Tampere University

MIKKO SAUNI

# Application of Track Geometry Deterioration Modelling and Data Mining in Railway Asset Management

Tampere University Dissertations 737

Tampere University Dissertations 737

MIKKO SAUNI

# Application of Track Geometry Deterioration Modelling and Data Mining in Railway Asset Management

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Built Environment
of Tampere University,
for public discussion in the auditorium RG202
of the Rakennustalo, Korkeakoulunkatu 5, Tampere,
on 3 February 2023, at 12 o’clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Built Environment
Finland

| | | |
|---|---|---|
| *Responsible supervisor and Custos* | Professor<br>Pauli Kolisoja<br>Tampere University<br>Finland | |
| *Supervisors* | Professor<br>Pauli Kolisoja<br>Tampere University<br>Finland | DSc (Tech)Heikki Luomala<br>Tampere University<br>Finland |
| *Pre-examiners* | Associate Professor<br>Stefan Marschnig<br>Graz University of Technology<br>Austria | PhD<br>Eric Berggren<br>EBER Dynamics AB<br>Sweden |
| *Opponents* | Professor<br>Jens Nielsen<br>Chalmers University of Technology<br>Sweden | Associate Professor<br>Stefan Marschnig<br>Graz University of Technology<br>Austria |

Cover design: Roihu Inc.

ISBN 978-952-03-2739-2 (print)
ISBN 978-952-03-2740-8 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-2740-8

PunaMusta Oy – Yliopistopaino
Joensuu 2023

# ACKNOWLEDGEMENTS

This dissertation was made possible by several people and organisations who I would like to acknowledge and thank.

My supervisors Dr. Heikki Luomala and Prof. Pauli Kolisoja have been guiding me since my masters' thesis in 2018, and they have been indispensable. Heikki, I value all our rambles in your office, whether they were on or off topic. I always came into your office with some problem and always left with either an answer or the motivation to find the answer no matter what. Pauli has been the guiding light, putting my results and also just about everything else into context. I also appreciate how you showed me the ropes in scientific publishing and conferences.

I was extremely fortunate to have distinguished professionals as my examiners and opponents. I consider my pre-examiners Prof. Stefan Marschnig from TU Graz and Dr. Eric Berggren the best experts to review my work. I thank you for taking the time and for the valuable comments I received. In February 2023 the opponents in the public defence will be Prof. Stefan Marschnig and Prof. Jens Nielsen, both of who I respect and value their work. I look forward to a lively conversation with you.

The financial support for this dissertation was provided by the Tampere University doctoral school program and the Finnish Transport Infrastructure Agency (Väylävirasto). I am forever grateful that I had the privilege of stable funding throughout my doctoral studies so I could concentrate on my studies and research.

I would also like to thank Prof. Esko Turunen and Dr. Tapio Nummi for their collaboration and sharing me their mathematical expertise. A big thank you also goes to Mika Silvast from Loram Finland Ltd for sharing data and software that made much of the research possible.

The people at the research centre Terra deserve high praise for creating the best work environment imaginable. From the energetic conversations at coffee breaks to the cheerful get-togethers, you have been the best work community I could have ever wished for. I hope you always keep the door open and the coffee warm when I come to visit.

I also feel grateful for having such a supportive family and friends, who have kept me on track during these busy years.

I have worked in the railway industry for many years now. I have met plenty of wonderful people who have helped me forward in my career and studies. This dissertation would not have been possible if it were not for all the people encouraging me to take on new challenges. For everyone I have had the pleasure of working with during my years in the railway industry, I would like to quote Yogi Berra regarding my upcoming public defence: Thank you for making this day necessary.

My next big project will be to take what I have learned and put it to good use. I have started working at Väylävirasto, where I have begun the process of driving this branch of the railway industry forward in Finland. This work will not be completed in a day, nor will it be easy. Achieving tangible change within the industry actually reminds me of when I was learning to code; if it feels comfortable and easy, you are probably doing something wrong.

Tampere 6.12.2022

Mikko Sauni

# ABSTRACT

In the management of a modern European railway system, spending is predominantly allocated to maintaining and renewing the existing rail network rather than constructing completely new lines. In addition to major costs, the maintenance and renewals of the existing rail network often cause traffic restrictions or line closures, which decrease the usability of the rail network. Therefore, timely maintenance that achieves long-lasting improvements is imperative for achieving competitive and punctual rail traffic. This kind of maintenance requires a strong knowledge base for decision making regarding the current condition of track structures.

Track owners commission several different measurements that depict the condition of track structures and have comprehensive asset management data repositories. Perhaps one of the most important data sources is the track recording car measurement history, which depicts the condition of track geometry at different times. These measurement results are important because they offer a reliable condition database; the measurements are done recurrently, two to six times a year in Finland depending on the track section; the same recording car is used for many years; the results are repeatable; and they provide a good overall idea of the condition of track structures. However, although high-quality data is available, there are major challenges in analysing the data in practical asset management because there are few established methods for analytics. Practical asset management typically only monitors whether given threshold values are exceeded and subjectively assesses maintenance needs and development in the condition of track structures. The lack of advanced analytics prevents the full utilisation of the available data in maintenance planning which hinders decision making.

The main goals of this dissertation study were to develop track geometry deterioration modelling methods, apply data mining in analysing currently available railway asset data, and implement the results from these studies into practical railway asset management. The development of track geometry deterioration modelling methods focused on utilising currently available data for producing novel information on the development in the condition of track structures, past maintenance effectiveness, and future maintenance needs. Data mining was applied in investigating the root causes of track geometry deterioration based on asset data. Finally, maturity models were applied as the basis for implementing track geometry deterioration modelling and track asset data analytics into practice.

Based on the research findings, currently available Finnish measurement and asset data was sufficient for the desired analyses. For the Finnish track inspection data, robust linear optimisation was developed for track geometry deterioration modelling. The modelling provided key figures, which depict the condition of structures, maintenance effectiveness, and future maintenance needs. Moreover, visualisations were created from the modelling to enable the practical use of the modelling results. The applied exploratory data mining method, General Unary Hypotheses Automaton (GUHA), could find interesting and hard-to-detect correlations within asset data. With these correlations, novel observations on problematic track structure types were made. The observations could be utilised for allocating further research for problematic track structures, which would not have been possible without using data mining to identify these structures. The implementation of track geometry deterioration and asset data analytics into practice was approached by applying maturity models. The use of maturity models offered a practical way of approaching future development, as the development could be divided into four maturity levels, which created clear incremental goals for development. The maturity model and the incremental goals enabled wide-scale development planning, in which the progress can be segmented and monitored, which enhances successful project completion.

The results from these studies demonstrate how currently available data can be used to provide completely new and meaningful information, when advanced analytics are used. In addition to novel solutions for data analytics, this dissertation research also provided methods for implementing the solutions, as the true benefits of knowledge-based decision making are obtained in only practical railway asset management.

# TIIVISTELMÄ

Modernin rautatiejärjestelmän hallinnassa rahankäyttö kohdistuu valtaosin nykyisen rataverkon korjauksiin ja parannuksiin ennemmin kuin uusien ratojen rakentamiseen. Nykyisen rataverkon kunnossapitotyöt aiheuttavat suurten kustannusten lisäksi myös usein liikennerajoitteita tai yhteyksien väliaikaisia sulkemisia, jotka heikentävät rataverkon käytettävyyttä. Siispä oikea-aikainen ja pitkäaikaisia parannuksia aikaansaava kunnossapito ovat edellytyksiä kilpailukykyisille ja täsmällisille rautatiekuljetuksille. Tällainen kunnossapito vaatii vankan tietopohjan radan nykyisestä kunnosta päätöksenteon tueksi.

Ratainfran omistajat teettävät päätöksenteon tueksi useita erilaisia radan kuntoa kuvaavia mittauksia ja ylläpitävät kattavia omaisuustietorekistereitä. Kenties tärkein näistä datalähteistä on koneellisen radantarkastuksen tuottamat mittaustulokset, jotka kuvastavat radan geometrian kuntoa. Nämä mittaustulokset ovat tärkeitä, koska ne tuottavat luotettavaa kuntotietoa: mittaukset tehdään toistuvasti, 2–6 kertaa vuodessa Suomessa rataosasta riippuen, mittausvaunu pysyy useita vuosia samana, tulokset ovat hyvin toistettavia ja ne antavat hyvän yleiskuvan radan kunnosta. Vaikka laadukasta dataa on paljon saatavilla, käytännön omaisuudenhallinnassa on merkittäviä haasteita datan analysoinnissa, sillä vakiintuneita menetelmiä siihen on vähän. Käytännössä seurataan usein vain mittaustulosten raja-arvojen ylittymistä ja pyritään subjektiivisesti arvioimaan rakenteiden kunnon kehittymistä ja korjaustarpeita. Kehittyneen analytiikan puutteet estävät kuntotietojen laajamittaisen hyödyntämisen kunnossapidon suunnittelussa, mikä vaikeuttaa päätöksentekoa.

Tämän väitöskirjatutkimuksen päätavoitteita olivat kehittää ratageometrian heikkenemiseen mallintamismenetelmiä, soveltaa tiedonlouhintaa saatavilla olevan omaisuusdatan analysointiin sekä jalkauttaa kyseiset tutkimustulokset käytännön rataomaisuudenhallintaan. Ratageometrian heikkenemisen mallintamismenetelmien kehittämisessä keskityttiin tuottamaan nykyisin saatavilla olevasta datasta uutta tietoa radan kunnon kehityksestä, tehdyn kunnossapidon tehokkuudesta sekä tulevaisuuden kunnossapitotarpeista. Tiedonlouhintaa sovellettiin ratageometrian heikkenemisen juurisyiden selvittämiseen rataomaisuusdatan perusteella. Lopuksi hyödynnettiin kypsyysmalleja perustana ratageometrian heikkenemisen mallinnuksen ja rataomaisuusdatan analytiikan käytäntöön viennille.

Tutkimustulosten perusteella suomalainen radantarkastus- ja rataomaisuusdata olivat riittäviä tavoiteltuihin analyyseihin. Tulokset osoittivat, että robusti lineaarinen optimointi soveltuu hyvin suomalaisen rataverkon ratageometrian heikkenemisen mallinnukseen. Mallinnuksen avulla voidaan tuottaa tunnuslukuja, jotka kuvaavat rakenteen kuntoa, kunnossapidon tehokkuutta ja tulevaa kunnossapitotarvetta, sekä muodostaa havainnollistavia visualisointeja datasta. Rataomaisuusdatan eksploratiiviseen tiedonlouhintaan käytetyn GUHA-menetelmän avulla voitiin selvittää mielenkiintoisia ja vaikeasti havaittavia korrelaatioita datasta. Näiden tulosten avulla saatiin uusia havaintoja ongelmallisista ratarakennetyypeistä. Havaintojen avulla voitiin kohdentaa jatkotutkimuksia näihin rakenteisiin, mikä ei olisi ollut mahdollista, jollei tiedonlouhinnan avulla olisi ensin tunnistettu näitä rakennetyyppejä. Kypsyysmallin soveltamisen avulla luotiin puitteet ratageometrian heikkenemisen mallintamisen ja rataomaisuusdatan analytiikan kehitykselle Suomen rataomaisuuden hallinnassa. Kypsyysmalli tarjosi käytännöllisen tavan lähestyä tarvittavaa kehitystyötä, kun eteneminen voitiin jaotella neljään eri kypsyystasoon, jotka loivat selkeitä välitavoitteita. Kypsyysmallin ja asetettujen välitavoitteiden avulla kehitys on suunniteltua ja edistystä voidaan jaotella, mikä antaa edellytykset tämän laajamittaisen kehityksen onnistuneelle läpiviennille.

Tämän väitöskirjatutkimuksen tulokset osoittavat, miten nykyisin saatavilla olevasta datasta saadaan täysin uutta ja merkityksellistä tietoa, kun sitä käsitellään kehittyneen analytiikan avulla. Tämä väitöskirja tarjoaa datankäsittelyratkaisujen luomisen ja soveltamisen lisäksi myös keinoja niiden käytäntöönpanolle, sillä tietopohjaisen päätöksenteon todelliset hyödyt saavutetaan vasta käytännön radanpidossa.

