the estimated second level parameters and ARMA models to generate the degradation parameters of the track line. The rest is similar to the prediction.

Remark 1: Since the least square algorithm is equivalent to the maximum likelihood estimation for a linear model with Gaussian noise, \( \hat{\alpha}_i, \hat{\beta}_i, \hat{\sigma}_i \), and \( \hat{\rho}_i \) are all the maximum likelihood estimators, as the maximum likelihood estimation is invariant under any function transformation of the parameters (see Shao, 2003).

Remark 2: Note that \( \hat{\beta}_{j,i} \) is the unbiased estimator of \( \beta_{j,i} = e^{-\gamma_{j,i} \mu_i} \) in the linear model with Gaussian noise; thus, \( \hat{\beta}_{j,i} \) follows a normal distribution \( \bar{\beta}_{j,i} = \beta_{j,i} \):

\[
E[\hat{\beta}_{j,i}] = E[Y_s(t_{m_1}) - Y_s(t_{m_1+1}) + \hat{\beta}_{j,i} t_{m_1+1} - \hat{\beta}_{j-1,i} t_{m_1}]
\]

\[
= a_j - \sum_{j=1}^{m_1} r_j + \beta_{j-1,i} \gamma_{j,i} - (a_j - \sum_{j=1}^{m_1} r_j + \beta_{j-1,i} \gamma_{j,i})
\]

\[
+ \beta_{j,i} t_{m_1+1} - \beta_{j-1,i} t_{m_1}
\]

\[
r_{j,i}
\]

Therefore, \( \hat{r}_{j,i} \) follows a normal distribution with mean \( r_{j,i} \). This guarantees that the estimates for the recovery model using the least squares method are the best linear unbiased estimates (see Shao, 2003), even if we replace the true value \( r_{j,i} \) by the estimate \( \hat{r}_{j,i} \) and introduce heteroscedasticity into the model. This result does not hold for the change of rate model, as the log transformation does not preserve the unbiasedness of the estimator. Since both \( \hat{\gamma}_i \) and \( \hat{\sigma}_t \) are maximum likelihood estimators, they follow normal distributions with the true rate as the mean asymptotically. Therefore, the estimates for the change of rate model might be biased but remain consistent.

Remark 3: The observed values \( Y_s(t_0) \) and \( Y_s(t_0) \) are used, not the fitted values \( \hat{Y}_s(t_0) \) and \( \hat{Y}_s(t_0) \) when calculating the estimated recovery value \( \hat{r}_{j,i} \) and fitting the ARMA model for the initial value \( \{ a_0 \} \). In particular, \( Y_s(t_0) = a_j + \epsilon_i \) is the true underlying process \( \{ a_j \} \) with a zero mean Gaussian noise process \( \{ \epsilon_i \} \). When the noise process \( \{ \epsilon_i \} \) is suppressed, the estimates given by auto.arima can be regarded as pseudo maximum likelihood estimator; this is also consistent and asymptotic normal (Azrak & Mérard, 2006). Both the estimated standard deviation process \( \{ \hat{\sigma}_t \} \) and the estimated log-rate process \( \{ \hat{\gamma}_i \} \) are consistent and asymptotic normal, as they are maximum likelihood estimates for the true process \( \{ \gamma_i \} \) and \( \{ \hat{r}_i \} \).

\[
\hat{\gamma}_i = \ln \hat{\beta}_{j,i} = \ln \left( \hat{\beta}_{0,i} + n_i^{-1} \epsilon_i \right) = \ln \left( e^{\gamma_i} + n_i^{-1} \epsilon_i \right)
\]

\[
= \gamma_i + \ln \left( 1 + n_i^{-1} e^{\gamma_i} \epsilon_i \right) \approx \gamma_i + n_i^{-1} e^{\gamma_i} \epsilon_i - O(e^{\gamma_i} n_i^{-2})
\]

(14)\n
where \( n_i \) is the number of measurements, i.e. the sample size, to estimate \( \hat{\beta}_{j,i} \). Since \( \gamma_i > 0 \), it is obvious that \( e^{\gamma_i} n_i^{-1} e^{\gamma_i} \epsilon_i < n_i^{-1} \epsilon_i \) \( < e_i \). Even with a small sample size \( n_i \), the second-order remainder term in Equation (14) is very small and
can be neglected. The first-order error term in Equation (14) is zero mean and satisfies the sub-Gaussian property, as it strictly dominated by the Gaussian process \( \{ \varepsilon_i \} \). Furthermore, it will almost surely converge to 0 when \( n_i \to \infty \). Therefore, these two estimated process can be treated as the true underlying processes with a zero mean Gaussian noise process, asymptotically. When the uncertainties introduced by the estimation procedure are ignored, the estimates given by \texttt{auto.arima} can also be regarded as pseudo maximum likelihood estimators; again, this is consistent and asymptotic normal. However, the estimation efficiency of PMLE will be lower than that of MLE in general.

4. Case study

To demonstrate the application of the proposed model, a case study is performed using data from line section 414, part of the Main Western Line in Sweden (Västra Stambanan). The data were collected from the Optram\(^1\) system and contain track geometry measurements and maintenance history data collected from 2007 to 2015. The maximum speed of trains on the Main Western Line is around 200 km/h. Line 414 is 82 km long, divided into 411 track sections, each 200 m in length. During the observation period, 30 inspections were carried out for most of the track sections. However, for a few sections, one to three measurements are missing. The tamping interventions in line 414 are performed based on the track geometry condition. Although no specific tamping threshold has been set, the time between two major tamping interventions is about five years. In total, 511 tamping interventions were performed on the track line from 2007 to 2015. Figure 1 depicts an example of the data for a typical track section.

4.1. Pre-analysis

Pre-analysis of data shows that for 39 track sections (39 maintenance cycles out of a total of 922 maintenance cycles) the track geometry condition exhibits an abrupt change in its degradation path after passing a specific accumulated load (MGT). Figure 2 displays the track geometry degradation path for the track section with the starting point of 80 km + 400 m; it is obvious that there is an unusual trend in the degradation path.

To identify the reasons for an abrupt change in the degradation path, we monitored the changes in measured geometry values for every 25 cm over time. It was found that the bumps (peaks) in the geometry data for every 25 cm correspond to sudden changes in the degradation path. The bumps may occur for different reasons. For some track sections, they appear in the location of a sleeper replacement or a steel drum installation. It is necessary to compact the ballast under the sleeper properly after replacing it. Otherwise, the replacing activity will cause irregularities in track geometry; i.e. it may change the current degradation condition. It usually causes an increase in degradation rate. The installation or replacement of steel drums may also cause a change in track geometry condition. On Line 414, drums are generally installed 1.8 m under the track surface. When we look at the data, we may infer that this process damages the condition of the track geometry and results in abrupt changes in the degradation path. We do not observe that one maintenance cycle of a section contains more than one change point. For those sections with an identified change point in a maintenance cycle, we fit two separate lines to model degradation using the piecewise linear function.

4.2. First level model – piecewise linear degradation model

This subsection explains the exploratory data analysis of the degradation data to address uncertainty in degradation parameters over track sections. More specifically, we estimate the first level parameters of the linear degradation model and investigate the variations in the parameters over the track sections. In the following sections, these estimates are used to estimate the second level parameters in three ARMA models and the parameters in the maintenance effect models. The left panels of Figure 3 present the estimates for all \( \{ a_i \}, \{ \gamma_i \}, \) and \( \{ \rho_i \} \) over the track sections obtained by applying the least squares algorithm. The values of the model parameters obviously change over track sections. The right panels of Figure 3 show the histograms of \( \{ a_i \}, \{ \gamma_i \}, \) and
4.3. Maintenance effect model

To construct the recovery model, we consider several factors in the initial analysis, i.e. track geometry condition before tamping, speed class, the number of tamping interventions and tamping type (complete or partial). However, since trains on all track sections operate at the same speed, the factor of speed is not considered in the proposed model. In addition, since the data were collected from 2007 onward, the complete historical information about the number of accumulated tamping interventions on each track section is not available. Therefore, this factor is not considered. In this study, the recovery model is constructed by considering the track geometry condition before tamping as the dominant explanatory variable and the tamping type as the other explanatory variable. The interaction effect between the track geometry condition before tamping and the tamping type is also considered. According to Equation (5) the model is presented as follows:

\[
\begin{align*}
\text{(15)}
\end{align*}
\]

where \( b_{1,0} \) is the coefficient for the variable representing the track geometry condition before tamping, \( b_{1,1} \) is the coefficient for the variable representing the tamping type, and \( b_{1,2} \) is the coefficient for interaction of the two mentioned variables.

By utilising Equation (5), the first index of the \( b \) shows that the coefficient is for the recovery model. In addition, \( X_{1,s,t} \) is a dummy variable with value 0 when tamping is done partially on section \( s \) at inspection time \( t \) and value 1 when tamping is done completely. Naturally, when a longer segment of a track section is tamped, there is a greater chance of a larger recovery value. Therefore, we should distinguish between partial and complete tamping interventions when estimating the recovery values after tamping. In addition, we see an increasing trend between the recovery value and the tamped length for a specific section. As information about the exact length of

\[
\begin{align*}
\text{Figure 3. Line plots over } S \text{ and histograms of } (\alpha_s), (\gamma_s) \text{ and } (\rho_s). \\
\text{Table 1. Results of AD test and KS test for parameters of degradation model.} \\
\text{\begin{tabular}{|l|l|l|l|}
\hline
\( \alpha_s \) & \( \gamma_s \) & \( \rho_s \) \\
\hline
\text{AD Test} & 0.681(0.075) & 0.723(0.059) & 0.702(0.066) \\
\text{KS Test} & 0.035(0.68) & 0.037(0.62) & 0.035(0.67) \\
\hline
\end{tabular}}\]
\]

\*For hypothesis tests, the first number is the test statistic and the number in brackets is the \( p \)-value.
Figure 8 presents the relation between recovery values and track geometry conditions before tamping and also the type of tamping. As shown in this plot, the recovery values after partial tamping is not complete in our data, we simply use the dummy variable and do not consider the effect of this factor in our recovery model.

Figure 4. Histogram of $p$-values for the normality of residuals using KS test and AD test (Left panel: KS test; Right panel: AD test).

Figure 5. Sample ACF and PACF for $\alpha_c$.

Figure 6. Sample ACF and PACF for $\gamma_c$.

Figure 7. Sample ACF and PACF for $\rho_c$. 
tamping are, on average, smaller than those after complete tamping. In fact, the partial tamping improves the track geometry condition less than the complete tamping. Table 2 presents the estimated parameters of the recovery model by fitting the proposed model using the least squares algorithm. Although it was not possible to consider other covariates in this case study, e.g. the number of previous tampings, the tamping length, and the speed class, for a general recovery model, they should be considered.

Furthermore, tamping may cause a change in the degradation rate. Therefore, by considering the degradation rate before tamping as the dominant factor, the change in rate model can be given as follows:

\[ \omega_{j,s} = a_2 + b_{2,0} \Gamma_s \left( t_{m_j} \right) + \delta_s \]  

Figure 9 illustrates the changes in rate vs. the degradation rate before tamping; Table 2 presents the estimated parameters of the model. There is an inverse linear relationship between the degradation rate before tamping and the changes in rate. When the degradation rate before tamping is high, there is a high possibility that tamping will reduce the degradation rate. In fact, when the degradation rate is high, there may be a problem in part of the track section that tamping intervention could repair.

As noted in several studies (Audley & Andrews, 2013; Quiroga & Schnieder, 2012), the number of accumulated tamping interventions affects the degradation rate. A higher degradation rate is expected after a number of tamping interventions. However, as mentioned, as full information about the number of accumulated tamping interventions is not available in this case study, we do not consider this factor. If the information is available, this factor must be considered as a covariate in the proposed model. The possible interaction of the covariate with the degradation rate before tamping can also be considered in the model.

The normality assumption of the residuals of the recovery and change in rate models are tested using the AD and KS tests, and the results are provided in Table 3. As can be seen, the \( p \)-values of both tests are larger than the significance level (0.05) for both models, and there is no evidence to reject the null hypothesis on the normality of the residuals.

4.4. Spatial model for degradation parameters

To properly fit ARMA models on \( s_1 \), \( s_2 \), and \( s_3 \), we perform several different tests. To determine whether the time series are stationary, we use the Augmented Dickey–Fuller test (ADF). The null hypothesis is that the process needs to be differenced to make it stationary. The results of the ADF test for the three processes are shown in Table 4. As the table shows, all \( p \)-values are smaller than the significance level. Therefore, it can be concluded that the three time series are stationary, and there is no need to consider the differencing operations.

In order to choose the best fitted model among different ARMA models, we compare the models with respect to AIC and BIC as the model selection criteria. Both of AIC and BIC are composed of two parts; a goodness of fit term and a penalty term for the number of estimated parameters in the model. The difference between these two criteria is that BIC penalises more than AIC and hence tends to give simpler models. The models with the lowest AIC and BIC values for the three time series are chosen. The optimal AIC and BIC values for the three time series are, on average, smaller than those after complete tamping. In fact, the partial tamping improves the track geometry condition less than the complete tamping. Table 2 presents the estimated parameters of the recovery model by fitting the proposed model using the least squares algorithm. Although it was not possible to consider other covariates in this case study, e.g. the number of previous tampings, the tamping length, and the speed class, for a general recovery model, they should be considered.