# CONTENTS

# KEY TERMINOLOGY

| | |
|---|---|
| Asset management | In the context of this dissertation, asset management refers to managing the maintenance, renewals, and life cycle of infrastructure assets, such as track structures, bridges, or signalling equipment. |
| Data | The word data is purposefully used in the uncountable form, as in data *is* available. It is acknowledged that the word data has origins in the Latin word datum, whose plural is data, but current dictionaries, like the Cambridge Dictionary (2022a) and Merriam-Webster (2022), accept the use of the word data as a singular and uncountable noun due to the diminished use of the original singular form, datum. |
| Framework | "A system of rules, ideas, or beliefs that is used to plan or decide something" (Cambridge Dictionary, 2022b) |
| Long term and short term (behaviour, maintenance, predictions, etc.) | Long term is used to refer to time periods lasting one or more year(s). Conversely, short term refers to time periods lasting less than one year. |
| Maintenance and tamping | Maintenance is used to refer to any methods of improving or sustaining the condition of in-use track structures. Tamping is regarded as one maintenance method, which can be accompanied or preceded by any number of other maintenance actions as well. When wording such as "tamping predictions" is used, it implies that at least tamping will be required, but other maintenance methods could be needed as well. |

# ABBREVIATIONS

| | |
|---|---|
| ANN | Artificial Neural Network |
| CMM | Capability Maturity Model |
| CMMI | Capability Maturity Model Integration |
| EN | European Standard |
| ETRS | European Terrestrial Reference System |
| FEM | Finite Element Method |
| FTIA | Finnish Transport Infrastructure Agency (Väylävirasto) |
| GPR | Ground Penetrating Radar |
| GNSS | Global Navigation Satellite System |
| GUHA | General Unary Hypotheses Automaton |
| ISO | International Organization for Standardization |
| ITDM | Integrated Track Degradation Model |
| LL | Longitudinal Level |
| MGT | Million Gross Tons |
| MSI | Maintenance Success Indicator |
| PI | Prediction Interval |
| $R^2$ | Coefficient of Determination |
| RAMS | Reliability, Availability, Maintainability and Safety |
| RO | Research Objective |
| SD | Standard Deviation |
| TGDR | Track Geometry Deterioration Rate |
| TGDA | Track Geometry Deterioration Analyses |
| UIC | International Union of Railways (Union Internationale des Chemins de Fer) |

# LIST OF ORIGINAL PUBLICATIONS

| | |
|---|---|
| Publication I | Sauni, M., Luomala, H., Kolisoja, P. and Nummi, T. (2022), “Advancing railway asset management using track geometry deterioration modelling visualization”, Journal of Transportation Engineering, Part A : Systems, Vol. 148 No. 2, available at: 10.1061/JTEPBS.0000626. |
| Publication II | Sauni, M., Luomala, H., Kolisoja, P. and Turunen, E. (2020), “Investigating Root Causes of Railway Track Geometry Deterioration – A Data Mining Approach”, Frontiers in Built Environment, Vol. 6, p. 122. |
| Publication III | Sauni, M., Luomala, H., Kolisoja, P. and Turunen, E. (2020), “Determining Sampling Points Using Railway Track Structure Data Analysis BT - Information Technology in Geo-Engineering”, in Correia, A.G., Tinoco, J., Cortez, P. and Lamas, L. (Eds.), Springer International Publishing, Cham, pp. 841–856. |
| Publication IV | Sauni, M., Luomala, H., Kolisoja, P. and Vaismaa, K. (2022), “Framework for implementing track deterioration analytics into railway asset management”, in Built Environment Project and Asset Management, Vol. 12 No. 6, pp. 871–886, available at: 10.1108/BEPAM-04-2022-0058. |

# AUTHOR CONTRIBUTIONS

| | |
|---|---|
| Publication I | The author of this dissertation was responsible for investigating the relevant mathematical models, writing the code for modelling, performing the modelling, analysing the results, and writing the paper as the corresponding author. |
| Publication II | The author of this dissertation was responsible for connecting different data sources to form the initial data, performing the data mining, analysing the results, and writing the paper as the corresponding author. |
| Publication III | The author of this dissertation was responsible for connecting different data sources to form the initial data, performing the data mining, assigning the sampling points, evaluating the laboratory test results, and writing the paper as the corresponding author. |
| Publication IV | The author of this dissertation was responsible for investigating the maturity models; designing and performing the interviews; designing; arranging and managing the workshops; formulating the framework; and writing the paper as the corresponding author. |

# 1 INTRODUCTION

## 1.1 Background and motivation

More than half of all rail infrastructure expenditure in the European Union is spent on maintenance and renewals (European Commission, 2021). Therefore, increasing the efficiency of maintenance and renewals has a great impact on railway infrastructure spending, which is mostly (72%) funded by national budgets (European Commission, 2021). Therefore, saving taxpayer money on railway maintenance and renewals has widespread beneficial effects, as the savings can be reallocated to other areas of society as well.

Railway maintenance and renewals have unique characteristics that are very different from, for example, roadworks. Rail traffic can rarely be rerouted to other track sections without causing major traffic capacity issues and temporarily diverted routes are immensely expensive to build. This means that the railway industry has a special interest in avoiding situations, in which maintenance and renewals are performed after structures have become defective, as this would lead to traffic disruptions or even line closures before track works are completed. Instead, railway maintenance and renewals must be planned ahead and performed within the time slots when trains are not running, so as not to disturb the rail traffic.

Proactive and preventative maintenance is based on knowing the condition of track structures and being able to estimate the development in their condition. Perhaps the most crucial data source in railway asset management is the track recording car (Esveld *et al.*, 1988). These cars are used to periodically measure the relative position of the rails because the rails do not remain in the same position as where they were initially installed but move due to deterioration and accumulated settlements in the structures below. Therefore, track (recording car) inspections are required to monitor, whether the track geometry is smooth enough to ensure safe rail traffic and detect irregularities that require maintenance, as these irregularities denote realised deterioration in the track structure.

Track inspections are conducted regularly, thus producing time series data on the track geometry condition. Past research has shown that this track inspection time series data can be used to investigate development in the condition of track structures (Andrade and Teixeira, 2015; Quiroga and Schnieder, 2012; Tanaka *et al.*, 2018), assess maintenance effectiveness (Audley and Andrews, 2013; Martey and Attoh-Okine, 2018; Soleimanmeigouni *et al.*, 2018), and predict future maintenance needs (Caetano and Teixeira, 2016; Lee *et al.*, 2020; Soleimanmeigouni *et al.*, 2020). These research outputs enable proactive and preventative maintenance. Subsequently, the tracks can be maintained before dangerous track geometry irregularities occur, and track works can be planned to be performed when there is no train traffic.

Nevertheless, the track inspection time series data from Finland has not been utilised in practical asset management to its full potential yet. Finnish practical asset management has had only static inspection reports (summaries provided in PDF format) to use for planning maintenance and renewals. Using only these types of reports makes maintenance planning difficult because assessing development in the condition of track structures is difficult and can only be done subjectively. Therefore, research on establishing suitable modelling methods for Finnish track inspection data is required to increase the efficiency of railway maintenance and renewals in Finland.

With track geometry deterioration modelling, the locations and severity of defective track structures can be identified. Nevertheless, different types of defects require different repair methods. To select the correct repair method, the root causes of track geometry deterioration should be known to avoid treating the symptoms instead of the disease, figuratively. The root causes of track geometry deterioration can be very complex, as a defect in one part of a track structure might weaken another part, thus deluding analyses attempting to find the original cause of deterioration. Track geometry deterioration root cause investigations require rigorous data on various features of the track structure. However, as the amount of data sources increases, subjective data analyses become more and more challenging. Therefore, the use of novel data analysis methods is required to explore the available track asset data to investigate the root causes of deterioration.

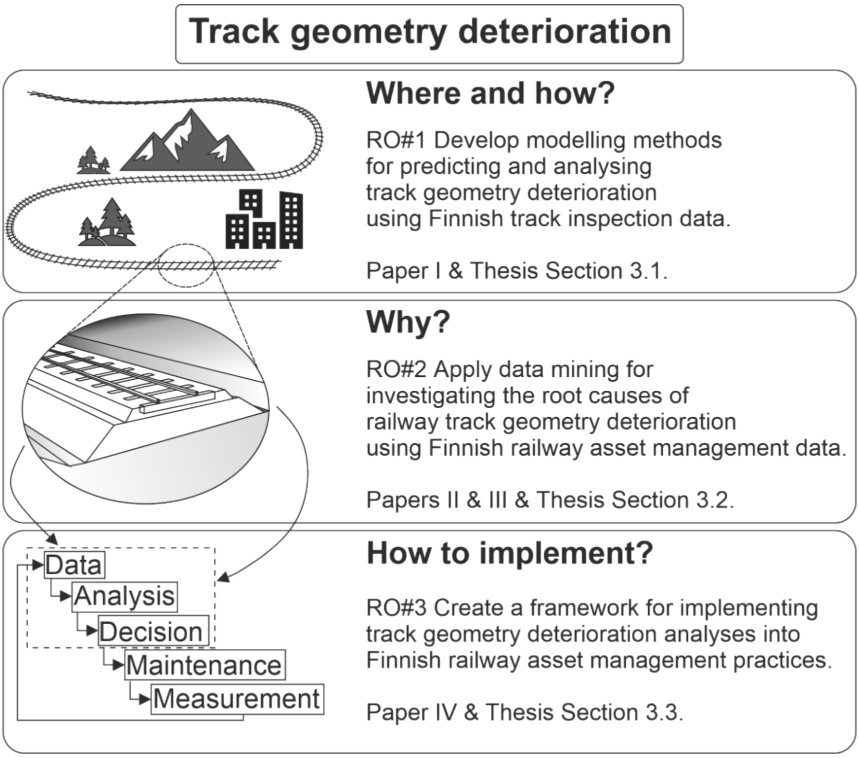
Finally, once track geometry deterioration modelling practices and data analysis methods are researched, they must be implemented into practical railway asset management. This is especially important as the greatest and most tangible benefits from these results can be obtained in practical operations. However, the implementation is not simple, because it will require large-scale software and organisational development. Therefore, the implementation of track geometry deterioration modelling and data analytics is required.

## 1.2 Research objectives and scope

The three research objectives (ROs) of this study were as follows:

- RO#1: Develop modelling methods for track geometry deterioration predictions and analyses using Finnish track inspection data.
- RO#2: Apply data mining for investigating the root causes of railway track geometry deterioration using Finnish railway asset management data.
- RO#3: Implement track geometry deterioration analyses into Finnish railway asset management practices.

The ROs were investigated in Papers I–IV and discussed in this thesis in the order presented in Figure 1.



RO#1 Develop modelling methods for predicting and analysing track geometry deterioration using Finnish track inspection data.

Paper I & Thesis Section 3.1.

RO#2 Apply data mining for investigating the root causes of railway track geometry deterioration using Finnish railway asset management data.

Papers II & III & Thesis Section 3.2.

RO#3 Create a framework for implementing track geometry deterioration analyses into Finnish railway asset management practices.

Paper IV & Thesis Section 3.3.

*Figure 1. Summary of the ROs.*

The first RO investigates answers to the following: *How can track geometry deterioration be modelled using Finnish track inspection data?* In practical terms, the modelling methods produce answers to where and how the track geometry deteriorates. These investigations have been reported in Paper I and section 3.1 of this thesis.

The second RO explores ways to answer the question: *How can data mining be used to investigate the root causes of track geometry deterioration using Finnish railway asset data?* The practical aspect of this RO is to use data to investigate why track geometry deteriorates. The available data depicts track geometry deterioration development and data on track structure features (see section 1.4). Studies relating to this RO are reported in Papers II–III and section 3.2 of this thesis.

The third RO studies *how to implement track geometry deterioration analyses into Finnish railway asset management*, which is also the practical perspective of this part of the study. Paper IV and section 3.3 of this dissertation report the studies related to this topic.

All the research objectives together enable more efficient planning and monitoring of track maintenance and renewals by utilising currently available data. All three research objectives investigate *how* to achieve this, meaning suitable methods are first investigated from a theoretical perspective and then their practical implementation is considered.

The scope of this research is focused on using only currently available railway structure data from Finnish railway asset management for analysing the condition of track structures. New data was gathered for the practical implementation of these analyses by using interviews and workshops held in cooperation with relevant railway industry stakeholders, as is elaborated further in Paper IV. As the research was limited to concern the Finnish state-owned rail network, it is important to consider the way Finnish railway ownership and management responsibilities are divided. Figure 2 presents a graph of this division.

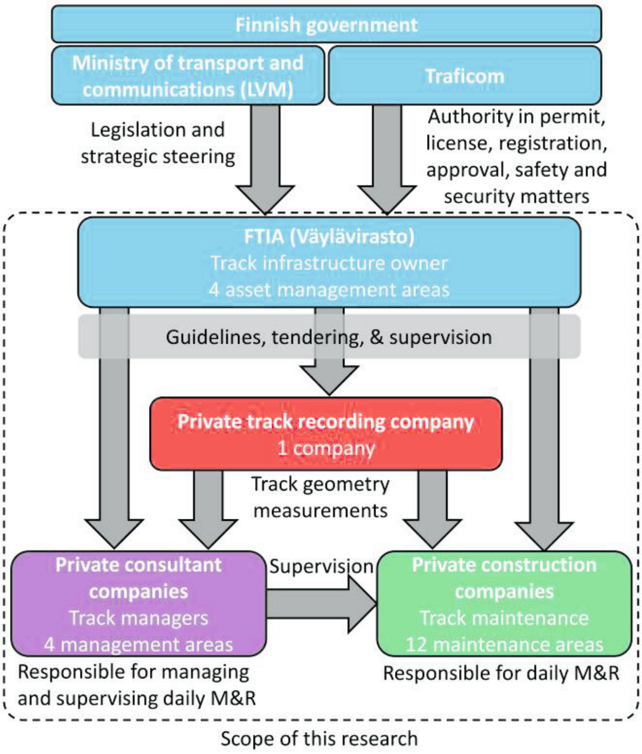


*Figure 2. Division of responsibilities in the Finnish state rail network ownership and management (Paper IV).*

The Finnish state-owned railway infrastructure is owned and controlled by the Finnish Transport Infrastructure Agency (FTIA, Väylävirasto in Finnish), which is a government agency. The FTIA is strategically steered by the Ministry of transport and communication, which is a branch of the Finnish government. The permits for owning and operating the railway system are controlled by another Finnish government agency, Traficom. The FTIA outsources much of the daily operations and focuses on managing the operations through contracts and guidelines. The daily track asset management is outsourced to private consultant companies, who control four track asset management areas. Similarly, track maintenance is divided into 12 areas, where private track construction companies are responsible for daily track maintenance operations. The track inspections are performed by one private company that uses one track recording car for the whole network.

The scope of researching the implementation of track geometry deterioration modelling and data analytics is limited to planning and roadmapping the future development in Finland. This excludes research on change management during the implementation, which is left out of scope as it will be a topic of future research.

## 1.3 Research methodology

The research strategy was based on applied empirical research, in which observations gathered during the research were used in the analyses. Nevertheless, the research design, type, and methods varied by RO. A summary of the used research designs, types, and methods is presented in Figure 3 and is subsequently discussed. The terminology is based on the book by Goddard and Melville (2001).