Furthermore, tamping may cause a change in the degradation rate. Therefore, by considering the degradation rate before tamping as the dominant factor, the change in rate model can be given as follows:

\[ \omega_{j,s} = a_2 + b_{2,0} \Gamma_s \left( t_{m_j} \right) + \delta_s \]  

Figure 9 illustrates the changes in rate vs. the degradation rate before tamping; Table 2 presents the estimated parameters of the model. There is an inverse linear relationship between the degradation rate before tamping and the changes in rate. When the degradation rate before tamping is high, there is a high possibility that tamping will reduce the degradation rate. In fact, when the degradation rate is high, there may be a problem in part of the track section that tamping intervention could repair.

As noted in several studies (Audley & Andrews, 2013; Quiroga & Schnieder, 2012), the number of accumulated tamping interventions affects the degradation rate. A higher degradation rate is expected after a number of tamping interventions. However, as mentioned, as full information about the number of accumulated tamping interventions is not available in this case study, we do not consider this factor. If the information is available, this factor must be considered as a covariate in the proposed model. The possible interaction of the covariate with the degradation rate before tamping can also be considered in the model.

The normality assumption of the residuals of the recovery and change in rate models are tested using the AD and KS tests, and the results are provided in Table 3. As can be seen, the \( p \)-values of both tests are larger than the significance level (0.05) for both models, and there is no evidence to reject the null hypothesis on the normality of the residuals.

### Table 2. Parameter estimates for recovery model and change of rate model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Recovery Model Estimates</th>
<th>Change in Rate Model Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_1 )</td>
<td>0.25(0.014)</td>
<td>−0.27(0.034)</td>
</tr>
<tr>
<td>( b_{1,0} )</td>
<td>0.39(0.039)</td>
<td>−0.86(0.026)</td>
</tr>
<tr>
<td>( b_{1,1} )</td>
<td>0.28(0.021)</td>
<td>−</td>
</tr>
<tr>
<td>( b_{1,2} )</td>
<td>−0.27(0.052)</td>
<td>−</td>
</tr>
<tr>
<td>( \sigma^2 )</td>
<td>0.05</td>
<td>0.17</td>
</tr>
</tbody>
</table>

*For the parameter estimates, the first number is the estimated value and the number in brackets is the standard deviation of the estimates.*
and the estimated parameters of the ARMA models for the three processes are given in Table 4.

To test the goodness of fit of the models, we apply the Box–Ljung (BL) test with 20 lagged observations to the residuals of the time series after fitting the ARMA models. The null hypothesis states that the residuals will follow a white noise process. The results of the BL test for the three ARMA models corresponding to $\gamma_s$, $\alpha_s$, and $\rho_s$ are presented in Table 4. As can be seen, all $p$-values are bigger than the significance level. Thus, there is not enough evidence to reject the null hypothesis.

The results of the analysis of the spatial correlation of $\{\gamma_s\}$, $\{\alpha_s\}$, and $\{\rho_s\}$ for adjacent track sections show that the parameters of the ARMA model for $\{\alpha_s\}$ and $\{\rho_s\}$ are significantly larger than for $\{\gamma_s\}$. Therefore, we may infer the spatial correlation of $\{\gamma_s\}$ for adjacent track sections is weaker than other parameters. In addition, a comparison of the coefficients of the three ARMA models shows the highest correlation is between the sections $\rho_s$.

As discussed in Remark 3 of Section 3, the data input for estimating the ARMA model is true parameter with a noise process, as $Y_s(t_s) = \alpha_s + \epsilon_s$ and $\hat{\gamma}_s \approx \gamma_s + n_s^{-1}e^{-\gamma_s}\epsilon_s$. Since $e$ depends on $\rho_s$, additional dependencies are introduced into the input. However, the additional dependencies are almost negligible. Take $Y_s(t_s) = \alpha_s + \epsilon_s$ as an example. We denote the auto-covariances of $Y_s(t_s)$, $\alpha_s$, and $\epsilon_s$ by $\sigma^2 = \text{Cov}(Y_s(t_s), Y_s(t_s'))$, $\sigma_{\alpha,\epsilon} = \text{Cov}(\alpha_s, \epsilon_s)$, and $\sigma_{\epsilon,\epsilon} = \text{Cov}(\epsilon_s, \epsilon_s)$, respectively, and their autocorrelation coefficients by $\phi_1 = \sigma^2 / \sigma_{\epsilon,\epsilon}$, $\phi_2 = \sigma_{\alpha,\epsilon} / \sigma_{\epsilon,\epsilon}$, and $\phi_3 = \sigma_{\epsilon,\epsilon} / \sigma_{\epsilon,\epsilon}$, respectively. As $\alpha_s$ is independent with $\epsilon_s$, we get the following result:

$$\phi_1 = \frac{\sigma_{\alpha,\epsilon}}{\sigma^2} = \frac{\sigma_{\alpha,\epsilon}}{\sigma_{\epsilon,\epsilon}} = \frac{\phi_{1,0}}{\sigma_{\epsilon,\epsilon}}$$

According to $\epsilon_s|\rho_s \sim N(0, e^{\rho_s})$ and the conditional variance law:

$$\varsigma_{0,2} = \text{Var}(\epsilon_s) = E[\text{Var}(\epsilon_s|\rho_s)] + \text{Var}(E[\epsilon_s|\rho_s]) = E[e^{\rho_s}]$$

where $e^{\rho_s}$ follows a lognormal distribution. This variance is the second-moment of the lognormal distribution, given as:

$$\varsigma_{0,2} = e^{2\mu + \sigma^2}$$

With the estimated ARMA process on $\rho_s$, we can estimate $C_{0,2}$ by plugging the estimated mean and variance of $\rho_s$ into Equation (18). The estimate of $\varsigma_{0,2} = 0.0026$ while the estimate of $\varsigma_{0,2} = 0.1295$. Substituting them into Equation (17), it is obtained:

$$\phi_1 \approx 0.98\phi_{1,1} + 0.02\phi_{1,2}$$

This means the correlation structure of the estimated parameter process is very close to the true process. A similar procedure can be applied to $\hat{\gamma}_s \approx \gamma_s + n_s^{-1}e^{-\gamma_s}\epsilon_s$, but will be a little more complex, as $\gamma_s$ is not independent with $n_s^{-1}e^{-\gamma_s}\epsilon_s$. Since $\text{Var}(\gamma_s)$ is much larger than $\sigma_{\epsilon,\epsilon}$, and the factor $n_s^{-1}$ further shrinks the magnitude of the variance of the noise term, the approximation of the estimated parameter should be even better.

The simulation based on the ARMA process is simple. After the orders $p$ and $q$, the AR and MA parameters, and the innovation in the ARMA process are specified, the function arima.sim in R package forecast can efficiently generate arbitrary numbers of realisations of this process. To study the spatial correlation structure more deeply, readers may refer to the matrix transformation technique in the Appendix 1; it gives the CAR representation for any stationary ARMA process.

### Table 3. Results of AD test and KS test for residuals of recovery and change in rate models.

<table>
<thead>
<tr>
<th>Test</th>
<th>$\delta_1$</th>
<th>$\delta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD Test</td>
<td>0.255(0.724)</td>
<td>0.327(0.516)</td>
</tr>
<tr>
<td>KS Test</td>
<td>0.028(0.809)</td>
<td>0.034(0.891)</td>
</tr>
</tbody>
</table>

*For hypothesis tests, the first number is the test statistic and the number in brackets is the $p$-value.

### Table 4. Summary of estimations of ARMA processes.

<table>
<thead>
<tr>
<th>Test</th>
<th>${\alpha_s}$</th>
<th>${\gamma_s}$</th>
<th>${\rho_s}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF Test</td>
<td>-5.501(0.01)</td>
<td>-6.57(0.01)</td>
<td>-5.66(0.01)</td>
</tr>
<tr>
<td>BL Test</td>
<td>7.44(0.99)</td>
<td>17.94(0.59)</td>
<td>20.92(0.402)</td>
</tr>
<tr>
<td>Model</td>
<td>ARMA(1,1)</td>
<td>AR(1)</td>
<td>ARMA(2,2)</td>
</tr>
<tr>
<td>Constant Term</td>
<td>$c$</td>
<td>-0.197(0.037)</td>
<td>-1.051(0.041)</td>
</tr>
<tr>
<td>AR Parameters</td>
<td>$\phi_s$</td>
<td>0.796(0.0611)</td>
<td>0.1945(0.0487)</td>
</tr>
<tr>
<td>MA Parameters</td>
<td>$\theta_s$</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Variances</td>
<td>$\sigma_s$</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>AIC</td>
<td>260.11</td>
<td>315.42</td>
<td>172.26</td>
</tr>
<tr>
<td>BIC</td>
<td>276.19</td>
<td>327.47</td>
<td>200.39</td>
</tr>
</tbody>
</table>

*For hypothesis tests, the first number is the test statistic and the number in brackets is the $p$-value. For the parameter estimates, the first number is the estimated value and the number in brackets is the standard deviation of the estimates.

5. **Conclusions**

The paper discusses the statistical modelling of the evolution of track geometry condition. The proposed model can be used to predict and simulate the evolution of track geometry condition over a track line. The contributions of this paper are threefold: first, it proposes a two-level piecewise linear framework to model the evolution of the track geometry degradation over a spatial and temporal space. A widely used simple linear model is used as the basic component for a single maintenance cycle in the proposed model. Two multivariate regression models are used to link the degradation parameters in two consecutive maintenance cycles. ARMA models are used to capture the possible spatial dependence structures within the degradation parameters of different sections. Second, to estimate the parameters, the paper proposes a fast procedure to estimate all the unknown parameters without employing a complex numerical algorithm and provides the necessary theoretical justifications. Third, it implements the model using a comprehensive case study with data from the Main Western Line in Sweden.

Findings show recovery after tamping intervention is strongly dependent on tamping type (partial or complete). In addition, the tamping effect on the degradation rate is dependent on the degradation rate right before the tamping intervention. Finally, the degradation parameters over the spatial interval may be generated by some Gaussian processes and the ARMA processes provide a satisfactory fit to them.

For the future research, it will be beneficial to consider abrupt changes in the degradation path which is ignored in our pre-analysis. Since an abrupt change can significantly increase the
degradation rate, it is risky to simply discard data after change. This problem is closely related to outlier detection and change point detection. Since it is necessary to test the existence of the change point in each maintenance cycle for every section, all the hypothesis tests comprise a multiple testing problem. To predict and simulate track geometry degradation with possible change points, it is essential to model the occurrence time of the change points and their effects on the degradation parameters. However, we may have only a few change points over the track line. The issue makes the problem an attractive but challenging topic.

**Note**

1. Optram is the system used since 2007 by the Swedish Transport Administration (Trafikverket) to study measurements of the track and overhead lines. This system visualises and graphically represents the track geometry measurements. Optram provides the functionalities for analysing and displaying data trends (Arasteh Khouy, 2013).

**Acknowledgements**

The authors wish to thank Trafikverket and Luleå Railway Research Centre (JVTC) for the technical and financial support provided during this project. Specific thanks are extended to Dr. Iman Arasteh Khouy from Trafikverket for his valuable support and helpful discussions.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**References**


where $Z_{-i} = \{Z_j | j \neq i\}$, $\mu$ is the mean of this CAR process, $c_j$ are constants where $c_0 = 0$ and $\sigma^2$ is conditional variance of $Z_j$ given $Z_{-i}$ such that

\[
(X_1, \ldots, X_s) \equiv (Z_1, \ldots, Z_s)
\]

where $\equiv$ means two sides follow the same joint distribution.

Proof: It is easy to check that when (A1) is stationary and by letting $\mu = E(X_1, \ldots, X_s)$ and $(Z_1, \ldots, Z_s)$ will have the same mean vector. Without loss of generality, assume the ARMA process and the CAR process have been centred, i.e. $\mu = 0$. Therefore, $(X_1, \ldots, X_s) \sim MVN(0, \Sigma)$, and the covariance matrix is given as

\[
\Sigma = \text{Var}(X_s) = \begin{pmatrix}
Y_0 & Y_1 & \cdots & Y_{s-1} \\
Y_1 & Y_0 & \cdots & Y_{s-2} \\
\vdots & \vdots & \ddots & \vdots \\
Y_{s-1} & Y_{s-2} & \cdots & Y_0
\end{pmatrix}
\]

where $Y_j = \text{Corr}(X_1, X_{s+j})$ is the value of the $j$th lagged autocorrelation coefficient of (A1).

Furthermore, $(Z_1, \ldots, Z_s)$ generated from a CAR process defined by (A2) follows another multivariate normal distribution, $MVN(0, \Sigma)$, and the covariance matrix is given as

\[
\tilde{\Sigma} = (I_s - C)^{-1} \Lambda
\]

where $\Lambda = \text{diag}(t_1^2, t_2^2, \ldots, t_s^2)$, $I_s$ is the identity matrix, and $C$ is a matrix with the $(i, j)$th entry $c_{ij}$ (Wall, 2004). Then, it is equivalent to show that $\Sigma$ can be represented in the form of $\tilde{\Sigma}$. First, $\Sigma$ is a covariance matrix, which is a positive definite matrix. Therefore, $\Pi = \Sigma^{-1}$ is also a positive definite matrix, and the $(i, j)$th entry of $\Pi$ can be expressed as $\pi_{ij}$. Subsequently, by letting $\tau^2 = \pi_{ii}c_i = 0$, $\tau_i = \pi_{i}/\pi_{ii}$, if $j \neq i$, it is derived that $\Pi = \Lambda^{-1}(I_s - C)$. Therefore, $\Sigma = (I_s - C)^{-1} \Lambda$, and this concludes the proof.
Prediction of railway track geometry defects: a Case Study

Prediction of railway track geometry defects: a case study
Iman Soleimanmeigouni¹, Alireza Ahmadi¹, Arne Nissen², Xun Xiao³

¹Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden
²Trafikverket, Luleå, Sweden
³School of Fundamental Sciences, Massey University, Palmerston North, New Zealand

Abstract
The aim of this study has been to develop a data-driven analytical methodology for prediction of isolated track geometry defects, based on the measurement data obtained from a field study. Within the study, a defect-based model has been proposed to identify the degradation pattern of isolated longitudinal level defects. The proposed model considered the occurrence of shock events in the degradation path. Furthermore, the effectiveness of tamping intervention in rectifying the longitudinal level defects was analysed. The results show that the linear model is an appropriate choice for modelling the degradation pattern of longitudinal level defects. In addition, a section based model has been developed using binary logistic regression to predict the probability of occurrence of isolated defects associated with track sections. The model considered the standard deviation and kurtosis of longitudinal level as explanatory variables. It has been found that the kurtosis of the longitudinal level is a statistically significant predictor of the occurrence of isolated longitudinal level defects in a given track section. The validation results show that the proposed binary logistic regression model can be used to predict the occurrence of isolated defects in a track section.

Keywords: Railway track maintenance, geometry defect, degradation, prediction, intervention limit, linear regression, binary logistic regression, shock events, tamping

1. Introduction
Nowadays railways are experiencing higher demands for the transportation of passengers and goods. This will in turn impose higher demands on the railway capacity and service quality. As a result, infrastructure managers are being compelled to make strategies and plans to meet new requirements, which include a higher level of resilience against failure, more robust and available infrastructure, and cost reduction. To achieve these goals, one of the key elements is the employment of an effective and efficient maintenance programme. A large part of the railway maintenance burden concerns track geometry maintenance. Maintenance actions are used to control the degradation of the track and restore the track geometry condition to an acceptable state. Activities that can be applied in track geometry maintenance include manual interventions, tamping and stone blowing. Three indicators are used to represent the track geometry quality, i.e. the standard deviation over a specific track length, the mean value, and the extreme value of isolated defects (EN 13848–5, 2008). Track quality indices (TQIs) are mostly formed based on the standard deviation of the vertical and lateral track irregularities (Soleimanmeigouni, Ahmadi, & Kumar 2018). Since a TQI aggregates the track geometry measurements to represent the overall condition of track sections, it may not provide complete information about severe isolated defects in the track sections. Isolated defects are short irregularities in the track geometry that can dramatically increase the dynamic forces between the wheel and rail, which in turn will accelerate the growth or occurrence of internal rail defects. These defects have been classified by the Swedish Transport Administration, Trafikverket, into three severity levels, i.e. UH1¹ defects, UH2² defects and critical defects (Trafikverket, 2015). UH1, UH2 and critical defects are those defects which have exceeded the lower bound of the intervention limit, the upper bound of the intervention limit, and the safety limit, respectively (see Figure 2 for details). The occurrence of severe geometry defects can cause comfort problems for passengers, damage to track components and an increase in the risk of derailment. Moreover, these defects cause unplanned maintenance activities on the railway track, which in turn decrease the track availability and safety and increase the maintenance cost. In order to develop a more efficient and effective maintenance plan, isolated defects must be considered in track geometry degradation modelling. There are a limited number of studies that have considered isolated defects in their degradation modelling and maintenance planning. Arasteh Khouy, Larsson-Kräik, Nissen, Juntti, and Schunnesson (2014) used survival analysis and considered isolated defects of the longitudinal level, twist 3 m, and twist 6 m for the purpose of predicting the need for corrective tamping actions. They found that, in their case study, the Weibull distribution was the best fitted distribution for modelling the probability of the occurrence of geometry defects. Later Alemazkoor, Ruppert, and Meidani (2018)
used survival analysis, taking into account a number of covariates, to predict the probability that a “yellow tag” defect would turn into a “red tag” defect in a given time. “Red tag” defects can be defined as isolated defects that exceed the maintenance limits of the Federal Railroad Administration (FRA). “Yellow tag” defects are those isolated defects which are as yet below the FRA limits, but will finally turn into “red tag” defects (Alemazkoo et al., 2018). They chose the Weibull distribution for modelling the time of the occurrence of “red tag” defects and modelled the scale parameter as a function of independent variables, i.e. initial absolute values of the amplitude and length of defects, the track code, the class of track, the operating speed, and the tonnage. Later they used their model to predict the probability that a track section would contain at least one “red tag” defect in a given time. Andrade and Teixeira (2013) predicted the probability that a given track section would need unplanned maintenance due to a severe geometry defect. They applied logistic regression to predict that probability and considered the standard deviation of the longitudinal level and alignment, and the existence of a switch or bridge in a section as explanatory variables. He, Li, Bhattachariya, Parikh, and Hampapur (2014) proposed a degradation model for capturing the changes of the amplitude of geometry defects and for predicting the probability that a “yellow tag” defect would turn into a “red tag” defect in a given time. They considered four explanatory variables in their model, namely the monthly traffic, the total monthly number of cars, the total monthly number of trains, and the number of inspections since the latest “red tag” defect. In addition, they applied the Cox proportional hazard model to predict the probability of derailment by considering the number and the amplitude of “yellow tag” defects. Cárdenas-Gallo, Sarmiento, Morales, Bolivar, and Akhavan-Tabatabaei (2017) proposed an ensemble classifier for predicting when a “yellow tag” defect would turn into a “red tag” defect. They studied the problem with regard to three different aspects, i.e. deterioration, regression and clustering. With regard to deterioration, they applied a gamma process to model the evolution of the amplitude of “yellow tag” defects over time. To identify the relationship between the explanatory variables and the future state of defects, they applied a logistic regression model. In addition, they used support vector machines (SVM) to predict the probability of the occurrence of a “red tag” defect. Sharma, Cui, He, Mohammadi, and Li (2018) proposed a Markov decision process for track maintenance decision making considering geometry defects. They used three algorithms, i.e. a random forest algorithm, a support vector machine algorithm and a logistic regression algorithm to predict the occurrence of severe isolated defects.

Although the literature addressed the track geometry isolated defects, still there is a need to a comprehensive study on the evolution of isolated defect to identify their degradation pattern. Moreover, although the effect of tamping on standard deviation of geometry parameters are well studied in the literature, its effect on isolated defects has not been extensively studied (Soleimanmeigouni, Ahmadi, Arasteh Khouy, & Letot, 2018). The other important point that must be considered for track geometry degradation modelling is considering the occurrence of shock events. Shock events in degradation path are assumed to be negative, causing the increase in degradation rate and even the failure of the system. Furthermore, according to the reviewed papers, different approaches are used to predict the probability of occurrence of isolated defects by considering aggregated TQIs as explanatory variables. However, to the best of our knowledge there is no study concerning the relationship between kurtosis of geometry parameters and the probability of occurrence of isolated defects. To this end, the present study was undertaken to explore the prediction of isolated track geometry defects and to address the abovementioned issues.

The aim of the present study has been to model the track geometry degradation and to predict the occurrence of UH2 defects. In order to model the track geometry degradation, the evolution of the amplitude of the longitudinal level defects within a maintenance cycle was modelled using simple linear regression. In addition to gradual degradation, there can be an abrupt change in the degradation path in which the amplitude of the defect will dramatically increase over time. We call this phenomenon a shock event, and this kind of event was also considered in our degradation modelling. Furthermore, the effectiveness of tamping intervention in rectifying the longitudinal level defects was analysed. In order to predict the probability of the occurrence of UH2 defects, a section-based model was developed using binary logistic regression, where the kurtosis of the longitudinal level also considered as an explanatory variable. In order to validate the model for the purpose of classification of track sections based on the presence of UH2 defects, the sensitivity and specificity of the developed model were calculated. Data obtained from the Main Western Line in Sweden were used for the purpose of model development and for the case study. The rest of the paper is organized as follows. Section 2 provides a background to the topics of track geometry parameters and maintenance limits. Section 3 presents information on the track line used for the case study and explains the data recording and preparation for the case study. Section 4 presents the process of developing the degradation model. The section-based model, which considers the relationship between common indicators for planned maintenance and the occurrence of UH2 defects, is presented in Section 5. Finally, Section 6 provides the conclusions and proposes directions for future studies.
### 2. Track geometry parameters and maintenance limits

Track geometry parameters are widely used to represent the track condition and to plan maintenance activities, and these parameters can be divided into five classes: longitudinal level, alignment, gauge, cant, and twist. Longitudinal level is the geometry of the track centreline projected onto the longitudinal vertical plane. Alignment is the geometry of the track centreline projected onto the longitudinal horizontal plane. Gauge is the distance between the gauge faces of two adjacent rails at a given location below the running surface. Cant (cross-level) is the difference in height between the adjacent running tables computed from the angle between the running surface and a horizontal reference plane. Twist is the algebraic difference between two cross-levels taken at a defined distance apart, usually expressed as the gradient between the two points of measurement (SS-EN 13848-1: 2004+A1, 2008). According to SS-EN 13848-1: 2004+A1 (2008), for each geometry parameter the indicators provided in Table 1 should be used to represent the track geometry quality.

<table>
<thead>
<tr>
<th>Geometry parameter</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longitudinal level</td>
<td>Isolated defects that exceed a prescribed threshold. The indicator is mean to peak value. Standard deviation over a defined length, typically 200 m, in the wavelength range $D_1$.</td>
</tr>
<tr>
<td>Alignment</td>
<td>Isolated defects that exceed a prescribed threshold. The indicator is mean to peak value. Standard deviation over a defined length, typically 200 m, in the wavelength range $D_1$.</td>
</tr>
<tr>
<td>Cross-level</td>
<td>Absolute value.</td>
</tr>
<tr>
<td>Twist 3 and 6 m</td>
<td>Isolated defects that exceed a prescribed threshold. The indicator is zero to peak value. Standard deviation over a defined length, typically 200 m.</td>
</tr>
<tr>
<td>Gauge</td>
<td>The identification of individual defects which exceed a prescribed threshold. The measured track gauge. The difference between the measured track gauge and the nominal track gauge. The mean track gauge over a specified distance. The variation of the track gauge over a specified distance.</td>
</tr>
</tbody>
</table>

According to EN 13848-5 (2008) there are three limits for maintenance actions: the immediate action limit (IAL), the intervention limit (IL) and the alert limit (AL). If the IAL or safety limit is exceeded, there is a potential risk of derailment, and consequently a speed reduction or line closure must be imposed before conducting a corrective maintenance action. If the IL or corrective maintenance limit is exceeded, a corrective maintenance action is required before the IAL is reached. If the AL or preventive maintenance limit is exceeded, the track geometry must be analysed for the planning of future maintenance actions. The European standard EN 13848-5 (2008) provides the IALs, ILs and ALs for isolated defects and the ALs for standard deviations. Generally, track quality indicators based on the standard deviation of the track geometry parameters are used to plan and perform preventive maintenance actions. On the other hand, the execution of corrective maintenance actions is based on the severity of isolated defects. Whenever the amplitude of an isolated defect exceeds the IL or IAL, corrective maintenance should be conducted on the track. Figure 1 shows the different maintenance zones based on the above-mentioned limits.
EN 13848-5 (2008) reflect the common practice of most European infrastructure managers. In alignment with the European standard EN 13848-5 (2008), Trafikverket has defined four main limits, namely the planning limit, the UH1 limit, the UH2 limit, and the critical limit, as can be seen in Figure 2. In Trafikverket (2015) the intervention limit is expressed as a range rather than a discrete value. Track irregularities that exceed the UH1 limit must be assessed for conducting maintenance before the UH2 limit is exceeded. For track irregularities exceeding the UH2 limit, maintenance action must be taken without unnecessary delay. Therefore, track irregularities must be corrected before the UH2 limit is reached. Table 2 shows the relation between the limits defined in EN 13848-5 (2008) and those defined in Trafikverket (2015).

<table>
<thead>
<tr>
<th>EN 13848-5 limits</th>
<th>Trafikverket’s limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert limit</td>
<td>Planning limit</td>
</tr>
<tr>
<td>Intervention limit</td>
<td>UH1 limit (lower bound for corrective maintenance)</td>
</tr>
<tr>
<td></td>
<td>UH2 limit (upper bound for corrective maintenance)</td>
</tr>
<tr>
<td>Immediate action limit</td>
<td>Critical limit</td>
</tr>
</tbody>
</table>

Based on the above-mentioned limits, geometry defects can be classified according to their severity into three groups, i.e. UH1, UH2 and critical defects. UH1, UH2 and critical defects occur when track irregularities exceed the UH1, UH2 and critical limits, respectively. Examples of these three categories of defects are shown in Figure 2.