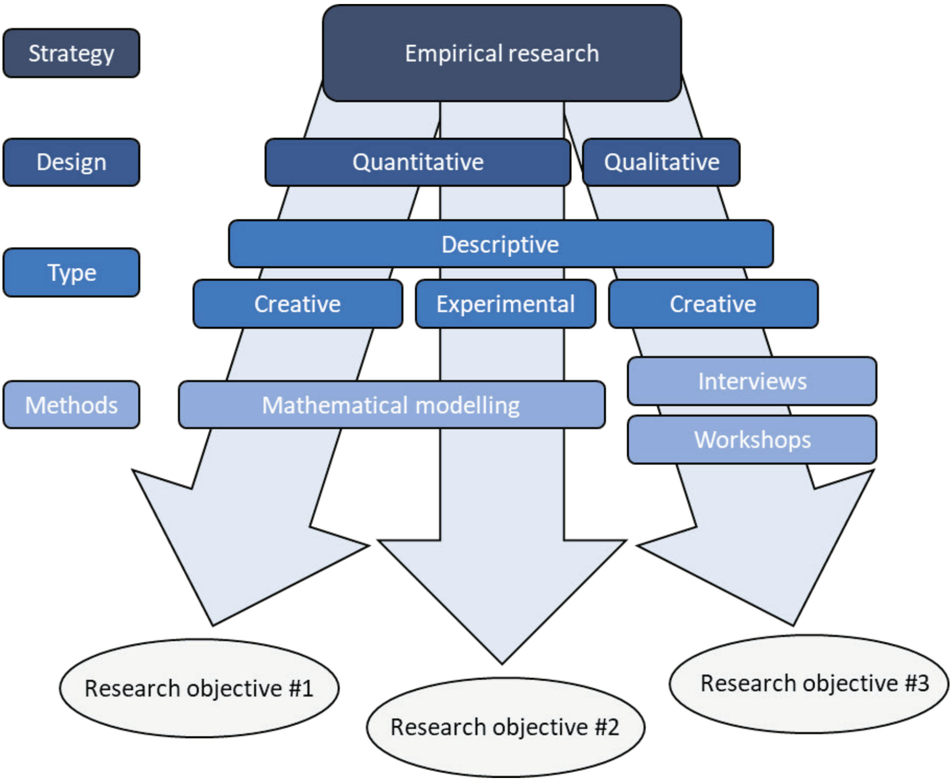


*Figure 3. Summary of the used research methods.*

The first RO, track geometry deterioration modelling, was approached with a quantitative research design; Data from track inspections was mathematically modelled to produce numeric data depicting the deterioration behaviour. The research type was considered descriptive and creative. Descriptive research studies specific situations to generate theories based on those situations. Correspondingly, actual track inspection data from multiple track sections was used to study a suitable modelling method. Additionally, the research was creative, as novel visualisations were generated from the modelling results to benefit practical railway asset management. As for the research methods, mathematical models were used both in generating the models and creating the results visualisations. The assumption in this part of the research was that Finnish track inspection data could exhibit similar deterioration behaviour as has been noticed in past research on track geometry deterioration in other countries. It was also assumed that the modelling results would provide intuitive information and that only the means to produce this information were missing.

The second RO, investigating the root causes of track geometry deterioration, was also approached with a quantitative research design. Data depicting different track structure features was used to find correlations between track structure types and track geometry deterioration behaviour. The descriptive part of the research concerned the initial data, which was actual measurement and asset data from in-use railway sections. Moreover, a case track section was used in testing the data mining results. The data mining used to find the correlations within the data was considered experimental research because the independent and dependent variables could be identified. The independent variables explaining the behaviour were the track structure features, and the dependent variable depicting the behaviour was the deterioration rate. The dependence between these two parameters was investigated using data mining, which is a form of mathematical modelling. The assumption was that the available track structure data would be sufficient to reveal interesting correlations between different track structure types and different deterioration behaviour. However, it was acknowledged that track geometry deterioration is a highly complex phenomenon and that the results will only imply correlations between structure types and observed deterioration behaviour. These correlations will only provide interesting hypotheses to investigate further with other research methods instead of elaborating on the actual structural mechanics.

For the third RO, the implementation of track geometry deterioration analytics, a qualitative research strategy was employed. The research included creative and descriptive elements. In the creative research part, maturity models were applied to form a basis for a track geometry deterioration analysis development framework. The descriptive part of the research included investigating the development framework contents in cooperation with stakeholders in Finnish railway asset management. The research methods used to formulate the framework included semi-structured expert interviews and workshops with the stakeholders. Creating a framework for planning future development can also be considered as roadmapping in other terminology (Hirose *et al.*, 2020; Kostoff and Schaller, 2001). In this dissertation, these are considered synonymous, as per the example of Kostoff and Schaller (2001), where it is mentioned that "*roadmaps present a framework*". In analysing the interview and workshop results, quantitive research elements were used when analysing the coded interview and workshop outputs. In studying the third RO, it was assumed that a maturity model could create a structure for the framework and that industry experts could provide the necessary domain knowledge to form its content. Additionally, it was acknowledged that the framework will be limited to only laying out a strategy for development, and that individual steps within the framework will require further defining and planning before practical implementation.

## 1.4 Research data

The research data included data depicting track geometry deterioration development, data on track structure features, and data produced in the interviews and workshops held as part of investigating the third RO. The research scope limited the use of track structure data to only currently available data. With that said, the case track sections were selected on the basis that sufficient data was available, and it should be considered that not all Finnish track sections have this much data readily available.

The data depicting track geometry deterioration development originated from periodical track inspections of the Finnish state rail network. Three track sections' data was acquired and used for different types of analyses. These track sections included Kouvola–Kotka (data from 2004–2017), Luumäki–Imatra (data from 2008–2018), and Karjaa–Ervelä (data from 2007–2015). The track section between Kouvola and Kotka is approximately 54 km in length, and both cargo and passenger traffic are operated on the track section. The yearly traffic volume on Kouvola–Kotka track section is 14–18 MGT depending on the location within the section. The Luumäki–Imatra (abbreviated as LUIMA) is approximately 65 km long and the yearly traffic volume is approximately 11 MGT. Lastly, the Karjaa–Ervelä (abbreviated as Rantarata) track section is about 29 km long and the yearly traffic volume is approximately 2 MGT. All track sections were inspected with the same track recording car, Plasser & Theurer EM120. This track recording car uses chord measurements from three bogies spaced 5 and 7 m apart, respectively. The track inspection data was acquired from Loram Finland Oy (formerly Roadscanners Oy), and the different measurement run signals had been aligned by Loram Finland Oy before handing over the data.

The data on the features of track structures originated from multiple sources and concerned the previously mentioned track sections. Track geometry deterioration rates were calculated from the track geometry history to provide a parameter for track deterioration behaviour. Track structure layer thicknesses and moisture indices were derived from ground penetrating radar (GPR) measurements. Laser scannings provided point clouds of the embankment surface, which was used to calculate ditch depth to provide data on drainage conditions. The GPR measurements and laser scannings were performed and analysed by Loram Finland Oy. Track deflection was measured using a continuous track stiffness device, which was developed in the Tampere University Research Centre Terra (Luomala *et al.*, 2017). The deflection measurements also provided cant data, which could be used to differentiate straight track elements from curves. Other data included asset data from railway asset data warehouses, soil maps, maintenance records, and data identified from a video feed of the track sections.

The data from interviews and workshops included their outputs and codings from the outputs. The interview outputs included the memos of the conversations, while the workshop outputs included mind maps, frameworks, and comments.

# 2 THEORETICAL BACKGROUND

This section presents the relevant background and previous research for the studies presented in Papers I–IV and section 3 of this dissertation.

## 2.1 Definition and measurement of track geometry

Track geometry can be defined and measured in two parallel systems: absolute and relative (Sánchez *et al.*, 2017; UIC, 2008). Absolute track geometry is measured and defined in a specified coordinate system, for example, the European Terrestrial Reference System (ETRS). This coordinate-based approach to track geometry is used mainly when designing and constructing railways when the shape and location of the track are created. In turn, relative track geometry is used in monitoring the condition of track geometry, meaning the deviations from an ideal shape. The relative track geometry is used for inspecting only the shape of the track geometry, not absolute coordinates (Sánchez *et al.*, 2017). As demonstrated in Figure 4, the major difference between these absolute and relative systems is that relative measurements do not recognise uniform settlements or movements, as long as the original shape remains (Esveld, 2001). Other differences between absolute and relative track geometry include the measurement reference systems, measurement loading conditions, and measurement speed.

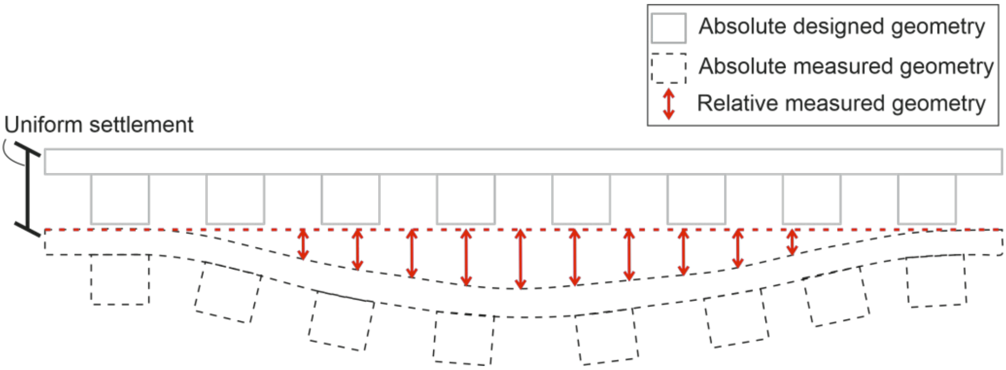


*Figure 4. Difference between absolute track geometry and relative track geometry.*

Absolute track geometry measurements need to be fixed in a coordinate system via known reference points near the tracks (UIC, 2008). The fixed points are required because track geometry measurements based solely on the Global Navigation Satellite System (GNSS) are typically not accurate enough (UIC, 2008). Nevertheless, recent advances in GNSS systems have shown measurement results with high accuracy (Szmagliński *et al.*, 2022), and it may be possible to obtain absolute geometry measurements using satellite technology in the future. The current devices used to measure the absolute geometry include total stations and lightweight surveying trolleys (Figure 5). These measurements provide very accurate information on the absolute track geometry (Chen *et al.*, 2018; Sánchez *et al.*, 2017). The downside is that the measurement speed is limited to walking pace and total station repositioning takes time (Sánchez *et al.*, 2017). Therefore, it is impractical to frequently measure the whole rail network with absolute geometry measurements. Furthermore, absolute geometry measurements are performed using lightweight equipment, thus the geometry under seating load (Li *et al.*, 2015; Sussmann *et al.*, 2001) is not recorded. As mentioned before, the best uses of absolute track geometry measurements are in designing and constructing railways but using them for network-wide condition monitoring is not currently feasible.

*Figure 5. Lightweight track geometry measurement trolley and total station. Photo credit: Marko Happo and Mikko Sauni.*

The relative track geometry measurements depict the shape of the rails using the parameters described in European Standard (EN) 13848-1 (2019). The measurements must be fixed to known coordinate and track location systems, but the accuracy requirements are lower than those of absolute geometry measurements. Therefore, GNSS accuracy coupled with reference points from track assets, for example, catenary poles, is sufficient. The relative geometry is usually measured using a specific track recording car (Esveld, 2001; SFS-EN 13848-2, 2020). This enables high measurement speeds of 160 km/h and up to 250 km/h (Lichtberger, 2005; Mermec, 2022; Plasser & Theurer, 2022a). The track recording cars measure loaded track geometry, which shows the effects of unsupported sleepers. Track geometry condition monitoring is generally based on recurrent relative track recording car measurements, because of the high measurement speeds. However, these measurements only provide information on the condition of track geometry, and absolute measurements are required for designing new track geometry for improvements and absolute tamping (UIC, 2008).

From here on, this dissertation will discuss only relative track geometry, if not otherwise mentioned, because track geometry is periodically inspected only with the track recording car, which measures the relative track geometry. Expectedly, future condition monitoring will be able to incorporate absolute track geometry monitoring via GNSS (Szmagliński *et al.*, 2022) or developed track recording cars (Plasser & Theurer, 2022b), but current collected data concerns only the relative track geometry.

## 2.2 Track geometry maintenance

Track geometry deteriorates when the forces subjected to the track structure exceed the capability of the track structure to resist those forces (Li *et al.*, 2015). Track geometry deterioration is observed in track recording car measurements as irregularities in a designed ideal geometry. Track geometry irregularities exceeding given thresholds must be corrected to ensure the safety of train traffic. Track geometry is predominantly corrected using a tamping machine (Esveld, 2001; Li *et al.*, 2015). Figure 6 presents the working principle of a single tie tamper. There are also multiple tie tampers whose working principle is similar to the single tie tamper, but, as per the name, more than one sleepers are tamped simultaneously (Plasser & Theurer, 2022c).