3. Data collection and preparation

Track geometry data for line section 414 between Järna and Katrineholm Central Station, collected from January 2015 to July 2018, were used to perform the case study presented in this paper. Line section 414 is part of the Main Western Line in Sweden (Västra Stambanan). The line section data were taken from Optram, which is the system used by Trafikverket for the study of measurements performed on the track and overhead lines. The maximum speed of trains on the Main Western Line is around 200 km/h. Line section 414 is 82 km long and consists of UIC 60 and SJ 50 rails, M1 ballast, Pandrol e-Clip fasteners, and concrete sleepers. The annual passing tonnage of the line section is around...
20 MGT. This line section is divided into 440 track sections with different lengths, mostly between 100 m and 300 m. It must be mentioned that, generally, the track is divided into small sections of the same length (usually 200 m) to evaluate the behaviour of the track geometry better and to plan for maintenance activities. However, in this case study the segmentation performed based on track objects and attributes. In this approach different track sections may have different length. This segmentation provides the opportunity to analyses and compares the degradation behavior of different section types. Table 3 summarize the abovementioned information about the line section 414 which is used to perform the case study presented in this paper.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>length</td>
<td>82 km</td>
</tr>
<tr>
<td>Number of track sections</td>
<td>440</td>
</tr>
<tr>
<td>Ballast type</td>
<td>M1</td>
</tr>
<tr>
<td>Rail type</td>
<td>UIC 60 and SJ 50</td>
</tr>
<tr>
<td>Sleeper type</td>
<td>Concrete</td>
</tr>
<tr>
<td>Fastener type</td>
<td>Pandrol e-Clip</td>
</tr>
<tr>
<td>Annual passing tonnage</td>
<td>20 MGT</td>
</tr>
<tr>
<td>Maximum allowable speed</td>
<td>200 km/h</td>
</tr>
</tbody>
</table>

In order to measure the vertical and lateral deviation of the track, geometrical parameters are measured and recorded by two inspection cars, i.e. the IMV 100 with a speed up to 80 km/h and the IMV 200 with a speed up to 160 km/h. In order to model the track geometry degradation and predict track geometry defects, the longitudinal level measurements were considered in this case study. Concerning longitudinal level irregularities, a total of around 1,300 UH1 defects and 450 UH2 defects are recorded in the dataset. By analysing the measurement data we found that there is a slight difference between the measurements obtained by different types of measurement cars. In fact, using the measurement data gathered by different measurement cars might have resulted in significant errors in the prediction of the occurrence of isolated defects. Therefore, to avoid this kind of error, measurements were carried out by the IMV 200 inspection car were used in this case study to model the track geometry degradation. During the observation period, 15 inspections were carried out for most of the track sections. The data used for the case study include measurement dates, measurement speeds, measurement IDs, measurement car IDs, TQI data, isolated defect data, and data on infrastructure-related features such as the presence of special assets, e.g. switches and crossings.

### 3.1. Data alignment

Generally, track geometry measurement data obtained by inspection cars suffer from measurement errors and positional errors. The positional error can be as much as 100 metres in some cases (Wang, Wang, Wang, & Liu, 2018). Positional errors negatively affect the accuracy of the track irregularity predictions achieved by track geometry degradation models. A main source of positional errors is the presence of error in reference mileposts installed along the track or the Global Positioning System (GPS), while another source of positional error is errors in the measured travelling distance. The accuracy of the measured travelling distance is affected by various factors, including the condition of the wheel-rail contact (Xu, Sun, Liu, & Wang, 2013). In order to correct the positional measurement data, absolute position-based (APB) and relative position-based (RPB) methods can be applied. APB methods correct the positional measurement data based on the estimated absolute position information of the physical reference position. RPB methods correct the positional measurement data according to the estimated positional shift relative to historical inspection data. RPB methods are used to align measurement data collected in different inspections (Xu, Sun, Liu, & Souleyrette, 2016). For more studies regarding the APB methods and RPB methods, readers are referred to (Wang et al. 2018, Xu et al. 2016). Since the aim of our study has been to analyse the evolution of location-specific defects over time, the possession of accurate positions of the defects is of crucial importance. In order to align the measurement data of different inspections, OPTRAM uses a cross-correlation algorithm at specified intervals. Data alignment can be accomplished by cross-correlation analysis performed on the measured waveforms of inspections for different numbers of channels. The first channel roughly aligns the measurement data and the other channels refine the alignment (Selig, Cardillo, Stephens, & Smith, 2008). In the present study, the data alignment was performed separately using three different channels, i.e. short wavelength longitudinal level (wavelength 1-25 m), medium wavelength longitudinal level (wavelength 25-70 m), and gauge. If two of the alignments are within the same range, this is proof that the alignment is correct. The main advantage of using the medium wavelength longitudinal level is that it is not affected by tamping. However, at lower speeds the medium wavelength is not recorded and is not available for data
alignment. By analysing the measurement data, it was observed that before data alignment the positional errors were as much as 30 m, whereas after data alignment the positional errors were in a range of 3 m. Figure 3 shows an example of the cross-correlation analysis applied on measurement data in Optram to correct positional errors. Figure 3a shows the deviations in the longitudinal level over a track section recorded on two different inspection dates before data alignment. As can be seen, the measurements recorded in the second inspection do not represent the same positions as those recorded in the first inspection. By performing a cross-correlation analysis between the two measured waveforms, the position of the maximum value of the cross-correlation indicates sample leads or lags. Figure 3b shows cross-correlation between the two measured waveforms, $C_{21}$. By considering the location of the maximum value of the cross-correlation, it can be inferred that the location of the waveform measured in the inspection performed in February 2017 leads the location of the waveform measured in the inspection performed in February 2016 by 66 sample intervals. By considering the fact that the sampling interval for the IMV200 measurement car is 25 cm, there is a shift between the two measured waveforms corresponding approximately to $66 \times 0.25 = 16.5$ m.

Therefore, to deal with the position shifting error, the data must be aligned by shifting the measurements along the track position as shown in Figure 3c.

3.2 Identification of tamping cycles

In order to model the track geometry degradation, the time and position of the tamping interventions conducted on the track should be extracted to identify the maintenance cycles. Later this information will be used to model the evolution
of the amplitude of the longitudinal level defects within a maintenance cycle. The information about the time and position of the tamping actions performed on the track is extracted from the BIS\textsuperscript{3} database. BIS is an asset register database which contains information on infrastructure and facilities, agreements, and the history of tamping and grinding. However, in a number of cases the tamping interventions are not registered in the BIS system, and for these cases Trafikverket’s experts use certain criteria to identify the unregistered tamping actions. Figure 4 shows the tamping zone based on the standard deviation of longitudinal level before tamping and the tamping ratio for the identification of unregistered tamping interventions. Track sections with a standard deviation of longitudinal level less than 80\% of the UH1 limit are considered as good track sections, and track sections with a standard deviation of longitudinal level larger than 80\% of the UH1 limit are considered as poor track sections. For a good track section, a 15\% reduction in the standard deviation of longitudinal level is considered as a tamping intervention. For poor track sections, equation 1 is used to identify unregistered tamping interventions.

\[
TR = \frac{SDLL_a}{SDLL_b} < 0.9 - \frac{0.16}{SDLL_b} \tag{1}
\]

where TR is the tamping ratio, \(SDLL_a\) is the standard deviation of longitudinal level after tamping, and \(SDLL_b\) is the standard deviation of longitudinal level before tamping. A tamping intervention was conducted on the track section if the tamping ratio satisfies the inequality in formula (1).

By applying the mentioned criteria and using information from the BIS database, the positions and times of the tamping interventions conducted on line section 414 are identified. Figure 5 illustrates the variation of the standard deviation of longitudinal level over time for different track sections in line 414. By considering that a reduction in the evolution path of the standard deviation of longitudinal level represents a tamping intervention, Figure 5 can be used to identify the time and position of tamping interventions.

By considering the maintenance limits, Figure 5 can be used for an overall evaluation of the track geometry quality of the line section. Trafikverket (2015) stipulates that for the speed class of line section 414 the planning limit and the UH1 limit for the standard deviation of longitudinal level are 1.15 mm and 1.6 mm, respectively. Figure 5 shows that most of the track sections have a standard deviation of longitudinal level below the UH1 limit over the time period of the study. In addition, it can be seen that line tamping was not performed on line section 414 in the period of the study.

\textsuperscript{3} Baninformationssystem (in English: Track Information System)
4. Modelling the evolution of geometry defects

The track geometry parameter which usually drives the need for maintenance activities is the short-wavelength longitudinal level irregularities (UIC 2008). In order to model the track geometry degradation, the evolution of the amplitude of longitudinal level defects was studied.

4.1. Degradation model

The standard deviation of longitudinal level is widely used to represent the overall quality of railway track and to plan preventive maintenance actions. However, this aggregated indicator will not provide detailed information about the track condition and the severity of isolated defects in a track section. Therefore, there is a need to monitor and analyse the changes in the amplitude of isolated defects over time, to prevent the occurrence of UH2 defects in a track section.

Figure 6 is a heat map of longitudinal level measurements at each sampling interval (every 25 cm) over time and within a given track section. Figure 6 can be used to observe the evolution of longitudinal level defects over time and to find the positions prone to the occurrence of UH2 defects. As can be seen, there are a few positions for which an increasing trend in the amplitude of the defects can be observed. These positions are prone to the occurrence of UH2 defects and must be considered for further analysis.

In order to detect the positions prone to the occurrence of UH2 defects, the six measurements with the highest values and the six measurements with the lowest values are recorded in Optram in each inspection run for each section and for both the right rail and the left rail. Those defects which have exceeded the planning limit may turn into UH2 defects in a short period of time and must be considered carefully for planning maintenance activities. Therefore, the changes in the amplitude of these defects over time must be analysed. The trend of the changes in the amplitude of the defects can be used to predict when a defect which has exceeded the planning limit will turn into a UH2 defect. The positions of defects are checked in each inspection run and defects within a distance of 3 m from each other are considered as belonging to the same defect. Figure 7 shows the evolution of the amplitude of the two longitudinal level defects at the positions 59.050 km and 59.074 km. In order to identify the evolution pattern of the amplitude of the defects in line section 414, the same type of plot as that used for Figure 7 was used to provide plots for all the longitudinal level defects in this case study.
By analysing the changes in the amplitude of the longitudinal level defects, it was observed that the degradation within a maintenance cycle had a linear pattern. Therefore, the linear regression model presented in equation 2 was applied to model the evolution of isolated longitudinal level defects:

\[
A(t) = A_0 + \beta(t - t_{\text{tamp}}) + \varepsilon 
\]  

(2)
where $A(t)$ is the absolute value of the amplitude of the longitudinal level defect in time $t$, $A_0$ is the absolute amplitude of longitudinal level defect after the latest tamping intervention, $\beta$ is the degradation rate, and $t_{\text{tamp}}$ is the latest tamping time, see Figure 8. Finally, $\varepsilon$ is the Gaussian random error term with a mean value equal to zero:

$$\varepsilon \sim N(0, \delta^2)$$

(3)

By using the measurement data recorded by the inspection cars, the parameters of the proposed model are estimated. When the amplitude of a defect exceeds the planning limit, the proposed model can be used to predict the time of the occurrence of a UH2 defect. In order to check the normality assumption for the residuals of the simple linear model, the Kolmogorov-Smirnov (KS) test is applied. The results of the KS test applied in the present study are summarized in Figure 9 in a histogram of the p-values. As can be seen, all the p-values of the KS test were larger than the significance level (0.05). Consequently, it could be concluded that the normality assumption for the residuals of the simple linear model was suitable for our case study.

Figure 9. Histogram of the p-values for the normality of the residuals using the KS test

Figure 10 illustrates the distribution of the degradation rates with a histogram of the degradation rates. It can be observed in this figure that most of the defects have a very small degradation rate, but the tail of the histogram indicates that there are a number of defects with a high degradation rate. These defects must be monitored and analysed as they may turn into UH2 defects in a short period of time.
In order to analyse the effect of the presence of a special asset in a track section, e.g. a switch and crossing, on the degradation rates, a box plot was created, see Figure 11. Figure 11 shows that defects located in a track section with a special asset have a higher degradation rate on average. Therefore, when an isolated defect exceeds the planning limit in these sections, special consideration should be given to them.

4.2. Effect of shock events on the degradation path

By analysing the degradation path of geometry defects in line section 414, it is observed that in addition to gradual degradation, there can be an abrupt change in the degradation path in which the degradation level dramatically increases over time. We call this phenomenon a shock event. Although the occurrence of an abrupt change in the degradation path is not common and may only occur in a small number of sections, considering them when modelling track geometry degradation is very important (Soleimanmeigouni, Xiao, Ahmadi, Xie, Nissen, & Kumar, 2018). Track sections with an unusual trend in the degradation path normally have a significantly higher degradation rate after a shock event, causing a shorter maintenance cycle compared to that of track sections with a normal degradation path.
Shock events in the degradation path may cause safety problems and, in the worst-case scenario, lead to derailment. To analyse the occurrence of shock events in the degradation path, the changes in the evolution of the standard deviation of the longitudinal level and amplitude of the longitudinal level defects over time are monitored and analysed. Figure 12a shows the changes in the standard deviation of longitudinal level for a track section with a shock event. When there is a shock event in the degradation path of the standard deviation of longitudinal level, there is also a change point in the degradation path of the isolated defect (see Figure 12b).

Figure 12. Irregular trend in the track geometry degradation path

Figure 13 illustrates in detail the occurrence of a shock event by showing the longitudinal level measurements obtained in the 5th, 6th and 7th inspection of the track section concerned in Figure 12. As shown in Figure 13, there is a dramatic increase in the amplitude of the two isolated defects. The defect amplitude was below the planning limit in the 5th inspection, while it turned into a UH2 defect in the 7th inspection. Therefore, it can be concluded that the reason for the unusual trend observed in the evolution of the standard deviation of the longitudinal level was a problem that happened at the location within the track section marked with red cycle in Figure 13.

Shock events in the degradation path can occur for different reasons. For a number of track sections that have a change point in the degradation path, it has been observed that the increase in the defect amplitude occurred at exactly the same position as where a sleeper was replaced or a steel drum was installed. If after a sleeper replacement the ballast is not compacted well, this causes a change point in the track geometry degradation path. In addition, an inadequate installation or replacement of steel drums may also cause some damage to the track geometry condition and result in a change point in the degradation path. In the event of a shock event occurring in the degradation path, the regression
model presented in equation (2) should be slightly modified to model the track geometry degradation in the way presented in equation (4):

\[ A(t) = A_0 + \beta(t - t_s) + \varepsilon \] (4)

where \( t_s \) is the time of the occurrence of the shock event. Since after a shock event there is a new degradation pattern with a higher degradation rate, one must only use the data recorded after the shock event to model the evolution of the amplitude of the defects. After identifying a change point in the degradation path, the linear model expressed in equation (4) can be used to predict the time of the occurrence of UH2 defects.

Figure 13. Unusual growth of the amplitude of isolated longitudinal level defects

4.3. Effect of tamping on isolated defects

The effect of tamping on the degradation of defects whose amplitude has exceeded the planning limit can be seen by a sudden change in their amplitude after a tamping intervention, as is indicated by the vertical black dashed line in Figure 14. This figure illustrates in detail the effect of tamping on the amplitude of isolated defects by showing the longitudinal level waveform of a given track section before and after a tamping intervention. Figures 14a, 14b and 14c represent the longitudinal level waveform before the tamping intervention, one month after the tamping intervention, and one year after the tamping intervention, respectively. As can be seen in Figure 14a, there is a UH2 defect in the track section before the tamping intervention. Figure 14b shows that the tamping intervention rectified the UH2 defect and all the measurements are within the planning limits one month after the tamping intervention. As can be seen in Figure 14c, although the tamping rectified the UH2 defect, after one year a UH1 defect occurred at the same position.

This issue was analysed for all the track sections in line section 414 and it was observed that for around 35% of the sections with a UH1 defect or UH2 defect, after a tamping intervention a UH1 defect or UH2 defect occurred again at the same position. This shows that any rectification of isolated defects through the current maintenance practice is not durable. In fact, correcting isolated defects using tractors and lightweight machines or tamping machines cannot remove the root causes of defects. Many problems in the track can be considered as root causes of isolated defects, e.g. broken sleepers, track substructure problems and drainage problems.
Figure 14. Evolution of longitudinal level irregularities for a single track section

5. Section-based model

The proposed degradation model can be used to predict the time of the occurrence of UH2 defects. However, generally railway infrastructure managers prefer to plan maintenance activities based on overall condition of track sections rather than isolated defects. Moreover, generally TQIs based on the standard deviation of longitudinal level is used for planning track geometry maintenance activities. Therefore, we developed a section-based model which considers the standard deviation and kurtosis of longitudinal level to predict the probability of the occurrence of UH2 defects in a given period of time.

5.1. Binary logistic regression

Regression is a well-known statistical method in data analysis for describing the relationship between a response variable and a set of explanatory variables (Hosmer & Lemeshow 2000). However, in many cases the response variable is binary or dichotomous. Modelling a binary response variable using a linear regression which assumes that the response is normally distributed will result in biased parameter estimates. In the case that the response is a binary variable the logistic regression which links the probability of a binomial distribution to the explanatory variables after a suitable transformation has been used as a standard method. By considering \( Y \) as the binary response variable, the aim is to model the conditional probability \( P(Y = 1|x) = \pi(x) \) as a function of a set of explanatory variables. In logistic regression the logit transformation of the probability \( \pi(x) \) is modelled as the linear function of the explanatory variables, as in equation 5:

\[
g(x) = \logit(\pi(x)) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{5}
\]

where \( n \) is the number of explanatory variables, \( x = (x_1, \ldots, x_n) \) are explanatory variables, \( \beta_0, \beta_1, \ldots, \beta_n \) are the model coefficients, the logit denotes the logit transformation, \( \frac{\pi(x)}{1-\pi(x)} \) is the odds of \( Y \), and \( g(x) \) is the linear function associated
to the logit model. By solving the logit equation with respect to $\pi(\bar{x})$, the probability of the occurrence of an event $P(Y = 1|\bar{x})$ can be obtained using equation 6:

$$P(Y = 1|\bar{x}) = \pi(\bar{x}) = \frac{e^{g(\bar{x})}}{1 + e^{g(\bar{x})}}$$  \hspace{1cm} (6)

In order to estimate the coefficients of the model, the maximum likelihood method can be applied. Given $m$ observations as $(x_j, y_j), j = 1, \ldots, m$, the likelihood function is given as equation 7:

$$L(\alpha, \beta_1, \ldots, \beta_n) = \prod_{j=1}^{m} \pi(x_j)^{y_j}(1 - \pi(x_j))^{1-y_j}$$ \hspace{1cm} (7)

The estimation of the model coefficients $\beta_0, \beta_1, \ldots, \beta_n$ will be obtained by maximization of the likelihood function or equivalently the log-likelihood function.

5.2. Relationship between the standard deviation and kurtosis of longitudinal level and the occurrence of UH2 defects

The standard deviation is used to show the variation or dispersion of track geometry measurement data. A low standard deviation indicates that the geometry measurements are close to the mean and a high standard deviation indicates that the geometry measurements have a high variation around the mean value. Since the presence of extreme values in the geometry measurements is of importance, then kurtosis can provide useful information. Kurtosis is a measure of tailedness of data. Higher kurtosis is the result of infrequent extreme observations (or outliers), as opposed to frequent moderate deviations. Therefore, it is expected that a track section where most of the geometry measurements are low, but which contains a UH2 defect, will have a high kurtosis. Figure 15 presents the relationship between the standard deviation and the kurtosis of the longitudinal level with the presence of a UH2 defect in the track section. The left panel of Figure 15 shows the waveforms of the longitudinal level for three track sections in the same kilometre of line section 414. The right panel of the figure shows the estimated density of the longitudinal level measurements of the track sections obtained with a kernel smoothing function.

![Figure 15. Relationship between the standard deviation and the kurtosis of the longitudinal level with the presence of a UH2 defect in the track section](image-url)
In both track sections "a" and "b" there is a UH2 defect. However, in section "a" the longitudinal levels of most of the sample points are below the planning limit, whereas in section “b” there are a number of defects which have exceeded the planning limit. As a result, section "a" has a smaller standard deviation than section "b", but has a higher kurtosis than section "b". This point is clear from the density function of the two waveforms in that the density function of section "a" has a sharper peak and longer tails than that of section "b". When studying the longitudinal level waveform of section “c” in Figure 15, one can observe that all the longitudinal level measurements are below the planning limit and there is no UH2 defect in that section. As expected, the density function for this section has a sharper peak and shorter tails than that for the other sections, which indicates that section “c” has a smaller kurtosis than the other two. In addition, the density functions of sections “a” and “b” are flatter than the density function of section "c", which indicates that section "c" has a smaller standard deviation.

In order to find out how standard deviation and kurtosis of longitudinal level are related to the occurrence of UH2 defects, the binary logistic regression was applied. The response variable takes a value of 1 with the occurrence of at least one UH2 defect, and otherwise it takes a value of 0. The results for the fitting of the model are summarized in Table 4.

Table 4. Results of estimation of binary logistic regression for the longitudinal level defects

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Standard deviation of longitudinal level</th>
<th>Kurtosis of longitudinal level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>1191</td>
<td>520</td>
<td>769</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt;0.005</td>
<td>&lt;0.005</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Coefficient (β)</td>
<td>-13.77</td>
<td>5.82</td>
<td>0.59</td>
</tr>
</tbody>
</table>

In this study, the statistical significance of the individual regression coefficients was tested using the chi-square statistic. As can be seen in Table 4, the p-values of all the predictors are less than the significance level. Therefore, it can be concluded that all the variables in the model have a significant effect on the probability of the occurrence of UH2 defects.

Figure 16. Relationship between the standard deviation and kurtosis of longitudinal level and the occurrence of UH2 defects in the whole of line section 414

The Hosmer-Lemeshow (H-L) test was applied to assess the fit of the applied logistic regression model against actual outputs. A p-value for the H-L test less than the significance level (p-value <0.05) indicates a poor fit to the data and leads to the conclusion that the predicted probabilities from the logistic regression model deviate from the observed
proportion of the events. The value obtained for the H-L goodness-of-fit statistic was 10.63 and the corresponding p-value from the chi-square distribution with 8 degrees of freedom was 0.23. Therefore, it can be inferred that the proposed model was properly fitted.

According to Table 4, both the standard deviation and the kurtosis of longitudinal level are statistically significant and have positive coefficients. This means that the higher the standard deviation and kurtosis of longitudinal level is, the higher is the probability of the occurrence of UH2 defects. This finding is illustrated in Figure 16, which shows the relationship between the standard deviation and kurtosis of longitudinal level measurements and the presence of UH2 defects in the whole of line section 414.

Figure 16 shows that when the standard deviation or the kurtosis of longitudinal level is low, there is no UH2 defect in the track section in question. Moreover, whenever the standard deviation of the longitudinal level is higher than the UH1 limit and the kurtosis is low (close to zero), a very small number of UH2 defects have occurred. Similarly, when the kurtosis is high and the standard deviation is low, a small number of UH2 defects have occurred. Therefore, both the standard deviation and the kurtosis of the longitudinal level must be higher than some specified value for a UH2 defect to be occurring in the track section. Therefore, considering both the standard deviation and the kurtosis as the explanatory variables in predicting the occurrence of UH2 defects may increase the prediction accuracy.

5.3. Prediction of the occurrence of UH2 defects

In order to predict the probability of the occurrence of UH2 defects in a track section in a given period of time, binary logistic regression was applied to develop the section-based model. The response variable $Y$ takes a value of 1 whenever there is at least one UH2 defect in the track section and a value of 0 whenever the section only contains defects which have exceeded the planning limit or the UH1 limit. In the section-based model, we predict the probability that a section which contains defects which have exceeded the planning limit or the UH1 limit will turn into a section containing at least one UH2 defect in a given period of time. The explanatory variables considered in the model are the following four variables: (1) the standard deviation of longitudinal level, (2) the kurtosis of longitudinal level, (3) the presence of defects which exceeded the planning limit or the UH1 limit in the latest measurement, and (4) the time interval (in years). The binary logistic regression for the section-based model is presented in equation 8:

$$P(Y = 1) = \frac{e^{\beta_0 + \beta_1 SDLL + \beta_2 \gamma_2 + \beta_3 \Delta t + \beta_4 c}}{1 + e^{\beta_0 + \beta_1 SDLL + \beta_2 \gamma_2 + \beta_3 \Delta t + \beta_4 c}}$$

(8)

where $SDLL$ is the standard deviation of longitudinal level, $\gamma_2$ is the kurtosis excess of longitudinal level, $\Delta t$ is the time interval, and $c$ is a categorical variable which takes a value of 0 if the section only contains a defect which exceeded the planning limit in the latest measurement and a value of 1 if the section also contains at least one UH1 defect. Maximum likelihood estimation method is used to estimate the model parameters. The key results for the fitting of binary logistic regression to the data are summarized in Table 5.

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>SDLL</th>
<th>$\gamma_2$</th>
<th>$\Delta t$</th>
<th>$c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>415.4</td>
<td>95.9</td>
<td>156.4</td>
<td>5.3</td>
<td>74.1</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt;0.005</td>
<td>&lt;0.005</td>
<td>&lt;0.005</td>
<td>0.022</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Coefficient ($\beta$)</td>
<td>-12.68</td>
<td>3.49</td>
<td>0.38</td>
<td>3.59</td>
<td>3.315</td>
</tr>
</tbody>
</table>

As can be seen in Table 5, the p-values of all the predictors are less than the significance level. Therefore, it can be concluded that all the variables in the model have a significant effect on the probability of the occurrence of UH2 defects. The value obtained for the H-L goodness-of-fit statistic was 7.21 and the corresponding p-value from the chi-square distribution with 8 degrees of freedom was 0.514. Therefore, it can be inferred that the proposed model is properly fitted. According to Table 5, the coefficient of time is positive, which means that the probability of the
occurrence of a UH2 defect in a track section is higher in a longer period of time. In addition, the coefficient of the
categorical variable \( c \) is positive. This means that when a track section has been found to have a UH1 defect in the
latest measurement, the probability of the occurrence of a UH2 defect in the time period \( \Delta t \) is higher than when the
section only contains a defect which has exceeded the planning limit. In addition, the standard deviation and kurtosis
of longitudinal level have a positive coefficient. This means that the higher the standard deviation and kurtosis are, the
higher is the probability of the occurrence of a UH2 defect in the time period \( \Delta t \).