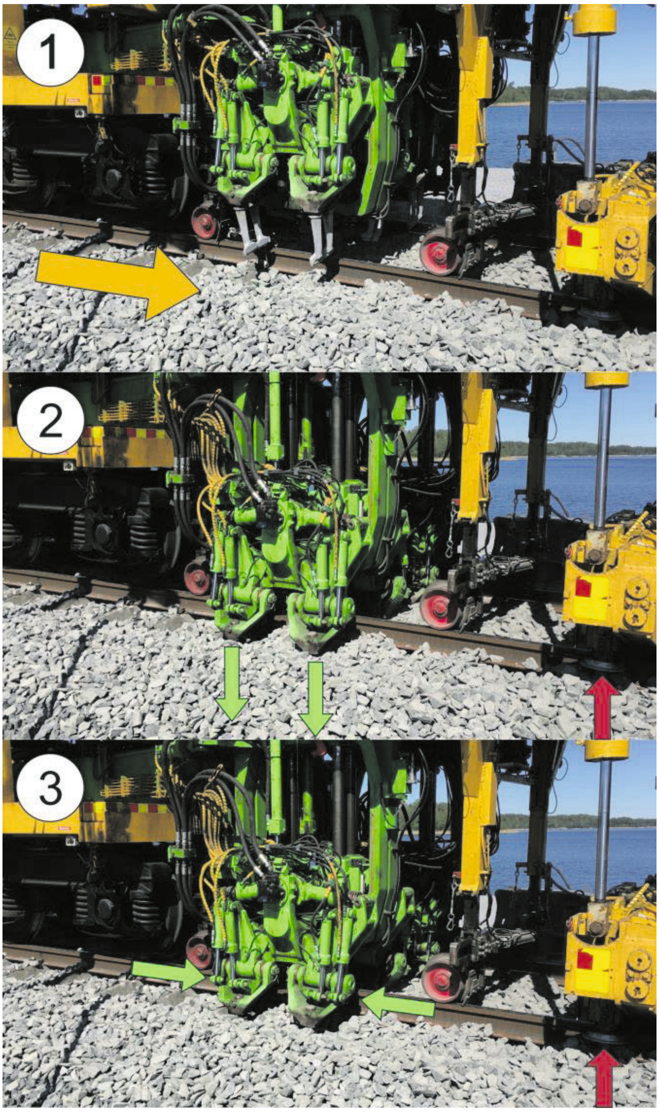


*Figure 6. Tamping machine working principle. Photo credit: Mikko Sauni.*

The working principle of a tamping machine (presented in Figure 6) is described as follows. First (1) the tamping machine moves into place, and designed displacement and lift values are set as input data for the machine. Secondly, (2) the rails are lifted to the designed position using a roller, indicated by the red arrow, and tines are pushed into the ballast around a sleeper. Lastly, (3) the tines are pushed together and vibrated to compact the ballast material underneath the sleeper to attain a permanent lifted position for the sleeper and rails. Once the tines are lifted from the ballast, the machine moves over to the next sleeper and repeats the process. Tamping can also be done with hydraulic jacks and handheld tamping devices, or tamping heads attached to excavators, but their use is not advisable for normal maintenance. These tamping devices are suitable when tamping only a few sleepers after minor track work. The handheld tamping devices may also be used together with a tamping machine when tamping turnout areas.

Tamping can be preceded by any number of track works ranging from replacing the rails to renewing the whole superstructure of the track. At the very least, tamping often requires adding ballast material to account for the lifted track geometry. Regardless of the maintenance works preceding tamping, it is the predominant method for implementing the desired track geometry (Esveld, 2001; Li *et al.*, 2015).

Tamping is followed by profiling and, in some cases, stabilising (Lichtberger, 2005). Dynamic stabilising of the ballast layer is performed using a special track work machine, a dynamic track stabiliser (Lichtberger, 2005). This machine induces vibrations to the track structure which causes the ballast particles to rearrange and lock into place (Li *et al.*, 2015). Dynamic stabilising machines are reported to have a positive effect, especially on horizontal track stability, which is beneficial when traffic can resume at full speed after tamping instead of speed restrictions (Esveld, 2001; Li *et al.*, 2015). However, dynamic track stabilisers are rare in Finland, and this phase is usually skipped, yet, speed restrictions after tamping are generally not applied. Profiling, conversely, can be performed with high rail excavators or special ballast regulators and is required almost always after tamping. Profiling forms the ballast shoulders and slopes that provide necessary horizontal track stability.

Alternatives to tamping include stone blowing, use of adjustment plates inserted below the rails, and ballast layer adhesion. Stone blowing is a similar operation to tamping, but instead of compacting the ballast material beneath the sleepers, gravel or sand is blown in the void space under the sleepers to attain a lifted geometry (Li *et al.*, 2015). Stoneblowing is criticised for inserting small grain material into the ballast layer; therefore, some consider that it fouls the ballast layer (Lichtberger, 2005). However, if the ballast layer is already fouled, stoneblowing may provide a better alternative for maintenance than tamping. Nevertheless, no stoneblowing machines are currently available in Finland. Adjustment plates are inserted between the sleeper and rail to correct spot geometry defects. The adjustment plates are meant to be only a temporary solution until the track is properly tamped. Ballast adhesion should, in theory, retain the ballast layer in a fixed position and eliminate the need for subsequent tamping (D'Angelo *et al.*, 2016; Jing *et al.*, 2019; Kennedy *et al.*, 2013). Nevertheless, if settlements do occur anywhere in the track structure or in the underlying subsoil after ballast adhesion, the adhesion can obstruct further traditional maintenance, for example, tamping, and novel maintenance methods are required (Jing *et al.*, 2019).

Tamping is not a repair method in the sense that it does not improve the track structure; it only rearranges the ballast material to form a stable base for the sleepers and resets the designed track geometry (Audley and Andrews, 2013; Li *et al.*, 2015). In fact, tamping is known to break the ballast material by crushing the ballast layer rock material, which weakens the condition of the ballast layer (Audley and Andrews, 2013; Kumara and Hayano, 2016). Therefore, unnecessary tamping must be avoided, not only due to unnecessary maintenance costs but also to preserve the ballast layer condition. To avoid unnecessary tamping, the root cause of track geometry deterioration must be realised. This topic will be discussed in the next section.

## 2.3 Causes of track geometry deterioration

Track geometry deterioration can be caused by several track structure features, external sources, or their combinations. The track structure features contributing to track geometry deterioration are categorised as follows:

- settlements in structures below the track
- track structure stiffness variations
- wear and damage in track components.

The main external sources for aggravating track geometry degradation are as follows:

- forces from rolling stock
- forces of nature.

### Settlements in structures below the track

Settlements in the earth structures below the track are reported to be the main cause of track geometry irregularities (UIC, 2008). Ballast layer settlements are often attributed to the rearrangement of ballast material, increased fines contents due to ballast fouling, mixing of structural layers, or fines infiltration from outside sources (Indraratna *et al.*, 2013; Li *et al.*, 2015). Subballast, a layer between the ballast and subsoil, is installed to separate and reduce stress on the subsoil and drain and insulate the track structure (Li *et al.*, 2015). Subballast layers are usually made of granular material and can be reinforced with, for example, bitumen, geogrids, and geomembranes (Biabani and Indraratna, 2015; Ferreira and Teixeira, 2012; Li *et al.*, 2015). Subballast settlements are considered rarer than settlements in other layers, as long as the material used is granular and well compacted (Li *et al.*, 2015). By contrast, the subsoil (or subgrade) is generally considered the weakest and most variable part of the track structure and one that is difficult to repair (Li *et al.*, 2015; Selig and Li, 1994). Especially, fine graded subsoils are considered prone to settlements due to consolidation and repeated loading (Li and Selig, 1995, 1996).

Besides earth structure settlements, damage and settlements in built structures can cause track geometry deterioration (Gou *et al.*, 2019). These structures include mainly bridges and culverts, but other structures, such as pile foundations or retaining walls, can also exhibit similar problems. These structures may be founded improperly resulting in uneven settling and track geometry deterioration (Gou *et al.*, 2019). This problem concerns mainly older structures, as new structures have strict design and construction guidelines for foundations (Liikennevirasto, 2017). In addition to settlements, damage to built structures can cause track geometry irregularities, a problem, which is common, especially with old culverts and bridges. For example, rock and concrete culverts are formed of many pieces (square rocks or concrete rings), which may move in relation to each other. This will result in earth material entering the culvert from the formed cracks causing settlements in the track above as well as issues to the track drainage (Figure 7).

*Figure 7. Concrete ring culvert where the rings have been displaced. Photo credit: FTIA asset management registry (Taitorakennerekisteri).*

Transitions from built structures to embankments are also prone to geometry deterioration because of differential settlements. Some of these settlements may be caused by subsoil consolidation or insufficient structural layer compaction causing consolidation after construction (Li and Davis, 2005). Besides consolidation settlements, stiffness variations have a major influence on transition zone track performance, and they are discussed next.

### Track structure stiffness variations

Track structure stiffness variations can lead to differential track structure settlements (Esveld, 2001; Li *et al.*, 2015; Sañudo *et al.*, 2016). In this dissertation, stiffness variations are considered a distinct phenomenon from track structure settlements, because stiffness variations can cause differential settlements in places where settlements would not otherwise pose a problem. Track structure stiffness varies along a track section due to transitions between different structure types, often referred to as transition zones (Sañudo *et al.*, 2016). For example, bridges provide an almost rigid foundation, whereas embankments built on soft soil are much more elastic (Li and Davis, 2005). Ideally, the track structure should be uniformly and sufficiently elastic as too rigid a foundation may cause excessive stresses to the superstructure, and too elastic a foundation may cause settlements in the earth structures. A uniform deflection value of <0.25 inches or 6 mm is given as a prerequisite for a well-performing track (Li *et al.*, 2015). Current Finnish guidelines determine that the maximum elastic vertical deflection should be <4 mm on structures not founded on peat (Liikennevirasto, 2018). However, smaller deflection values are more typical, for example, <2 mm deflection is common for good condition tracks in Finland (Luomala, 2019).

The main issue in transition zones is the vicious cycle between dynamic loads and plastic deformations. First, the dynamic loads from the rolling stock are increased due to stiffness variations, which causes plastic deformations in the transition zone. Later, the cumulative plastic deformations increase the dynamic loads even further, deteriorating the track geometry much faster than in areas with uniform track stiffness. (Indraratna *et al.*, 2019; Li and Davis, 2005; Sañudo *et al.*, 2016)

Transition zones are difficult to repair because only tamping the track does not usually cure transition zone problems. Instead, recurrent tamping may lead to aggravating the problem, as the rail height on the stiff structure side of the transition is increased and settlements continue on the soft side. An example of the end results of this type of action is demonstrated in Figure 8. A pile-founded rock culvert had been installed on soft soil, causing differential settlements around the culvert. The subsequent geometry faults have been corrected by tamping and adding ballast so many times that the ballast layer could not fit on top of the culvert anymore and ballast material had started to pour over the sides of the culvert. Instead of recurrent tamping, transition zones require much more extensive and costly repairs, for example, reinforced backfills, pile-supported slabs, or elastic superstructure elements (Li *et al.*, 2015; Paixão *et al.*, 2015; Sañudo *et al.*, 2016).

*Figure 8. Poor condition culvert with ballast falling over the edge. Photo credit: Mikko Sauni.*

Besides the structural layers, stiffness variations can be caused by the track superstructure (Andersson and Dahlberg, 2000; Dahlberg, 2010; Li *et al.*, 2015; Lundqvist and Dahlberg, 2005). Discontinuity in the rails can cause impact loads 1.5–3 times higher than the static wheel load (Andersson and Dahlberg, 2000; Li *et al.*, 2015). Discontinuity in the rails and high impact loads are often encountered in rail joints (Figure 9) and turnouts (Andersson and Dahlberg, 2000; Suzuki *et al.*, 2005). Even rail welds cause discontinuity in the rails, as the welded area is softer than the rails and may have a geometrical discontinuity, causing increased dynamic loading to the rails (Li *et al.*, 2015; Messaadi *et al.*, 2021). High impact loads can damage the track components and cause a void space under the sleeper, a phenomenon referred to as hanging sleepers, which further deteriorates the track geometry (Dahlberg, 2010; Lundqvist and Dahlberg, 2005).

*Figure 9. Poor condition insulated rail joint. Photo credit: Mikko Sauni.*

### Wear and damage in track components

Damage to rails, rail fastenings, and sleepers can have direct and indirect influences to track geometry deterioration (Ferdous and Manalo, 2014; UIC, 2002; Williams *et al.*, 2016). Rails can be damaged in several different ways (UIC, 2002). Some rail defects are internal and do not appear in track inspections until the rail breaks (Kumar *et al.*, 2008). Other rail defects appear on the surface of rails, for example, rail squat defects (Figure 10) (Grassie, 2012; Li *et al.*, 2008). These can increase the dynamic forces from rolling stock as the contact surface between the rail and wheel is disturbed (Li *et al.*, 2008). Therefore, many rail surface defects are detectable in track inspections.

*Figure 10. Rail squat defect. Photo credit: Mikko Sauni.*

Rail fastenings and sleepers are also prone to failures noticeable in track inspections (Ferdous and Manalo, 2014; Williams *et al.*, 2016). Poor condition rail fastenings can be noticed in track inspections as gauge widening. Furthermore, loose rail fastenings can induce high impact loads between the rail and sleeper. (Williams *et al.*, 2016)

### Forces from rolling stock

Forces from rolling stock can be static, quasi-static, or dynamic. The static component is the weight of stationary railway stock. Static loads are controlled by assigning allowable axle weights to track sections. In Finland, the allowable axle weights vary from 200 to 250 kN on different track sections, but the is aim to increase the maximum axle weights of certain track sections up to 275 in the future (Väylävirasto, 2020). The quasi-static component includes the uncompensated wheel loads in curves and increased wheel loads due to crosswinds. The dynamic component is the force increase due to abrupt changes in track geometry, discontinuity in the rail running surface leading to wheel motion in relation to the rail running surface, and rolling stock defects, such as wheel flats. (Esveld, 2001)

The dynamic loads can be difficult to estimate as they are dependent on the wheel–track interface and rolling stock speeds and weight (Li *et al.*, 2015). Some studies suggest that dynamic loads can be 1.5–3 times that of static loads (Andersson and Dahlberg, 2000; Li *et al.*, 2015). As track geometry irregularities cause increased dynamic loading, a vicious circle is formed, as mentioned earlier.

The loads from moving rolling stock are not only static nor individual impacts but cyclic. Cyclic loading is applied every time a train axle passes a point in the track. The repetitive nature of the loading is rapid when compared with the reaction from the earth structures to the loads (Li *et al.*, 2015). This means that the settlement and possible excess pore water pressure in the earth structure are not recovered before the next loading, which leads to increased cumulative effects in successive load repetitions (Li *et al.*, 2015), as shown Figure 11.