The probabilities predicted using binary logistic regression can be used to classify the outputs by using a single cut-
point (\( \omega \)) to compare each estimated probability with respect to \( \tau \). If the estimated probabilities are greater than \( \omega \), the
predicted class will be equal to 1, \( \hat{Y} = 1 \), and otherwise the predicted class will be equal to 0, \( \hat{Y} = 0 \). In order to
establish the final prediction we estimated an optimal cut point value using the sensitivity and specificity. When
classification is the main goal of the analysis, the sensitivity and specificity can be used to assess the model
performance (Hosmer & Lemeshow, 2000). The sensitivity and specificity are statistical measures of the performance
of a binary classification test. The sensitivity or true positive rate (TPR) measures the proportion of correctly classified
events. The specificity or true negative rate (TNR) measures the proportion of correctly classified non-events (Peng,
Lee, & Ingersoll, 2002).

\[
\text{Sensitivity} = \frac{\text{Number of true positives}}{\text{Number of true positives} + \text{Number of false negatives}}
\]

\[
\text{Specificity} = \frac{\text{Number of true negatives}}{\text{Number of false positives} + \text{Number of true negatives}}
\]

In order to test the performance of the model using the available data, 70% of the data were used as a training set and
30% of the data were used as a test set. Sensitivity and specificity are dependent on the value of cut-point. Figure 17
show the changes in sensitivity and specificity with respect to different cut-point values. Particularly, the sensitivity is
monotonic decreasing in \( \omega \) while the specificity is monotonic increasing in \( \omega \) as shown in Figure 17. Therefore, it is
impossible to maximize both of sensitivity and specificity simultaneously. In order to select the optimal cut-point, one
should optimize a suitable measure which balances the sensitivity and specificity well. For example, the arithmetic
mean of sensitivity and specificity is the simplest measure. One can also use a weighted average by incorporating
different costs related to false positives and false negatives. Another widely used measure is the F-score which is the
harmonic mean of sensitivity and specificity. In this regard, the sensitivity and specificity for different cut-points value
are calculated and we select the cut-point which balances both of them, i.e. the sensitivity and specificity are equal, to
classify data.

![Figure 17. Plot of the sensitivity and specificity for different cut-point values](image)

As can be seen in Figure 17, the optimal value for the cut-point should be \( \omega^* = 0.23 \), as this is the value at which the
sensitivity and specificity curves cross each other. By considering \( \omega^* \) as the optimal cut-point, the model sensitivity
and specificity are 89%, which seems reasonable for this case study.
6. Conclusions

The aim of this study has been to develop a data-driven analytical methodology for the prediction of track geometry defects by performing an extensive case study on line section 414 of the Main Western Line in Sweden. In this study, particular emphasis has been placed on the prediction of UH2 defects, which entail great costs for the maintenance of railway tracks. In order to identify the degradation pattern of isolated defects, a detailed analysis on foot-by-foot track geometry measurement data has been performed. It is found that isolated longitudinal level defects have a linear degradation pattern. The modelling methodology considered the occurrence of shock events and their associated change points in the degradation path. Shock events in the degradation path may cause safety problems and, in the worst-case scenario, lead to derailment. In addition, the effectiveness of tamping intervention in rectifying the longitudinal level defects was analysed. It was observed that for around 35% of the sections with a UH1 defect or UH2 defect, after a tamping intervention a UH1 defect or UH2 defect occurred again at the same position. This shows that in many cases, spot tamping using tractors and lightweight machines or tamping machines cannot remove the root causes of defects. In order to predict the probability of the occurrence of UH2 failures in a track section, binary logistic regression is applied in the model developed in this study. Four variables, namely the standard deviation and kurtosis of the longitudinal level, the time interval and the presence of defects in the latest measurement, are selected as the explanatory variables in the model. The results show that the probability of the occurrence of UH2 defects in a track section is linked to the standard deviation and kurtosis of the track section. Therefore, it can be concluded that the aggregated TQIs which are statistically significant predictors for the occurrence of UH2 defects. Using the proposed section-based model railway companies can predict which track section will need a maintenance due to the occurrence of a severe isolated defect. This will enable the railway companies to perform section-wise preventive maintenance. In addition, it is found that the kurtosis, which is a measure of the tailedness of the data, can be used efficiently to capture the information about the occurrence of UH2 defects. This is important because it affects the design and interpretation of future studies on track quality indices. The applied binary logistic regression model shows a satisfactory performance in predicting the occurrence of UH2 defects, which further justifies our approach. One possible future research direction would be to explore the prediction of other defect types, e.g. twist and alignment defects. Another possible direction would be to include more explanatory variables, e.g. the soil type, traffic information and the track type, in the logistic regression model to improve the prediction accuracy. Finally, a further research direction would be to combine the predictions produced through the degradation model and the section-based model to establish an integrated framework for efficient maintenance planning.

Acknowledgements

The authors would like to thank Trafikverket, SIMTRACK project, and Luleå Railway Research Center (JVTC) for their technical and financial support during this project.

References


Investigation of the effect of the inspection intervals on the track geometry condition

Investigation of the effect of the inspection intervals on the track geometry condition

Iman Soleimanmeigouni1, Alireza Ahmadi1, Hamid Khajehei1, Arne Nissen2

1Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden
2Trafikverket, Luleå, Sweden

Abstract

In order to evaluate the railway track geometry condition and plan maintenance activities, track inspection cars run over the track at specific times to monitor it and record geometry measurements. Applying an adequate inspection interval is vital to ensure the availability, safety and quality of the railway track, at the lowest possible cost. The aim of this study has been to investigate the effect of different inspection intervals on the track geometry condition. To achieve this, an integrated statistical model was developed to predict the track geometry condition given different inspection intervals. In order to model the evolution of the track geometry condition, a piecewise exponential model was used which considers break points at the maintenance times. Ordinal logistic regression was applied to model the probability of the occurrence of severe isolated defects. The Monte Carlo technique was used to simulate the track geometry behaviour given different inspection intervals. The results of the proposed model support the decision making process regarding the selection of the most adequate inspection interval. The applicability of the model was tested in a case study on the Main Western Line in Sweden.

Keywords: Inspection, track geometry, maintenance, multivariable linear regression, isolated defects, ordinal logistic regression

1. Introduction

The railway track geometry is subjected to degradation mainly due to the forces induced by traffic, which cause deviations from the designed vertical and horizontal alignments. When a geometry defect occurs, the consequences can be significant, including economic loss, damage to the environment and, in severe cases, the possible loss of human lives. In order to prevent against unacceptable consequences, track geometry maintenance actions are planned to retain the geometry condition in or restore it to an acceptable state with respect to ride comfort and the safety limits. In order to have an effective and efficient maintenance plan, the track geometry condition must be inspected on a regular basis. The frequency of inspection is highly important as it affects the detectability of geometry defects and the associated maintenance requirements. As a result, a proper inspection interval is a key factor for determining the performance of line sections, controlling and reducing the risk of derailment and the maintenance cost, and keeping the punctuality of train operations at the highest level. Various factors are involved in the selection of an applicable and effective inspection interval, including the traffic profile, the regulatory requirements, the level of risk in train operations, and the maximum allowable speed of trains (American Railway Engineering and Maintenance-of-Way Association (AREMA), 2006).

In recent years, a number of researchers have tried to determine an optimal inspection interval with respect to different objective functions. Arasteh Khouy, Larsson-Kräik, Nissen, Junitti, and Schunnnesson (2014) optimized the track geometry inspection intervals with the objective of minimizing the total maintenance cost. In the proposed model, preventive and corrective tamping is executed based on the Q-value indicator and severe defects of twist, respectively. A similar approach can be found in Soleimanmeigouni, Ahmadi, Letot, Nissen, and Kumar (2016), who attempted to determine a cost-effective track geometry inspection interval. They considered the standard deviation of the longitudinal level as the dominant factor for evaluation of the track condition. They used the Wiener process to model the track geometry degradation. Lyngby, Hokstad, and Vatn (2008) conducted research to optimize the intervals of track geometry inspection with the same objective. They applied the Markov methodology to model the track geometry degradation. In their study, twist was considered as a track quality indicator. In the model proposed by Lyngby et al. (2008), the corrective and preventive activities are assumed to be perfect. Meier-Hirmer, Sourget, and Roussignol (2005) conducted a case study and determined a trade-off between the inspection interval and the maintenance threshold that leads to the minimum track maintenance cost. The track geometry degradation and tamping effectiveness were modelled using the gamma process and linear regression, respectively. In this study, the maintenance was carried...
out on track sections with a fixed delay after detection. More recently, Osman, Kaewunruen, Jack, and Sussman (2016) conducted research on the possibility of considering a ‘plan B’ or contingency plan in track inspection schedules. This ‘plan B’ would reschedule the inspection plan in the event of an incident happening which would affect the current inspection schedule. A different approach can be seen in a study presented by Andrews, Prescott, and De Rozieres (2014), who did not include an objective function in their work. They studied the effect of different inspection strategies on the quality of the track geometry. The standard deviation of the longitudinal level was used as a measurement of the track geometry quality. They reported that, despite their expectations, different inspection intervals did not have a significant effect on the total number of maintenance actions. However, they ascertained that the longer the inspection intervals were, the larger was the proportion of maintenance which had to be carried out on track with poorer quality. Moreover, Andrews et al. (2014) pointed out that a decrease in the frequency of inspections would increase the percentage of time during which the track would need emergency maintenance.

Obviously, an effective inspection regime can be developed according to an assessment of the effect of different inspection intervals on the track performance. The performance of the track can be measured by determining the percentage of time spent by the track sections in different geometry states. These states can be defined based on the maintenance limits. The aim of the present study has been to develop an integrated approach to investigate the effect of different inspection intervals on the performance of the track. A fundamental requirement for achieving this aim is the long-term prediction of the track geometry condition by modelling and integrating: (1) the track geometry degradation; (2) the tamping recovery; (3) the occurrence of severe isolated defects. In this study, to characterize the track geometry degradation and restoration, a piecewise exponential model was applied which considers break points at the maintenance times. A multivariable linear regression model was used to link different covariates to the tamping recovery. In addition, ordinal logistic regression was applied to predict the probability of the occurrence of isolated defects. Due to the existence of variation in the model parameters, the Monte Carlo technique was used to estimate the percentage of time spent in different track geometry states. In order to verify the applicability of the model, a case study was performed with data collected from the Main Western Line in Sweden. The rest of this paper is organized as follows. Section 2 contains background information on the track geometry parameters and maintenance limits. Section 3 describes the process of track geometry measurement. Section 4 deals with the analytical models and the proposed approach. The case study is presented in Section 5 and, finally, Section 6 provides the conclusions.

2. Track geometry parameters and maintenance limits

Track geometry parameters are widely used to represent the track condition and to plan maintenance activities. Track geometry parameters can be divided into five classes, i.e. the longitudinal level, alignment, gauge, cant, and twist. The longitudinal level is the geometry of the track centreline projected onto the longitudinal vertical plane. The alignment is the geometry of the track centreline projected onto the longitudinal horizontal plane. The gauge is the distance between the gauge faces of two adjacent rails at a given location below the running surface. The cant (cross-level) is the difference in height between the adjacent running tables computed from the angle between the running surface and a horizontal reference plane. The twist is the algebraic difference between two cross-levels taken at a defined distance apart, usually expressed as a gradient between the two points of measurement (SS-EN 13848-1: 2004+A1, 2008).

According to EN 13848-5 (2008), there are three limits for maintenance actions. The immediate action limit (IAL) or safety limit refers to the value which, if exceeded, due to the potential risk of derailment, requires that a speed reduction or line closure be imposed before a corrective maintenance (CM) action is conducted. The intervention limit (IL) or CM limit refers to the value which, if exceeded, requires a CM action before the immediate action limit is reached. The alert limit (AL) or preventive maintenance (PM) limit refers to the value which, if exceeded, requires that the track geometry be analysed for the planning of future maintenance actions. The European standard EN 13848-5 (2008) provides the IALs, ILs and ALs for isolated defects and gives ALs for standard deviations. Generally, the track quality indices (TQIs) based on the standard deviation of track geometry parameters are used to plan and perform PM actions. On the other hand, the execution of CM actions is based on the severity of isolated defects. Whenever the amplitude of an isolated defect exceeds the IL or IAL, CM should be conducted on the track. Figure 1 shows the different maintenance zones based on the above-mentioned limits.