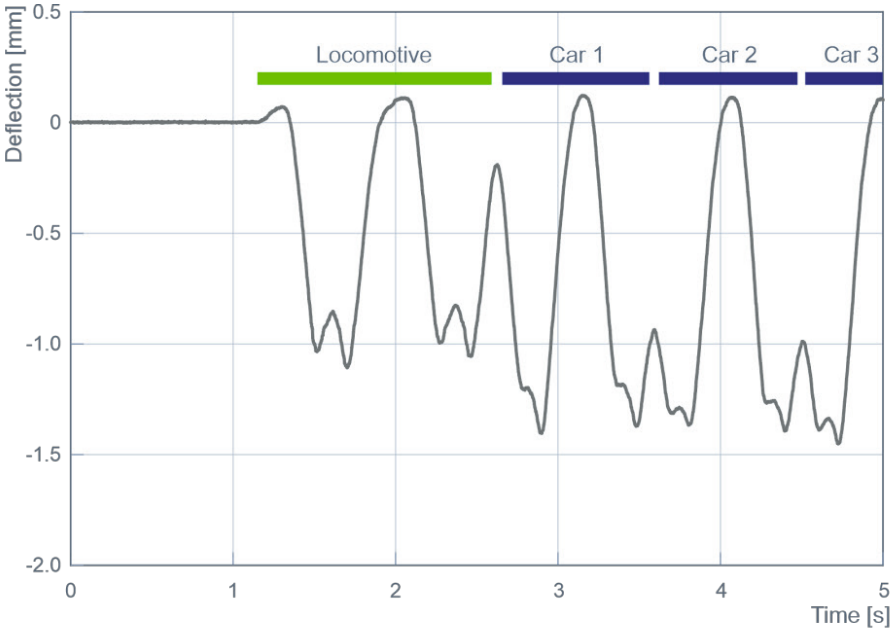


*Figure 11. Cyclic train loading induced deflection measured from a stationary track monitoring station.*

Figure 11 presents a time–deflection graph from a stationary track monitoring station on a Finnish track section between Pori and Mäntyluoto. The graph shows how the second axle from a bogie causes more deflection than the first because the track structure does not have the time to recover from the settlement caused by the first axle. Similarly, the deflection is not recovered before a following car begins to load the structure. The repeated load settlement behaviour is affected by traffic speed, rolling stock axle weight, and axle spacing. The effects of repeated loading can be reduced by lowering the speed or axle weight or by increasing axle spacing (Li *et al.*, 2015).

Besides traffic loads, vibrations from rolling stock can also degrade track structures (Mezher *et al.*, 2016). At very high speeds, usually near or in excess of 200 km/h, train vibrations can cause a shear wave that reaches a critical velocity on soft soils (Krylov *et al.*, 2000; Li *et al.*, 2015). At critical velocity, the train speed exceeds the speed at which the shear wave can advance in the soil which causes a sonic boom and results in ground settlements (Krylov *et al.*, 2000). In addition to these large vibrations, smaller vibrations can occur in the superstructure if the ballast layer has deformed and the sleepers are hanging from the rails. Passing traffic can cause vibrations to the hanging sleepers which can degrade concrete sleepers and loosen rail fastenings (Lundqvist and Dahlberg, 2005).

Lastly, irregular contact in the wheel–rail interface can cause rail faults, which can later manifest into track geometry defects (Barke and Chiu, 2005; Johansson and Nielsen, 2003). Especially, out-of-round wheels cause recurrent high impact loads that damage the track structure (Barke and Chiu, 2005; Johansson and Nielsen, 2003). The effects of the impact loads can be much greater than the allowable static axle loads, as has been researched by, for example, Johansson and Nielsen (2003).

### Forces of nature

In Finnish conditions, the major detrimental forces of nature to railway structures are seasonal frost (heave) and saturation of soils due to insufficient drainage. Frost is a particular feature of cold regions which does not necessarily degrade the structure. In fact, a solid frozen structure is very rigid and durable, more so than the equivalent un-frozen structure. However, frost can lead to two types of detrimental effects: frost heave and thaw-softening. (Andersland, 2004)

Frost heave can have a lifting effect of several centimetres on the track geometry, and, in most severe cases, even over 10 cm (Akagawa *et al.*, 2017; Miao *et al.*, 2020; Pylkkänen and Nurmikolu, 2015). Frost heave is rarely uniform, so the lift is usually noticed as track geometry irregularities. To countermeasure frost heave, Finnish requirements for the thicknesses of non–frost-susceptible track structures are exceptionally high: between 2.0 and 2.6 m depending on how far north the track section is located (Liikennevirasto, 2018). However, old track structures rarely have structures as thick as required for new structures. In these cases, frost insulation boards can be installed underneath the ballast or subballast layer to attenuate frost penetration (Figure 12) (Nurmikolu and Kolisoja, 2005). The downside of frost insulation boards is that the frost insulation boards can alter the stiffness properties of the track structure by increasing track deflection (Luomala *et al.*, 2017).

*Figure 12. Frost insulation board installation in progress. Photo creadit: Mikko Sauni.*

Frost begins to thaw from the top down. This results in a situation called thaw softening, in which the top layer of the track structure has thawed and been saturated with water while the bottom part of the structure is frozen and prevents the water from draining out of the structure (Li *et al.*, 2015). Consequently, excess pore water pressure can build up in the track structure under repeated loading which can degrade the track geometry at a considerable pace. Similar detrimental effects caused by trapped water in the track structure can result from heavy rains and poor drainage conditions (Latvala *et al.*, 2016). Track structures on low embankments or in cuttings are especially prone to suffer from drainage issues (Latvala *et al.*, 2016).

## Ensemble of causes for track geometry deterioration

Finding the exact reason or reasons for an irregularity in track geometry is a difficult task, because there are several reasons for failure, as mentioned earlier. This task becomes even more challenging as deterioration can progress from a failure of one component to the next. For example, the wheel spin of a stationary locomotive might cause wheel–burn damage to the rails. The damage to the rails increases the dynamic loads of all passing trains, which increases loads to the ballast layer and eventually fouls the ballast aggregate. The fouled ballast layer can result in unsupported sleepers increasing the stiffness variations of the track, thus increasing track settlements. As track maintenance does not have access to continuous data streams regarding the condition of the track structure, the first observation of this chain of events may be a track geometry irregularity noticed in the track recording car measurements, although the fault did not originate from the track structure.

Moreover, neither the condition of the infrastructure nor the forces aggregating degradation are the only possible reasons for track geometry deterioration. Trackwork and routine maintenance can cause problems to track geometry unintentionally. For example, the installation of a new interlocking system requires constructing new cabling routes which means drilling pipes through the track embankment for the cables. If the embankment is disturbed during the drilling and settlements around the pipe occur, the deformation can be seen in track recording car measurements.

Considering the varying, and to some degree even random, reasons for track geometry deterioration, determining its root causes requires a tremendous amount of information regarding the track structure, the rolling stock, and maintenance history. Yet, even if all this information is obtained, figuring out the causality of the deterioration requires rigorous analysis to form a timeline of the deterioration. Therefore, all parts of track sections with deteriorated track geometry cannot be rigorously investigated to find their root causes in practice, as there is simply not enough time or resources to do so. Some obvious problems can be observed and identified, but only the track geometry is generally monitored, and irregularities are mostly addressed by means of tamping if no other obvious reasons are observed. Therefore, track geometry deterioration is the most important source of information in railway maintenance and asset management. Consequently, to effectively investigate track geometry deterioration, its modelling is very important.

## 2.4 Track geometry deterioration modelling

Track geometry deterioration is not an inherently random process but one that can be idealised and modelled (Esveld, 2001; Lichtberger, 2005). Certainly, it must be recognised that as demonstrated in section 2.3, there are many reasons why track geometry deteriorates and some of them may be random at times. Nevertheless, when completely random events are disregarded, the track geometry deterioration resulting from consistent traffic follows a theoretical path presented in Figure 13 (Esveld, 2001; Jovanovic, 2004; Lichtberger, 2005).

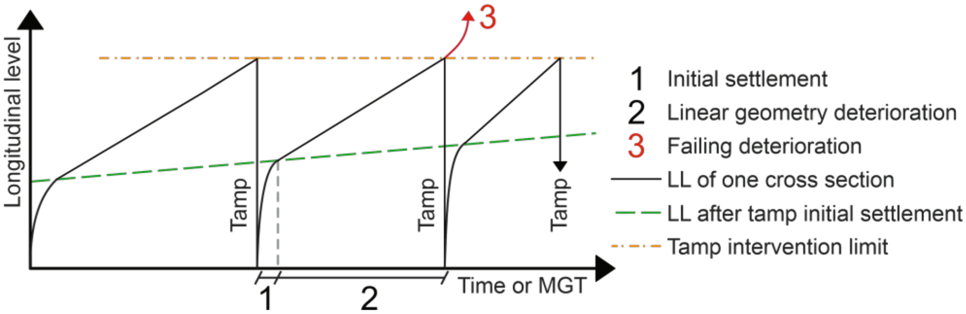


*Figure 13. Theoretical track geometry deterioration (Paper I).*

Theoretical track geometry deterioration is usually based on the ballast settlement behaviour (Dahlberg, 2001; Lichtberger, 2005). As demonstrated in Figure 13, the first stage in theoretical track geometry deterioration is the initial settlement after construction or maintenance (Lichtberger, 2005). At this stage (1), the tack geometry deteriorates at a high pace, usually described in logarithmic or exponential behaviour, for a short period of time (Jovanovic, 2004; Lichtberger, 2005; Soleimanmeigouni, Ahmadi and Kumar, 2018). The initial settlement is sometimes referred to as ballast memory (Lichtberger, 2005). This is followed by the linear deterioration stage (2) until a maintenance intervention limit is reached (Jovanovic, 2004; Lichtberger, 2005; Soleimanmeigouni, Ahmadi and Kumar, 2018). If the tracks are not maintained when the maintenance limit is reached, the deterioration, caused by dynamic load increases due to track irregularities, may begin to grow at a considerably faster pace (3) (Lichtberger, 2005). Theoretically, the amount of initial settlement and the slope of the linear regression will increase with each maintenance cycle, thus reducing the maintenance interval length and track availability. When the maintenance interventions become too frequent for efficient track use, the track structure must be renewed. After a renewal, the deterioration/maintenance cycle (Figure 13) is reset and reinitiated.

There are two main methodologies for modelling track geometry deterioration: **mechanistic** and **statistical** (Higgins and Liu, 2018; Soleimanmeigouni *et al.*, 2018). The main difference between the two can be coarsely described as follows: Mechanistic models use generalising formulae to create data imitating track geometry deterioration, whereas the statistical models describe actual measurement data on track geometry deterioration.

**Mechanistic** models are usually derived from a combination of mechanical models and soil models, usually depicting track structure settlements (Dahlberg, 2001; Ferreira and Murray, 1997; Soleimanmeigouni, Ahmadi and Kumar, 2018). To clarify, railway settlements are not synonymous with track geometry deterioration, but they both describe much of the same phenomenon (Esveld, 2001). The mechanistic models always produce the same results when initial data is constant. Therefore, the models require immense research on past behaviour for creating the formulae producing the results. Nevertheless, many mechanistic models have been created for railroad settlements (Dahlberg, 2001; Soleimanmeigouni, Ahmadi and Kumar, 2018). Traditional mechanical models implement initial parameters (such as traffic speed and volume, rail weight, and soil parameters) and apply them in generalising formulae (Dahlberg, 2001; Ferreira and Murray, 1997). Examples of such models are the Japanese settlement model (Sato, 1995) and the integrated track degradation model (ITDM) (Zhang *et al.*, 2000). More recent mechanical models have adopted modern tools in modelling the deterioration behaviour, for example, finite element models (FEM) (Indraratna *et al.*, 2019; Kalliainen *et al.*, 2016; Li *et al.*, 2016; Paixão *et al.*, 2021; Suiker and de Borst, 2003). These models can consider many features, such as three-dimensional loading (Kalliainen *et al.*, 2016), cyclic loading (Suiker and de Borst, 2003), transition zones (Indraratna *et al.*, 2019; Paixão *et al.*, 2021), and heterogeneous track characteristics (Li *et al.*, 2016).

Mechanistic models can be considered theoretical, as they do not depict specific measured behaviour of certain locations. Instead, the output is based on the generalised history of several locations (Soleimanmeigouni, Ahmadi and Kumar, 2018). Therefore, mechanistic models are best suited for studying different hypothetical structure types or loading conditions, for example, when planning new structures or increases to the allowable axle loads of current structures. Nevertheless, for asset management, these generalisations are not as helpful, as asset management is highly pragmatic and requires models depicting the actual measured deterioration of specific track locations, not generalisations based on past research (Soleimanmeigouni, Ahmadi and Kumar, 2018).

**Statistical** track geometry deterioration models are primarily based on time series data from track recording car measurements, which can be coupled with static asset data (Soleimanmeigouni, Ahmadi and Kumar, 2018). Many different statistical models have been applied to track geometry deterioration modelling, as presented in literature reviews by Elkhoury *et al.* (2018), Higgins and Liu (2018), and Soleimanmeigouni *et al.* (2018). The statistical models can be divided into simple and complex models. Simple models include deterministic models that apply curve fitting to time series data, whereas more complex models include, for example, artificial neural networks (ANNs), whose goal is to imitate past data patterns with great detail and investigate the influences of different features on the deterioration behaviour (Elkhoury *et al.*, 2018; Soleimanmeigouni, Ahmadi and Kumar, 2018).

Simpler statistical models are usually selected based on their goodness-of-fit, which can be parametrised with different approaches, for example, the (root) mean squared error. A typical approach is to fit linear models to maintenance intervals, as has been demonstrated by, for example, Andrade and Teixeira (2015), Audley and Andrews (2013), Caetano and Teixeira (2016), Vale and Lurdes (2013), Khajejhei *et al.* (2019), Lee *et al.* (2020), Li *et al.* (2019), Soleimanmeigouni *et al.* (2020), and Neuhold *et al.* (2020). Other studies have found exponential models to fit their data better, for instance, Famurewa *et al.* (2016) and Quiroga and Schnieder (2012).