The IALs are normative and take into account the track-vehicle interaction and the risk of unexpected events, whereas the ILs and ALs are informative and are mainly linked with the maintenance policy. In fact, the ILs and ALs provided in EN 13848-5 (2008) reflect the common practice adopted by most European infrastructure managers. In alignment with the European standard EN 13848-5 (2008), Trafikverket (2015) has defined four main limits, namely the planning limit, the UH1 and the UH2 limit, and the critical limit, as can be seen in Figure 2. In Trafikverket (2015) the
intervention limit is expressed as a range rather than as a discrete value. Track irregularities that exceed the UH1 limit must be assessed for conducting maintenance before the UH2 limit is exceeded. For track irregularities exceeding the UH2 limit, a maintenance action must be planned without unnecessary delay. Therefore, track irregularities must be corrected before the UH2 limit is reached.

Table 1 shows the relation between the limits defined in EN 13848-5 (2008) and those defined in Trafikverket (2015).

<table>
<thead>
<tr>
<th>Alert limit</th>
<th>Planning limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention limit</td>
<td>UH1 limit (lower bound for corrective maintenance)</td>
</tr>
<tr>
<td></td>
<td>UH2 limit (upper bound for corrective maintenance)</td>
</tr>
<tr>
<td>Immediate action limit</td>
<td>Critical limit</td>
</tr>
</tbody>
</table>

Based on the above-mentioned limits, geometry defects can be classified according to their severity into three groups, i.e. UH1, UH2 and critical defects. UH1, UH2 and critical defects occur when track irregularities exceed the UH1, UH2 and critical limits, respectively. Examples of these three categories of defects are shown in Figure 2.
3. Track geometry inspection

The main objectives of measuring the track geometry are to: (1) locate the occurrence of severe isolated defects and prioritize the CM actions required to rectify them, (2) to plan the PM actions, (3) to evaluate the effectiveness and quality of conducted maintenance activities, and (4) to establish the effectiveness and adequacy of financial expenditure (Zaayman, 2017). The track geometry measurement car is an automated track inspection vehicle whose purpose is to measure and record the track geometry in real time and to display the measurements obtained. The vehicle measures various geometry parameters, e.g. the longitudinal level, alignment, cant, gauge, twist, curve radius, and gradient at a specific sampling interval. The measured data are available in three forms, i.e. on-board network viewing of measurements, on-board real-time reports, and off-board post-processed reports. The on-board real-time reports can be used to address severe isolated defects and comprise strip charts and exceptions reports. The strip chart is a graphical representation of the overall track geometry condition in one picture and displays all the track geometry measurements and the corresponding maintenance limits. The strip chart shows when a particular measurement approaches or exceeds the maintenance limits. When the track geometry measurements exceed the maintenance limits, a geometry defect will appear in the exceptions report. The purpose of the exceptions report is to provide a list of all the geometry defects found by the track geometry measurement vehicle during the inspection of the track. The basic information provided in the exceptions report is the defect name, the position of the defect, and the defect length, defect magnitude, maintenance limit, and maximum allowable speed. Using the strip charts and the exceptions reports aids the process of conducting CM activities to avoid the occurrence of severe isolated defects which may cause catastrophic consequences. In addition to on-board reports, the geometry measurements are post-processed into different standard reports. These reports can be used for further analysis of the track geometry condition to support track maintenance decision making to achieve a robust and cost-effective maintenance plan. Analysing the track geometry measurements and aggregating them into a track quality index (TQI) can aid the planning of track geometry maintenance activities. By performing a statistical analysis on the evolution of the TQIs, the trends in the track geometry degradation can be identified. In addition, the measured data can be used to identify potentially hazardous track geometry conditions.

4. Proposed methodology

Figure 3 provides a schematic description of the proposed analytical model. As can be seen, the track geometry is inspected at discrete time intervals ($\tau$) to determine its condition. At each inspection, the standard deviation of the geometry parameters and the extreme values of isolated defects are recorded. According to the results of the track geometry inspection, different intervention decisions can be made to retain the geometry condition in or to restore it to an acceptable state. With reference to the maintenance limits explained in Section 2, after each inspection, one of the following three actions may be taken.

- **Preventive maintenance**: If the standard deviation of the geometry parameters exceeds the planning limit, then the track section is assessed for PM to be performed in the first available maintenance window. By scheduling the PM for the time periods planned for maintenance interventions, the maintenance will be performed in a way which will neither interrupt the normal traffic nor cause traffic delays.

- **Normal corrective maintenance (CM$_n$)**: When a UH2 defect occurs, a CM action is conducted on the track section. Railway companies need short-term plans for adjusting the traffic by cancelling, postponing or rerouting trains to provide time to conduct CM actions. Such short-term plans provide a period of time during which the track can be operated until maintenance takes place.

- **Emergency corrective maintenance (CM$_e$)**: When a critical defect occurs, an immediate CM action with a speed reduction or line closure is carried out on the track section.

Figure 3 depicts the three above-mentioned situations.

The aim of the proposed analytical model is to predict the percentage of time which the track spends in the different track geometry states using various inspection intervals. With reference to the maintenance limits explained in Section 2, Table 2 presents the different geometry states.

Owing to the complexity of the degradation process and the maintenance process, predicting the effect of employing different inspection intervals on the track geometry condition is a difficult task (Andrews et al., 2014; Quiroga & Schnieder, 2012). Therefore, the Monte Carlo simulation technique is used to handle the variation of the various
parameters within the proposed integrated model and to estimate the percentage of time spent in the different track geometry states.

![Figure 3. Schematic description of preventive and corrective maintenance actions](image)

Table 2. Three defined track geometry states based on the maintenance limits

<table>
<thead>
<tr>
<th>State</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>There is no UH2 defect or critical defect in the track section.</td>
</tr>
<tr>
<td>2</td>
<td>There is at least one UH2 defect in the track section, but there is no critical defect in that section.</td>
</tr>
<tr>
<td>3</td>
<td>There is at least one critical defect in the track section.</td>
</tr>
</tbody>
</table>

4.1. Track geometry degradation and restoration

In order to model the evolution of the geometry condition of a track line over time, three main challenges must be addressed: (1) modelling the track geometry degradation within a maintenance cycle, (2) modelling the effect of tamping on the geometry degradation pattern, and (3) modelling the section-to-section variation in the degradation parameters. In order to model the track geometry evolution, the first level of the framework proposed by Soleimanmeigouni, Xiao, Ahmadi, Xie, Nissen, and Kumar (2018) was adapted.

Exponential model, with time or tonnage as the explanatory variable, is widely used in the literature to model track geometry degradation within a maintenance cycle (Soleimanmeigouni, Ahmadi, & Kumar, 2018). The main advantage of exponential model lies in its simplicity and ability to represent the underlying geometry degradation path. In this study, in order to model track geometry degradation of the section $s$ within a maintenance cycle an exponential model with a multiplicative lognormal error term is used (see equation 1):

$$D_s(t) = D_{s,0} e^{eta_s t} e_{s}$$

Where

- $D_s(t)$ is the degradation value at time $t \in T$
- $D_{s,0} > 0$ is the initial degradation level
- $\beta_s > 0$ is the degradation rate
- $e_{s}$ is the random error term following a lognormal distribution with the location parameter zero and the scale parameter $\sigma_s$.

When the degradation characteristic reaches the maintenance threshold, tamping action will be scheduled to restore the track geometry condition. The geometry degradation path always has a break point after a tamping action. Although tamping actions will improve the track geometry condition, they cannot rejuvenate the geometry condition to an as-good-as-new state. Therefore, in order to predict the track geometry evolution in multiple maintenance cycles, the
The effect of tamping on track geometry degradation must be assessed. The recovery value after tamping intervention is a function of different covariates, e.g., degradation value before tamping and number of accumulated tamping interventions. In this study, multivariable linear regression was used to explain the relationship between the covariates and the recovery value after tamping, see equation 2.

\[ R_{js} = a + b_0 D_{s\left(t_{m_{j,s}}\right)} + \sum_{v=1}^{u} b_v x_{v,j,s} + \delta \]  

where

- \( D_{s\left(t_{m_{j,s}}\right)} \) is the degradation value just before the \( j \)th tamping
- \( \{x_{v,j,s}, v = 1, \ldots, u\} \) are covariates for the \( j \)th tamping of the \( s \)th section
- \( a, b_v, v = 1, \ldots, u \) are regression coefficients
- \( \delta \sim N(0, \tau^2) \) is a Gaussian random error term.

For a specific section, the degradation value before tamping is the critical covariate used to predict recovery after tamping. Therefore, in equation 2, the coefficient \( b_0 \) is listed separately to emphasize its importance.

By integrating the degradation model within a maintenance cycle and the recovery model, a piecewise exponential model was applied to characterize the degradation in multiple maintenance cycles (see equation 3).

\[ D_s(t) = \frac{D_{s,0}}{\prod_{j=1}^{k_s} \delta_{j,s}(t-t_{m_{j,s}})} e^{(\beta_s t) \varepsilon_s} \]  

where:

- \( t_{m_{j,s}} \) is the time of the \( j \)th tamping for the \( s \)th track section,
- \( k_s \) is the number of tamping interventions on the \( s \)th track section,
- \( I(\cdot) \) is the indicator function, and

Due to the variability of the track structure, traffic conditions, environmental conditions, and maintenance history, there is section-to-section variation in the degradation parameters. In the degradation model applied in this study, the initial degradation value and degradation rates are considered as log-normally distributed random variables. This is in accordance with the findings of a study by Soleimanmeigouni et al. (2018).

4.2. Occurrence of severe isolated defects

In order to consider CM actions in our model, we needed to model the occurrence of UH2 defects and critical defects. The probability of occurrence of severe isolated defects is dependent on the typical track quality indices used to plan maintenance activities (Andrade & Teixeira, 2013). Therefore, in order to estimate the probability of the occurrence of UH2 defects and critical defects, ordinal logistic regression with TQI as the explanatory variable was applied in this study.

4.2.1. Ordinal logistic regression

Logistic regression is a well-known statistical method in data analysis for describing the relationship between a discrete response variable, taking on two or more possible values, and a set of explanatory variables. (Hosmer Jr, Lemeshow & Sturdivant, 2013). When a response variable has only two possible values, the binary logistic regression model is widely used as a standard method of analysis. In the field of track geometry degradation modelling and maintenance planning, binary logistic regression was applied to predict the occurrence of isolated defects. Cárdenas-Gallo, Sarmiento, Morales, Bolivar, and Akhavan-Tabatabaei (2017) applied binary logistic regression to identify the relationship between a set of explanatory variables including; defect amplitude, traffic, and class of track; and the future state of the isolated defects. Andrade and Teixeira (2013), used binary logistic regression to predict the probability of the occurrence of isolated defects. They considered the standard deviation of the longitudinal level and alignment, and the existence of a switch or bridge in a section as explanatory variables. It must be noted that if the binary logistic regression be used, only two outcomes can be predicted: no isolated defect in the track section or at least one isolated defect in the track section. In the case of our paper, the isolated defects are divided into multiple
discrete categories based on their severity. In order to handle this situation an extension of typical logistic regression must be applied. When the response variable is rank-ordered or ordinal with more than two levels, ordinal logistic regression can be used to estimate the probability that the response variable is classified into one of the outcomes. Therefore, in order to estimate the probability of the occurrence of UH2 defects and critical defects, the ordinal logistic regression model is used.

In ordinal logistic regression, by considering \( Y \) as an ordinal response variable which can take on \( k + 1 \) values coded \( 0, 1, 2, \ldots, k \), the cumulative logit transformation of \( Y \) is modelled as the linear function of the explanatory variables, as expressed in equation 4:

\[
\text{logit}(Y \leq k) = g_k(x) = \log\left( \frac{P(Y \leq k)}{1-P(Y \leq k)} \right) = \theta_k + c_1x_1 + c_2x_2 + \ldots + c_nx_n, \quad k = 1, \ldots, K
\]  \hspace{1cm} (4)

where:

- \( n \) is the number of explanatory variables,
- \( x = (x_1, \ldots, x_n) \) are explanatory variables,
- \( \theta_k \) is the constant associated with the \( k^{th} \) distinct response category,
- \( c_1, \ldots, c_n \) are the model coefficients, with the logit denoting the logit transformation, and
- \( g(x) \) is the linear function associated with the logit model.

The cumulative event probability \( P(Y \leq k|x) \) can be obtained as equation 5:

\[
P(Y \leq k|x) = \sum_{i=0}^{k} \theta_i(x) = \frac{e^{\theta_k(x)}}{1+e^{\theta_k(x)}}
\]  \hspace{1cm} (5)

The cumulative probabilities reflect the order of the response, as shown in equation 6:

\[
P(Y \leq 1) < P(Y \leq 2) \ldots \leq P(Y \leq K) = 1
\]  \hspace{1cm} (6)

4.2.2. Imbalanced data

An important challenge that must be addressed for estimation of the probability of the occurrence of severe isolated geometry defects is the issue of imbalanced data. Since in the present study, the ratio of the inspection intervals with at least one UH2 defect or critical defect to the inspection intervals without geometry defects is very small (less than 5 percent), the dataset is imbalanced. Research on imbalanced data is a challenging topic in data mining and machine learning. A characteristic of an imbalanced dataset is that it has many more instances of some classes than it has of other classes (Sun, Kamel, Wong & Wang, 2007). Consequently, rare events are difficult to detect due to their infrequency. When the minority class is the class of greatest interest, misclassifying rare events may cause a big error cost from a learning point of view (Albisua et al., 2013; Haixiang et al., 2017). Two main approaches to addressing the imbalanced data problem are algorithmic approaches and data approaches. Data approaches including resampling techniques are more versatile as they are independent of the learning algorithm (Albisua et al., 2013). Resampling methods are used to rebalance the dataset to alleviate the effect of the skewed class distribution in the learning process (Haixiang et al., 2017). Resampling methods are categorized into three groups, i.e. over-sampling methods, under-sampling methods and hybrid methods.