More complex statistical methods include the use of different stochastic models such as the Gamma process, Markov process, classification methods, ANNs, Bayes models, and various data mining methods (Andrade and Teixeira, 2012; Bai *et al.*, 2015, 2016; Guler, 2014; Meier-Hirmer *et al.*, 2006; Mercier *et al.*, 2012). The Gamma process is used to estimate track geometry deterioration rates while considering the uncertainty regarding the deterioration (Meier-Hirmer *et al.*, 2006; Mercier *et al.*, 2012). Markov models have been utilised in predicting track deterioration based on four track irregularity categories (Bai *et al.*, 2015). ANNs have been applied in providing deterioration rates based on input data concerning the track features (such as the geometry, rail type, and sleeper type) and traffic conditions (traffic loads and speed) (Guler, 2014). Bayesian models have been used to predict the development in deterioration rates and assess the uncertainty of the predictions for different types of track sections: switches, bridges, stations, and plain tracks (Andrade and Teixeira, 2012). All these statistical models were used to create complex patterns from intricate data, whereas a mechanistic approach provides only generalisations.

Choosing the track geometry deterioration modelling approach is based on the available data and the desired information. Mechanistic models are best suited when designing new structures or comparing remediation methods to current structures, as these models produce estimations on deterioration even without measurement data from the actual structure. Contrarily, if the goal is to investigate how the track geometry is currently deteriorating and how maintenance is affecting it, statistical models and measurement data concerning the actual structure should be used. Therefore, the organisation modelling track geometry deterioration must clarify what the aims of the modelling are and what type of data is available. In the case of Finnish railway asset management, track geometry deterioration modelling has not yet been implemented. Therefore, the first steps were to identify and apply suitable modelling methods for the available data. These studies are reported in Paper I and section 3.1 of this dissertation.

## 2.5 Railway asset data analytics

Railway asset data analytics can be difficult to differentiate from track geometry deterioration modelling. In fact, track inspection history is railway asset data and deterioration modelling is data analytics. However, for clarity, this dissertation separates railway asset data analytics from track geometry deterioration modelling with the following definition: track geometry deterioration modelling investigates **how** the track geometry deteriorates, whereas railway asset data analytics investigate **why** the track geometry deteriorates. This definition leaves room for some research approaches to be considered either one of these or even both.

In practice, railway asset data can be analysed using available railway-secific data analysis software. For example, IRISSYS by ERDMANN Software, Optram by Bentley, RAMSYS by MERMEC, GeoEdit by ENSCO and Rail Doctor by Loram introduce the possibility for an asset manager to access their data easily. These programs are specialised in linear railway asset data analysis and deterioration analysis. In addition to practical asset management, railway data has been analysed in several scientific studies.

Some studies have investigated the reasons for track geometry deterioration from track inspection history alone. These types of studies have investigated the different wavelengths of track geometry defects using fractal analyses (Hyslip, 2002; Landgraf and Hansmann, 2019). Landgraf and Hansmann (2019) demonstrated that different wavelengths of the longitudinal level (LL) deviation are correlated with defects in different parts of the structure. More specifically, the mid-wave fractal range was found to correlate with ballast condition, and the long-wave fractal range correlated with the load bearing capacity, which implies the strength of the lower structural layers (Landgraf and Hansmann, 2019).

Other studies have investigated the causes of track geometry deterioration using additional data sources together with track inspection data. These studies include the use of railway **track stiffness** measurements*,* **GPR** measurements, and data on **track assets**.

Studies researching the effects of railway **track stiffness** on track geometry deterioration using data analytics include Grossoni *et al.* (2019), Megrali *et al.* (2020), and Nielsen *et al.* (2020). Grossoni *et al.* (2019) used log-linear models to assess the effects of vertical track stiffness variability on the deterioration of track geometry. Their models showed that support stiffness has the greatest influence on ballast layer settlement and deterioration rate (Grossoni *et al.*, 2019). Mehrali *et al.* (2020) applied classification, decision trees, clustering, and filtering data mining to investigate the correlations between track geometry and stiffness data. The study found many correlations between track geometry parameters and support stiffness. For instance, clustering results indicated the track alignment irregularities had the strongest correlations to vertical displacements, and filtering techniques produced a high correlation between the LL and stiffness index (Mehrali *et al.*, 2020). Nielsen *et al.* (2020) also investigated the correlations between track stiffness and geometry deterioration and found a strong correlation between low substructure stiffness and a high rate of track geometry deterioration.

Similarly, correlations between **GPR** measurements and track geometry deterioration have been investigated by, for instance, Sussmann *et al.* (2003), Scanlan *et al.* (2018), and Yurlov *et al.* (2019). Sussmann *et al.* (2003) found that in more than half of their test sites GPR data could be used to indicate substructure problems causing track geometry deterioration. Especially, problematic conditions in the substructural layer moisture and thickness could be interpreted from the GPR measurements (Sussmann *et al.*, 2003). Yurlov *et al.* (2019) created a probability model for track geometry defects utilising substructure parameters derived from GPR measurements. Yurlov *et al.* (2019) found that ballast fouling indices and the ballast layer thickness have strong correlations to track geometry irregularities. Conversely, Scanlan *et al.* (2018) did not find a large-scale correlation between track roughness (geometry quality) and ballast deterioration indices calculated from GPR measurements. However, Scanlan *et al.* (2018) mentioned that these results may be due to local variability in the ballast deterioration calculation outputs not corresponding to actual ballast condition. GPR measurements are used not only to investigate track geometry deterioration itself but also in many other indicators of track condition. For example, ballast and substructure condition and frost susceptibility have been investigated using GPR measurements (Silvast *et al.*, 2010, 2013).

**Track asset** data has also been used in investigating track geometry deterioration (Andrade and Teixeira, 2010; Sadeghi and Askarinejad, 2009). Andrade and Teixeira (2010) used asset data on turnouts, bridges, stations, and plain track to investigate whether the track sections with these assets exhibited differences in maintenance cycles. Their analyses were conducted using Monte Carlo analyses in the @risk software (Andrade and Teixeira, 2010). Andrade and Teixeira (2010) found that turnout areas and bridges have shorter maintenance cycles compared with plain track and stations, implying that areas with turnouts and bridges require maintenance more often than plain track or stations. Sadeghi and Askarinejad (2009) compared different defect densities of rails, sleepers, fastenings, and ballast with observed track geometry irregularities. They concluded that gauge deviations correlated with rail and fastening condition, whereas alignment deviations were correlated with sleeper and ballast condition, and the profile was influenced by ballast condition (Sadeghi and Askarinejad, 2009). However, variability in the results was rather high.

As presented above, past research shows many correlations between different data sources and track geometry deterioration. For example, low stiffness, high GPR signal attenuation, and turnouts have been individually linked to high track geometry deterioration (Andrade and Teixeira, 2010; Nielsen *et al.*, 2020; Sussmann *et al.*, 2003). However, as the cause of deterioration can be attributed to several sources, investigating the causes for deterioration in practice can be difficult when using only one data source at a time. Therefore, combining several data sources is very important for investigating all possible causes of deterioration. Unfortunately, combining all data available data sources make the analyses more complex (Berggren, 2010). Guler et al. (2011) investigated the causes of track geometry deterioration from multivariate data, revealing correlations between individual data sources and track geometry deterioration. However, investigating the effects of combinations of several different parameters was not in the scope of research by Guler et al. (2011); thus, the synergy of different features remained unknown. The synergic effects of several data sources can be subjectively analysed for individual spots on the track section, for example, by visualising all the data simultaneously. However, investigating the overall correlations of different structure types and track geometry deterioration becomes exhausting when done subjectively. Berggren (2010) approached the problem of investigating the root causes of track geometry deterioration from multivariate data using pattern recognition. A team of specialists categorised problematic track sections into three combined classes representing problems that were 1) rail-related, 2) ballast-related, or 3) soil- or embankment-related. With ample data from multiple sources, a pattern recognition algorithm was trained using specialist evaluations and the algorithm could indicate problem areas based on the training data. The research showed how the identification of known problem types can be performed using mathematics. (Berggren, 2010)

However, sometimes the problem with vast multivariate data is that we are not sure what we should even be looking for. We may have unexplored data and we do not know whether it contains something interesting or not. This was the case in Finland, where accumulated multivariate data had been used for visualisations and subjective assessments, but the correlations within the data seemed too complex to resolve by human efforts. More specifically, there were more than 20 attributes concerning different track structure features that were available, which makes statistical analyses rather complex. Therefore, explorative data mining methods were studied for investigating the causes of track geometry deterioration from multivariate data in Papers II and III and section 3.2 of this dissertation.

## 2.6 Maturity models

The implementation of track geometry deterioration modelling and data analytics into practical railway asset management requires development in the technological and organisational capabilities of the asset management organisation. Depending on the initial capabilities of the organisation, this development might take several years to succeed. This huge amount of work can seem overwhelming if attempted all at once. Therefore, the development process should be strategically planned and divided into smaller segments that provide incremental goals. Otherwise, the development might become too ambitious and fall under its own weight, as has happened to many software development projects in the past (Errida and Lotfi, 2021).

Maturity models were developed for managing large software development projects (Humphrey, 1988; Paulk *et al.*, 1993). The maturity models describe different maturity levels for a defined process, thus helping organisations control their development projects (Paulk *et al.*, 1993). The capability maturity model (CMM) established by Paulk et al. (1993) can be considered one of the original maturity models, from which several derivatives have been created for different purposes (Albliwi *et al.*, 2014; CMMI Product Team, 2010; Goncalves Filho and Waterson, 2018; Helgesson *et al.*, 2012; Herbsleb and Goldenson, 1995; Poeppelbuss *et al.*, 2011; Röglinger *et al.*, 2012).

Maturity models have been applied in the railway sector in asset management, safety, and security development. The International Union of Railways (UIC) (2016) has published a practical guide for International Organization for Standardization (ISO) 55001 asset management implementation, in which asset management maturity is depicted using a six-level maturity model. Fonseca and de Almeida Júnior (2005) extended the Capability Maturity Model Integration (CMMI) (CMMI Product Team, 2010) to incorporate European standards normalising railway system reliability, availability, maintainability, and safety (RAMS). Kour *et al.* (2019) applied a maturity model for assessing the cybersecurity capabilities of railway organisations.

The basic idea of maturity models is to identify organisational maturity and set sensible goals for process improvement (ISO 33004, 2020; Paulk *et al.*, 1993). This is achieved by creating a sequence of levels for maturity (Paulk *et al.*, 1993; Röglinger *et al.*, 2012). The CMM includes five maturity levels depicted in Figure 14 (Paulk *et al.*, 1993).

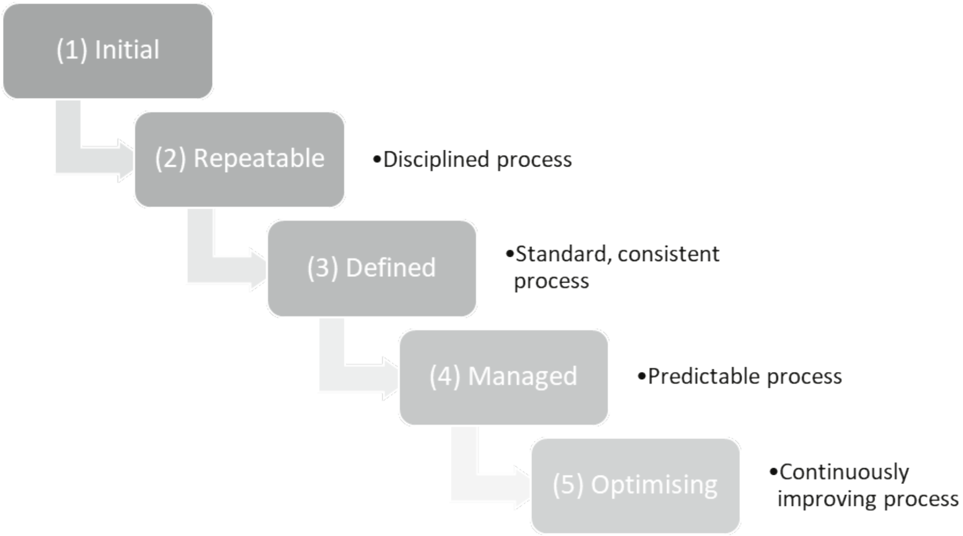


*Figure 14. CMM by Paulk et al. (1993).*

In the initial level (1) of the CMM (presented in Figure 14), processes are devised ad hoc during a project and there are no formal procedures or standard tools. The organisations or projects that operate on the initial maturity level tend to exceed their budgets and schedules, and the end-product quality is not consistent or predictable. However, projects or organisations at this level can produce successful results, but they are usually possible only due to employee heroics and competent management, which, if lost due to personnel changes, will demolish the success. (Humphrey, 1988; Paulk *et al.*, 1993)

The success of an organisation or a process on the repeatable level (2) is based on past experiences with similar tasks. As long as the workflow, tools, and products remain constant, past experiences will suffice for creating a viable product. However, if a new product has to be produced, new tools have to be implemented, or the process is otherwise disrupted, the process may fail and cannot be re-established until experience with the new way of working is gained. (Humphrey, 1988; Paulk *et al.*, 1993)

The defined level (3) differs from the previous level in that processes are defined and documented (Humphrey, 1988; Paulk *et al.*, 1993). However, although data concerning the process can be produced, its analysis is not common practice. Subsequently, in the managed level (4), an organisation will begin to analyse its processes based on the collected data and sets goals for products and processes (Humphrey, 1988; Paulk *et al.*, 1993). Lastly, in the optimising level (5), the organisation can identify its weakest links and address them to fully optimise their operations (Humphrey, 1988; Paulk *et al.*, 1993). The optimising level (5) can be summarised as the level of continuous improvement (Paulk *et al.*, 1993).

The process of creating a maturity model has been described in previous research (de Bruin *et al.*, 2005; Maier *et al.*, 2012). Maier *et al.* (2012) described the process in four stages, namely, 1) planning, 2) development, 3) evaluation, and 4) maintenance, whereas de Bruin *et al.* (2005) used six stages, namely, 1) scope, 2) design, 3) populate, 4) test, 5) deploy, and 6) maintain. The planning, scoping, and design phases have many similarities, in that the background of why and how a maturity model should be created is established. After planning, the model is developed and populated which is followed by deployment and evaluation. Finally, the maturity model should be maintained to evaluate its effects and set further goals.