- **Over-sampling methods**: The aim of these methods is to produce a new dataset with a balanced class distribution by creating new minority class samples (Haixiang et al., 2017; Loyola-González, Martínez-Trinidad, Carrasco-Ochoa & García-Borroto, 2016).
- **Under-sampling methods**: In contrast to the oversampling approach, these methods discard the intrinsic samples in the majority class to create a balanced dataset.
- **Hybrid methods**: These methods are a combination of under-sampling methods and over-sampling methods.

In this study, the adaptive synthetic (ADASYN) sampling method (He, Bai, Garcia & Li, 2008) was used to address the imbalanced data problem. The ADASYN sampling method belongs to the category of over-sampling methods. The basic idea of the ADASYN method is to generate samples adaptively for the minority class based on their distributions, to reduce the bias introduced by the imbalanced data distribution. In fact, using ADASYN method more synthetic
objects are generated for minority class objects that are harder to learn compared to those minority objects that are easier to learn. ADASYN improves the learning process regarding the data distribution in two ways: by reducing the bias and adaptively learning (He et al. 2008). It must be noted that the ADASYN sampling method was originally created to handle two-class imbalanced data. To handle a multiclass problem, class transformation is used; i.e. when oversampling one of the minority classes, the other classes are considered as the majority class. Readers are referred to He et al. (2008) for more details concerning this sampling approach.

5. Case study

In order to test the applicability of the proposed model, a case study was conducted using track longitudinal level data collected from line section 414 between Järna and Katrineholm Central Station during 2007 to 2018. The maximum speed of trains on this line section is around 200 km/h. Line section 414 is 82 km long and consists of UIC 60 and SJ 50 rails, M1 ballast, Pandrol e-Clip fasteners, and concrete sleepers. Line section 414 is divided into different track sections with different lengths, mostly between 100 m and 300 m. The annual passing tonnage of the line section is around 20 MGT. In order to measure the vertical and lateral deviation of the track, the geometrical parameters are measured and recorded by two inspection cars, i.e. the IMV100 with a speed up to 80 km/h and the IMV200 with a speed up to 160 km/h.

5.1. Model implementation using measurement data

Short wavelength longitudinal level irregularities (wavelength 1 m to 25 m) are widely used to represent the track geometry quality and to plan maintenance actions (UIC, 2008). Therefore, in this study we considered the standard deviation of the longitudinal level as the degradation characteristic. The model presented in Section 4.1 was used to model the track geometry degradation. As is mentioned in Section 4.1, the initial degradation level and degradation rate are considered as random variables following a lognormal distribution. Figure 4 shows the distribution of the degradation rates and the initial standard deviation of the longitudinal level with a histogram plot.

![Figure 4. Histograms of the initial degradation value and degradation rate](image)

In order to test the lognormality of the degradation parameters, the Anderson-Darling (AD) test was applied. The p-values for the AD tests performed on the initial degradation value and degradation rate are 0.35 and 0.14, respectively. Considering the significance level of 0.05, the results of the tests show that there is no reason to reject the lognormality
assumption for the initial degradation value and degradation rate. The estimated parameters of the lognormal distributions for the initial degradation value and degradation rate are presented in Table 3.

Table 3. The results of parameter estimation for lognormal distributions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Location parameters</th>
<th>Scale parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial degradation value</td>
<td>-0.17</td>
<td>0.26</td>
</tr>
<tr>
<td>Degradation rate</td>
<td>-2.77</td>
<td>0.61</td>
</tr>
</tbody>
</table>

To construct the tamping recovery model, the following explanatory variables were included: (1) $D_s(t_{m,i,s})$, which is the degradation level before a tamping intervention, (2) $x_{1,i,s}$, which is a dummy variable with value 0 when the $j^{th}$ tamping is a partial tamping performed on section $s$ and value 1 when the $j^{th}$ tamping is a complete tamping performed on section $s$, and (3) $x_{2,j,s}^{(2)}, x_{2,j,s}^{(3)}, x_{2,j,s}^{(4)},$ and $x_{2,j,s}^{(5)}$, which are dummy variables representing the second, third, fourth, and fifth interventions, respectively. The recovery model is formulated as equation 7:

$$R_{i,s} = a_1 + b_0 D_s(t_{m,i,s}) + b_1 x_{1,i,s} + b_2 x_{2,j,s}^{(2)} + b_3 x_{2,j,s}^{(3)} + b_4 x_{2,j,s}^{(4)} + b_5 x_{2,j,s}^{(5)} + \delta\quad(7)$$

where $a_1, b_0, b_1, b_2, b_3, b_4,$ and $b_5$ are the model coefficients. The estimated parameters of the recovery model obtained using the least squares algorithm are presented in Table 4.

Table 4. Parameter estimates for recovery model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$a_1$</th>
<th>$b_0$</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
<th>$b_4$</th>
<th>$b_5$</th>
<th>$\tau^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates</td>
<td>1.08</td>
<td>0.32</td>
<td>0.36</td>
<td>-0.13</td>
<td>-0.23</td>
<td>-0.30</td>
<td>-0.36</td>
<td>0.06</td>
</tr>
</tbody>
</table>

In order to show the effect of the tamping type and the number of accumulated tamping interventions on the recovery value after tamping, two box plots were created, see Figure 5.

Figure 5. (a) Box plot of the recovery values for different numbers of tamping interventions, (b) Box plot of the recovery values for different tamping types

Figure 5a shows that the recovery value after tamping decreases as the number of tamping interventions increases, while Figure 5b shows that the expected recovery values after complete tamping interventions are higher than the expected recovery values after partial tamping.

In order to estimate the probability of the occurrence of UH2 defects and critical defects, the model presented in equation 4 is used. In this study, the response variable $Y$ took a value of 0 whenever there was no longitudinal level
defect in the track section, a value of 1 whenever there was at least one UH2 defect in the track section, and a value of 2 whenever there was at least one critical defect in the track section. The explanatory variable considered in the model is the standard deviation of the longitudinal level. The ordinal logistic regression for the proposed model is presented in equation 8:

\[ P(Y \leq k|x) = \frac{e^{\theta_k + \theta_2 SDLL}}{1 + e^{\theta_k + \theta_2 SDLL}} \]  

where SDLL is the standard deviation of the longitudinal level. The results for the fitting of the model are summarized in Table 5.

Table 5. The ordinal logistic regression results for the longitudinal level defects

<table>
<thead>
<tr>
<th></th>
<th>(\theta_1)</th>
<th>(\theta_2)</th>
<th>SDLL</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-value</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Coefficient</td>
<td>9.8</td>
<td>13.02</td>
<td>-5.02</td>
</tr>
</tbody>
</table>

As can be seen in Table 5, the p-values of the standard deviation of the longitudinal level are less than the significance level. Therefore, it can be concluded that these variables have a significant effect on the probability of the occurrence of UH2 defects and critical defects. According to Table 5, the standard deviation of the longitudinal level has a negative coefficient. This means that the higher the standard deviation is, the higher is the probability of the occurrence of at least a UH2 defect or a critical defect. Figure 6 shows the relationship between the standard deviation of the longitudinal level and the probability of the occurrence of UH2 defects and critical defects.

![Figure 6. Relationship between the standard deviation of the longitudinal level and the probability of no defect, at least one UH2 defect, and at least one critical defect in a track section](image)

As can be seen in Figure 6, when the standard deviation of the longitudinal level increases, the probability that a track section does not have a UH2 defect or critical defect decreases. As is obvious from Figure 6, increasing the standard deviation of the longitudinal level increases the probability of the occurrence of at least one critical defect in a track section.

5.2. Simulation results

By considering the longitudinal level as geometry parameters and using the proposed model, the percentage of time spent in different longitudinal level states can be estimated for the purpose of comparing different inspection intervals.
In total, we assessed the effect of the following five different inspection intervals on the track longitudinal level: 1 month, 2 months, 3 months, 4 months, and 6 months. The Monte Carlo simulation technique was used to obtain the expected values for the percentage of time spent in different longitudinal level states. The simulations were run for a 10-year period. The maintenance planning time for the performance of normal CM was assumed to follow a normal distribution with a mean of $m = 5$ weeks and a standard deviation of $\sigma = 1$ week. In order to perform PM, the tamping windows were set to be available every 18 months. The alert limit was considered to be 1.5 mm. For each simulation, 80,000 runs were performed to make sure that the simulation would be converged. Table 6 shows the percentage of time spent in the three defined longitudinal level states for different inspection intervals.

Table 6. Effect of the inspection interval on the percentage of time spent in the different longitudinal level states

<table>
<thead>
<tr>
<th>Inspection interval</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Month</td>
<td>99.26%</td>
<td>0.69%</td>
<td>0.05%</td>
</tr>
<tr>
<td>2 Months</td>
<td>98.55%</td>
<td>1.35%</td>
<td>0.10%</td>
</tr>
<tr>
<td>3 Months</td>
<td>97.85%</td>
<td>2.01%</td>
<td>0.14%</td>
</tr>
<tr>
<td>4 Months</td>
<td>97.20%</td>
<td>2.62%</td>
<td>0.18%</td>
</tr>
<tr>
<td>6 Months</td>
<td>95.79%</td>
<td>3.91%</td>
<td>0.30%</td>
</tr>
</tbody>
</table>

It must be noted that the running behaviour of the train is influenced by the amplitude of the track irregularities. Therefore, the percentage of time that track sections spend in different longitudinal level states has a significant effect on the level of safety and the ride comfort. If the track longitudinal level degrades to state 2 and state 3, this will negatively affect the energy consumption, the degradation rate of other track components, and passenger satisfaction. From an availability point of view, the number of speed restrictions that a track section suffers in a given period of time is related to the track geometry state. In fact, when the track longitudinal level degrades to state 2 and state 3, there is a higher probability of the occurrence of temporary speed restrictions. From a safety point of view, the risk of derailment in a given track section is dependent on the number of isolated defects, the amplitude of those defects, and the length of the time during which those defects are not detected. Therefore, if the percentage of time spent in state 3 increases, the risk of train derailment increases. The results of the Monte Carlo simulation show that the inspection frequency has a significant effect on the time that a track section spends in different longitudinal level states. Based on the results obtained, changing the inspection interval from one month to six months increases the percentage of time that a track section spends in state 2 and state 3 by 3.22% and 0.25%, respectively. According to Table 6, the shorter the inspection interval is, the shorter is the length of time spent by a track section in state 2. In addition, as the results of the simulations indicate, decreasing the frequency of inspections will significantly increase the percentage of time spent in state 3. When the length of the inspection interval increases, UH2 defects will be detected later. Therefore, with longer inspection intervals there is a higher probability that a UH2 defect will turn into a critical defect.

6. Conclusions

The aim of the research presented in this paper has been to develop an integrated approach to investigation of the effect of different inspection intervals on the track geometry condition. The standard deviation of the longitudinal level and the extreme values of isolated defects of the longitudinal level were considered as quality indicators to assess the need for PM and CM activities, respectively. The developed approach involves the use of a random coefficient piecewise exponential model to characterize the track geometry degradation and restoration. In addition, ordinal logistic regression is used to estimate the probability of the occurrence of isolated defects. The proposed model can be used to make predictions of the percentage of time that a track section will spend in the different track geometry states in a given time horizon. Such predictions are very important, as they are a primary input for predicting the probability of the occurrence of temporary speed reductions and the risk of derailment. To verify the applicability of the proposed approach, a case study has been conducted using a data set collected from the Main Western Line in Sweden. Based on the results obtained, it was observed that varying the length of the inspection intervals has a significant effect on the percentage of time spent by the track in different longitudinal level states. It must be noted that the proposed approach can also be used to predict the number of different maintenance actions, i.e. PM, normal CM, and emergency CM. However, since in the present study it is assumed that the PM actions are carried out on the track at predetermined times, the number of PM actions and normal CM actions is not sensitive to different inspection intervals. In the case of a condition-based maintenance policy being implemented, the inspection frequency may affect the number of PM actions and normal CM actions. In this case, by determining the different direct and indirect costs for preventive
tamping actions, normal corrective tamping actions, emergency corrective tamping actions, inspection, derailment, and the loss of capacity, the life cycle cost of each inspection interval can be assessed for the purpose of choosing the most effective one.

**Acknowledgements**

The authors would like to thank Trafikverket, Bana Väg För Framtiden (BVFF), Luleå Railway Research Center (JVTC), and SIMTRACK project partners for their technical and financial support provided during this project.

**References**