Creating maturity models requires in-depth knowledge of current processes and future goals. This knowledge must be acquired from the industry experts using surveys, interviews, and workshops, especially, when the purpose of the maturity model is to raise awareness and improve performance (de Bruin *et al.*, 2005; Maier *et al.*, 2012). Some previous examples of gathering information concerning railway asset management practices include studies by Al-Douri *et al.* (2016), Dadashi *et al.* (2014), and Schraven *et al.* (2011). These studies approached data gathering using semi-structured interviews. Besides semi-structured interviews, workshops as research methods are especially attractive in railway asset management as they have been reported to be effective ways of communicating complex topics to a wide audience with very different backgrounds (Phaal *et al.*, 2007). The stakeholders in track geometry deterioration modelling include maintenance personnel, asset managers, and maintenance directors; therefore, universally understandable communication is paramount.

The perspective of maturity models can be descriptive, prescriptive, or comparative (de Bruin *et al.*, 2005). Solely descriptive models only describe the current maturity of an organisation without providing plans for improvement. A prescriptive model focuses on indicating how improvement in maturity could increase business and enables future development roadmapping. Comparative models are for benchmarking practices across organisations. These different perspectives can be considered different phases of one maturity model, as the use of the maturity model must begin with describing the current maturity of an organisation before comparing or advancing it (de Bruin *et al.*, 2005).

Some of the most crucial aspects for successful maturity model implementation have been identified to be senior management commitment, staff involvement, process clarity, and training (Herbsleb and Goldenson, 1995; Niazi *et al.*, 2005). Critical barriers preventing from realising the benefits of maturity models are (detrimental) organisational politics and a lack of awareness, support, formal methodology, or resources (Niazi *et al.*, 2005).

Maturity model evaluation is required to validate the model for practical use (de Bruin *et al.*, 2005; Helgesson *et al.*, 2012; Maier *et al.*, 2012). Helgesson *et al.* (2012) present three types of evaluation: 1) evaluation only by the authors of the maturity model, 2) evaluation involving practitioners, and 3) evaluation through practical case use. Evaluation or validation is required to obtain feedback from the stakeholders before the actual implementation of the maturity model to find the weak points and improve on them (Helgesson *et al.*, 2012).

The benefits of maturity models have been studied, and it has been found that the vast majority (86%) of organisations that have adopted maturity models, have found maturity models valuable in setting future goals for improvements (Herbsleb and Goldenson, 1995). Other claimed benefits include increased project performance, reduction in fault rates, and decreases in costs (Jiang *et al.*, 2004; Pitterman, 2000). Moreover, although maturity models aim to define and regulate processes, organisations have reported that maturity model implementation has actually reduced the amount of bureaucracy in higher maturity organisations (Herbsleb and Goldenson, 1995).

Regardless of the above-mentioned benefits, maturity models are not free of criticism. The criticism includes the difficulty of using maturity models based solely on their documentation, oversimplification of complex processes, oftentimes a lack of empirical background, and a large amount of only conceptual research (Albliwi *et al.*, 2014; Herbsleb and Goldenson, 1995; Niazi *et al.*, 2005; Poeppelbuss *et al.*, 2011). One of the major criticisms concerns the use of maturity models; it is believed that maturity models can identify only what needs to be done and not how it should be implemented (Albliwi *et al.*, 2014; Herbsleb and Goldenson, 1995; Niazi *et al.*, 2005). While it is true that maturity models include mostly information on what to do next, it must be recognised that a maturity model cannot be the only vehicle and driver for organisational change. Maturity models provide a way to formulate an achievable plan, but it should be up to the organisation to manage the change.

Change management is a whole branch of science in itself, which is not in the scope of this dissertation except for maturity models. In the research for this dissertation, the maturity models were used to roadmap future development by creating a development framework based on the maturity model. It is acknowledged that only the maturity model or the framework will not actualise the change within Finnish railway asset management. But the research for creating the maturity model is believed to provide a basis, from which the change can begin. The work on utilising maturity models for implementing track geometry deterioration modelling and data analytics into Finnish railway asset management is reported in Paper IV and section 3.3 of this dissertation.

# 3 MAIN RESULTS AND DISCUSSION

This section summarises and discusses the main results obtained in the studies reported in Papers I–IV. Section 3.1 presents the methods for modelling track geometry deterioration using Finnish data and utilising the results in practical asset management (Paper I). Section 3.2 elaborates on the means for investigating the root causes of railway track geometry deterioration (Papers II and III). Lastly, section 3.3 provides an outlook on the implementation of track geometry deterioration modelling and data analytics for practical asset management (Paper IV).

## 3.1 Track geometry deterioration modelling using Finnish data

The track recording car history provides a unique representation of what the condition of geometry has been at certain past instances when measurements have been performed. This history can be utilised as is for some condition analyses, like locating recurrent faults (Esveld *et al.*, 1988). However, because the measurements represent only intermittent moments in history, it is difficult to elaborate on the deterioration progress using only the measurements. For effective and objective analyses, the track geometry deterioration process must be idealised with mathematical models. This modelling is possible because track geometry deterioration is not an inherently random process but one that can be idealised (Esveld, 2001; Esveld *et al.*, 1988; Lichtberger, 2005). However, as elaborated in section 2.3, there are a great number of reasons that affect the way track geometry deteriorates. All the complexity and randomness in track geometry deterioration make it virtually impossible to reliably consider every influencing factor in the models. For example, no model can predict sudden extreme weather conditions or mistakes made in track maintenance. Therefore, a famous aphorism, generally credited to George Box, applies well here: *All models are wrong, but some are useful.* The task was, then, to find the most useful track geometry deterioration models for railway asset management with the least detrimental deficiencies.

### 3.1.1 Modelling process description

**Choosing the modelling method**

A statistical approach, robust linear optimisation, was selected for track geometry deterioration modelling of the Finnish state-owned rail network. The use of robust linear optimisation was based on the available data and the desired outputs. The available data included the track recording car measurement history. The goal was to obtain parameters describing the deterioration process, which meant choosing a statistical model that best describes the data. Choosing statistical methods over mechanistic was based on the domain, asset management. Mechanistic models are better suited for design and research purposes, as they provide only generalised information on the track performance, not the actual condition of structures (Soleimanmeigouni, Ahmadi and Kumar, 2018). A simple statistical model was preferred over complex models because the initial data and its behaviour could be elaborated using simple equations. Furthermore, it is almost impossible to observe the reasons why a complex model has produced its results, for example, in the case of neural networks. Thus, as the results are to be used in practice, it is important that the end results can be fully elaborated to verify the validity of the results.

The historical LL standard deviation (SD) data from Finland was found to follow a linear deterioration path (Paper I). As an example, the 10 year track geometry history from the LUIMA track section was modelled using both linear and exponential models. Within the track geometry history, there were some cross-sections, for which a reasonable exponential function could not be approximated due to fluctuation in the measurement results. When these cross-sections were ignored, the coefficient of determination ($R^2$) was found almost identical between linear and exponential modelling, with linear models having slightly higher $R^2$ values (Figure 15). In Figure 15, the boxplot demonstrates that the meadian and 25th and 75th percentiles are very similar, as are the outliers presented in red. There are many outliers presented in the boxplot because the data includes $R^2$ values from 682,233 maintenance cycles with varying deterioration trends.

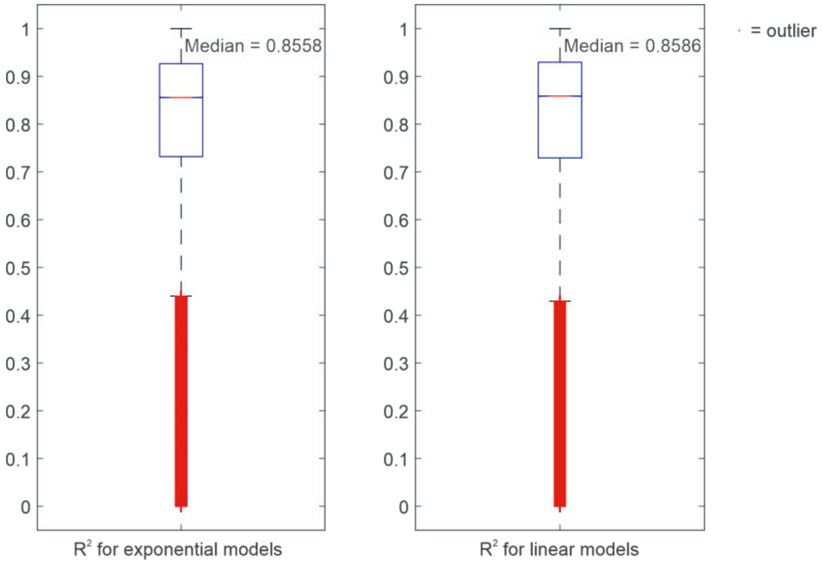


*Figure 15. LUIMA track section R$^2$ values for exponential and linear models.*

Obviously, the R$^2$ is not the only way to determine model suitability, and it was considered that because some of the data was unfeasible to model using exponential models, which did not depict the data more accurately, linear models were found to be better alternatives. Linear models also offer practical benefits, for example, using the slope of linear regression as the deterioration rate is an intuitive solution that can be communicated easily to a wide audience.

Similar to the decision to use linear deterioration modelling in Finland, many other researchers have based their models on linear deterioration paths (Andrade and Teixeira, 2015; Audley and Andrews, 2013; Caetano and Teixeira, 2016; Khajehei *et al.*, 2019; Lee *et al.*, 2020; Li *et al.*, 2019; Neuhold *et al.*, 2020). However, some have found that exponential models suit their data better (Famurewa *et al.*, 2016; Quiroga and Schnieder, 2012), especially when initial settlements are highlighted in the measurement data, for example, when the measurement frequency is high. It may be necessary to revisit evaluating exponential models in Finland, if the track geometry measurement frequency is increased, for example, by using measurements from in-service vehicles.

Robust linear optimisation was chosen over other linear regression methods because the data occasionally exhibited initial settlements after tamping (Paper I). Although recorded initial settlements were not common, when observed, simple linear models, for example, least squares, can overestimate the long-term deterioration trends because of the initial settlements (Paper I). The robust linear optimisation of a tamping cycle is not affected by the initial settlement as greatly as simple linear models are, making robust linear optimisation appropriate for long-term track geometry deterioration modelling (Paper I). Figure 16 demonstrates how in the second tamping cycle (2011–2017) the least squares algorithm is affected by the initial settlement, whereas linear optimisation characterises the long-term behaviour better. It should be recognised that the initial settlements are not outliers in the sense that they should not be removed from data. Recorded initial settlements provide valuable information when calculating, for example, the maintenance effects; however, for track geometry deterioration modelling, their influence should be limited.

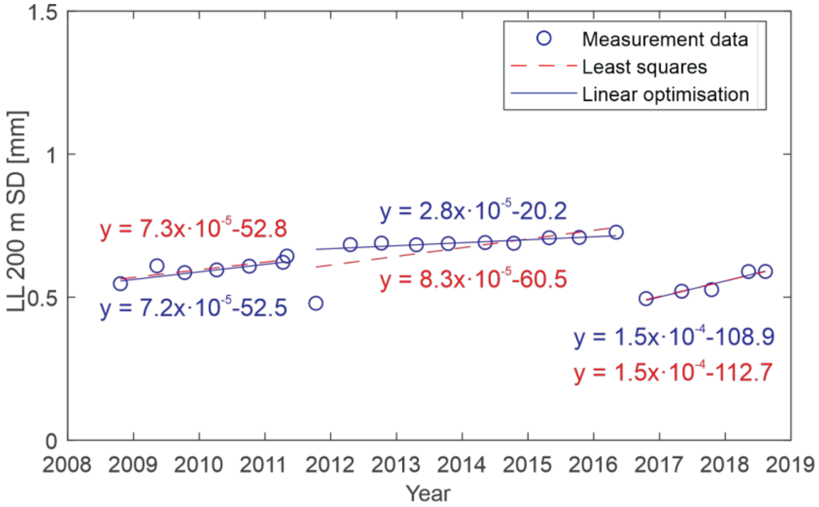


*Figure 16. Difference between linear optimisation and least squares in one cross-section (Paper I).*

The downside of robust linear optimisation is a longer calculation time compared with least squares, for instance. However, the calculation times are still reasonable when considering the measurement frequency. The calculation time for a 10 year biannual track geometry measurement history of the 65-km-long LUIMA track section was 1,5 h for robust linear optimisation versus 87 s for least squares. Calculations were performed with an office PC with an Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz and 32 GB of RAM. The most recent tamping cycle must be recalculated when new measurements are conducted, which means that recalculations should be done every two months on the most frequently measured lines. Considering this measurement frequency, even the longer calculation times are acceptable. However, if more frequent measurements were to be conducted, for instance, daily measurements, the importance of calculation times and the use of more powerful computers should be reassessed.

### Modelling process

Figure 17 presents the chosen track geometry deterioration modelling process. The deterioration modelling begins with acquiring time series data from track recording car measurements (Figure 17, Step 0). Of the parameters produced by the track recording car, the short-wave LL has been the most popular choice for track geometry deterioration modelling (Audley and Andrews, 2013; Khajehei *et al.*, 2019; Neuhold *et al.*, 2020; Nielsen *et al.*, 2020). This is because the settlements causing short–wave defects in the LL drive the need for track maintenance, especially tamping (Soleimanmeigouni *et al.*, 2020; UIC, 2008). Furthermore, some parameters like twist and cant can be considered derivatives of the LL, depicting the LL unevenness of one rail with the other rail as reference. However, although the modelling in this study concerned only the LL, other parameters following a linear deterioration path can be modelled with the presented methodology without reservation (Paper I). The LL 200 m SD is usually used for analytics and modelling, as the original LL signals are difficult to align and interpret due to signal fluctuation (Khosravi *et al.*, 2021; Neuhold *et al.*, 2020). The SD was calculated in sliding windows in Finland to provide the highest level of detail regarding the track geometry deterioration.

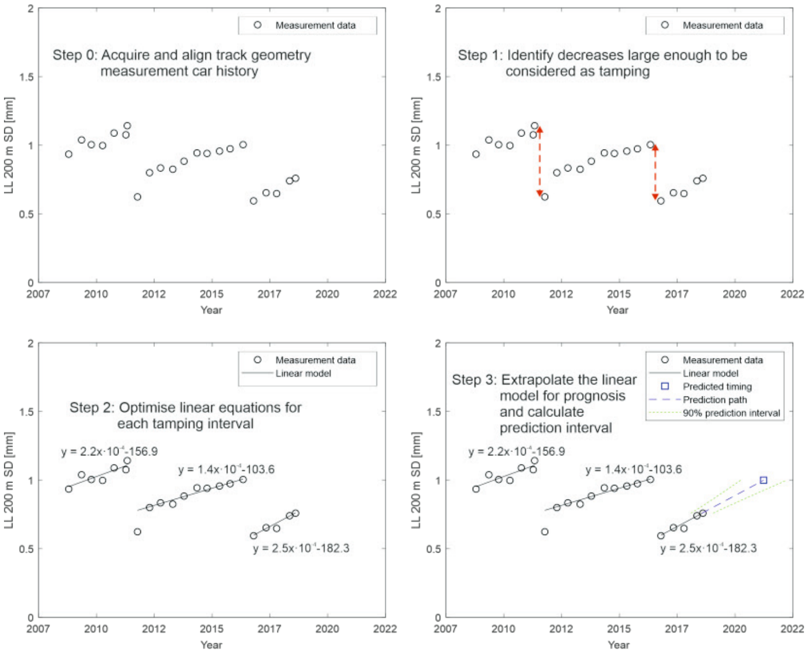


*Figure 17. Track geometry deterioration modelling process.*

Aligning different measurement runs' signals is perhaps the most important phase in track geometry deterioration modelling, as the reliability of any modelling result is based on correct input data. The differences in the alignment of different measurement runs' signals are inevitable, as the track recording car wheels can slip, slide, or lock up, causing differences in consecutive measurement run lengths (Wang *et al.*, 2018). Past research on signal alignment has shown many methods for alignment (Fellinger, 2020; Khosravi *et al.*, 2021; Tanaka *et al.*, 2018; Wang *et al.*, 2018; Xu *et al.*, 2015). These studies have shown how signals can be aligned with higher accuracy than the measurement interval (commonly 0.25 m), meaning that the measurements are practically perfectly aligned (Fellinger, 2020; Khosravi *et al.*, 2021; Li *et al.*, 2019; Tanaka *et al.*, 2018; Xu *et al.*, 2015). The track inspection data used for the studies reported in Paper I was aligned by Loram Finland Oy (formerly Roadscanners Oy) before delivering the data. Nevertheless, the importance of aligning the data is recognised, and there is ongoing research on different methods for aligning the data in Finland.

Before the actual deterioration modelling, track geometry measurement data must be divided into tamping cycles, meaning the time periods between maintenance actions (Figure 17, Step 1). This could be done easily with maintenance records if reliable repositories were available. However, because comprehensive repositories do not exist in Finland throughout the complete measurement history, maintenance actions had to be identified from the measurement history, as has been the case in similar previous studies (Audley and Andrews, 2013). Past maintenance can be identified as a significant decrease in the LL SD values. However, not all decreases in the LL SD values are caused by maintenance. Other major reasons for decreasing LL SD values are inaccuracy in the measurement results and frost thawing. The measurement inaccuracy accounts for only small, random decreases in the LL SD. Frost thawing can decrease the LL SD values if there has been significant uneven frost heave and track geometry measurements have been conducted at that time. These incidents can cause more significant decreases in the LL SD, but they generally do not concern very long areas. Therefore, a threshold is required to separate measurement inaccuracy and frost thawing from past maintenance (Paper I). The threshold must be used along with an algorithm that evaluates whether the area, where the LL SD has decreased, forms a logical tamping area. With the threshold and algorithm, it is possible to back-calculate the past maintenance action timings (Paper I). To clarify, the back-calculated past maintenance timing does not reveal the actual timing, but the measurement interval when maintenance was performed. Therefore, if the measurement interval is two months, the maintenance could have occurred on any given day within that timeframe. This causes uncertainty when estimating the numerical effects of maintenance. Moreover, the back-calculated maintenance timing does not express the types of performed maintenance actions. Furthermore, the algorithm accuracy could not be verified as there were no reliable maintenance records to use for the verifications. This is a source of future research, as comprehensive maintenance action reporting systems have been implemented into Finnish railway asset management in recent years. In the future, there will be enough maintenance data for testing the algorithm validity.

Next, the tamping cycles are modelled with the chosen modelling method, robust linear optimisation (Figure 17, Step 2). The track section is modelled one cross-section at a time, meaning that if track geometry data is collected every 25 cm, the modelling is done separately for every 25 cm (Paper I). The cross-section models are independent of one another making the modelling very responsive to abrupt changes in the deterioration behaviour along a track section.

Finally, the linear models created from the latest tamping cycles are used to make predictions on future maintenance timing (Figure 17, Step 3). The predictions are simply extrapolations of the latest tamping cycles (Paper I). However, these alone are not informative enough, as the confidence of the predictions must also be calculated. The prediction confidence is calculated using the prediction interval (PI) (Paper I). The PI depicts the range in which future observations should occur at a given confidence level (Agresti, 2015). The PI will be wide after maintenance when there are only few observations, on which to base the linear optimisation. As the amount of data within a tamping cycle increases, the PI narrows.

### Modelling outputs

As presented in Paper I, the outputs of the modelling include the following:

- the track geometry deterioration rate (TGDR)
- long- and short-term past maintenance effects
- predictions of the next maintenance timing and prediction confidence.

The TGDR is the slope of the deterioration regression within a maintenance cycle. This can be calculated in either the time dimension (mm/a) or the cumulative tonnage dimension (mm/MGT). When the traffic volumes are almost constant on a track section, the time dimension is preferred, because it is easier to communicate. For example, maintenance timing is more intuitive to communicate in time rather than in tonnage. However, if yearly tonnages have high fluctuation, the deterioration trend may not be linear if calculated in mm/a. In these cases, the deterioration rate must be calculated and presented in mm/MGT. The TGDR is not a standardised parameter, and there is no established method for attaining it. Nonetheless, some TGDR values presented in past research are collected in Table 1 for comparison. The deterioration rates in Table 1 concern the LL 200 m SD except for Audley and Andrews (2013) whose deterioration rates concern the very similar LL 220 yard SD. The TGDR values in Table 1 were converted to mm/MGT if enough information was available in the referenced articles for doing so.

*Table 1. Different TGDRs reported in the literature.*

| Reference | Reported TGDR | TGDR converted mm/MGT |
|---|---|---|
| (Andrade and Teixeira, 2010) | 0–8 mm/100 MGT (plain track) | 0–0.08 mm/MGT |
| (Audley and Andrews, 2013) | Mean 0.000259 mm/d = 0.095 mm/a | N/A |
| (Caetano and Teixeira, 2016) | 0–10.5 mm/100 MGT | 0–0.105 mm/MGT |
| (Khouy *et al.*, 2014) | 0–0.14 mm/MGT | 0–0.14 mm/MGT |
| (Li *et al.*, 2019) | –0.1–0.2 mm/month (approximate range)<br>0.03 mm/month (approximate mean)<br>Passing annual tonnage of 245 million tons | –0.005–0.01 mm/MGT (approximate range)<br>0.0015 mm/MGT (approximate mean) |
| (Soleimanmeigouni *et al.*, 2020) | 0–5.5 mm/year (plain track)<br>0.75 mm/year (plain track approximate mean)<br>Passing tonnage 20 MGT/year | 0–0.275 mm/MGT (range)<br>0.04 mm/MGT (plain track approximate mean) |

The TGDRs reported in the literature vary between roughly 0 and 0.3 mm/MGT. In Finland, the TGDR for LL 200 m SD was observed to be roughly between 0 and 0.5 mm/MGT (Figure 18). On a relatively good condition track section (LUIMA) with 11 MGT/a passing tonnage, the mean TGDR was 0.0074 mm/MGT. On a track section that is considered more problematic (Rantarata) and with 2.6 MGT/a passing tonnage, the mean TGDR was 0.0855 mm/MGT. The difference in these track sections' mean deterioration rates was tenfold, when calculated in mm/MGT. Even in deterioration rates in mm/a, the difference was almost threefold. However, it must be considered that the Luumäki–Imatra track section is 65 km long, whereas Karjaa–Ervelä is 29 km long. On short sections, the problematic areas are highlighted in percentual statistics, for example, the mean TGDR. Nevertheless, when compared with other research (Table 1), the Finnish TGDR values are only moderately different. However, this may be caused by modelling the deterioration of every cross-section in Finland as opposed to modelling 200 m segments as the smallest unit in some of the other studies. Modelling every cross-section increases the amount of significantly low and high TGDRs as the fixed 200 m windows can average out short sections of very low or high values. Furthermore, the different TGDRs found in the literature may imply different climate, loading, and structural conditions, or differences in the track recording cars. It is apparent from Table 1 and Figure 18 that the TGDR unit is not universal nor is the scale of the TGDR values. However, many of the TGDR figures from past research are within a similar range despite the differences. Nevertheless, the lack of consensus on the definition of the term TGDR and the lack of international statistics on the topic must be emphasised. These should be sources of future research and international collaboration. Moreover, the reasons why the TGDR values are so different on different track sections and different parts of the same track section should be investigated. These studies are reported in Papers II and II and section 3.2 of this thesis.

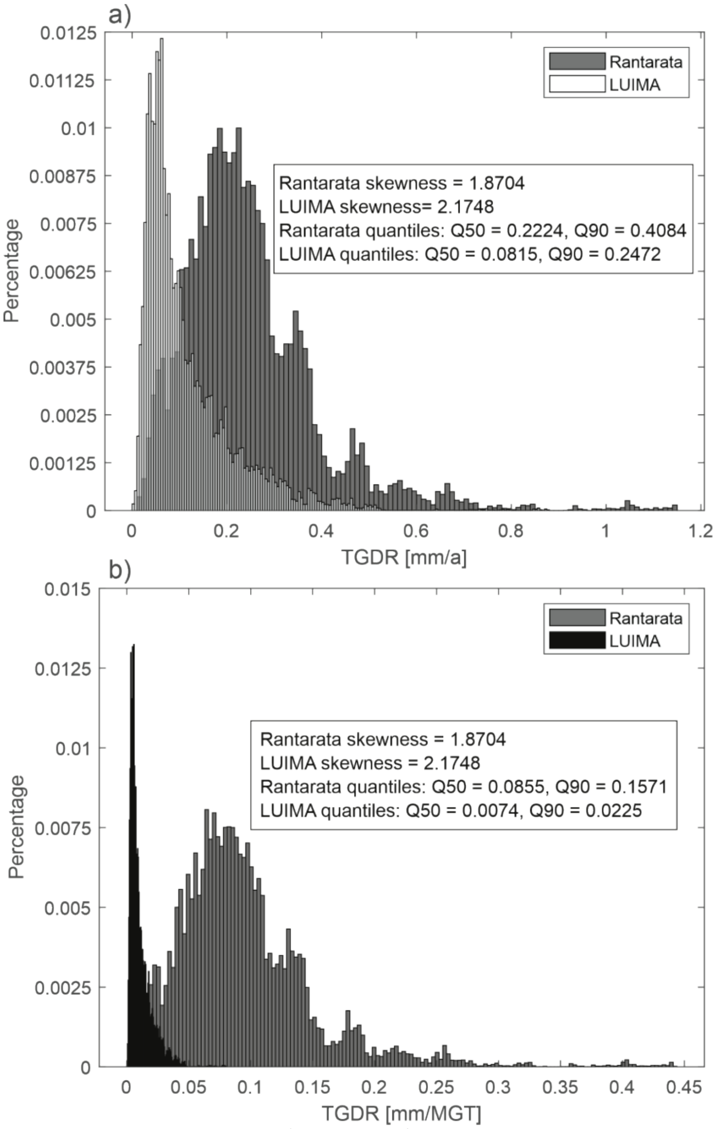


*Figure 18ab. TGDR of two track sections in a) mm/a and b) mm/MGT.*

The effects of past maintenance can be examined in the short and long term (Paper I). The short-term maintenance effect is the immediate improvement in track quality achieved by tamping. This is denoted by, for example, a decrease in the LL SD after tamping. This number indicates, above all, the purposefulness and success of maintenance. If the short-term effect is low, maintenance has not had a significant effect on the geometry quality. This can be attributed to either good quality geometry before tamping, ineffective tamping, or acute deterioration after tamping. The long-term maintenance effect is evaluated by comparing the TGDRs before and after tamping (Paper I). An increased TGDR after tamping implies deterioration in the track structure that cannot be repaired solely with tamping. Contrarily, if the TGDR after tamping is decreased or invariable when compared with the TGDR before tamping, it implies that maintenance can improve or sustain the durability of the track geometry condition, respectively.

Assessing the overall maintenance effect requires simultaneously examining both the short- and long-term maintenance effects as well as the prevailing condition before and after tamping (Paper I). However, it is impractical to subjectively examine multiple parameters at once. Therefore, an ensemble parameter, maintenance success indicator (MSI), was developed (Paper I). The MSI denotes whether maintenance has had 1) beneficial effects, 2) delaying effects, 3) no meaningful effects, or 4) negative effects on track geometry deterioration behaviour (Paper I). The MSI is based on the logical induction of four parameters: tamping effect on the TGDR and LL SD as well as the TGDR and LL SD after tamping (Figure 19).

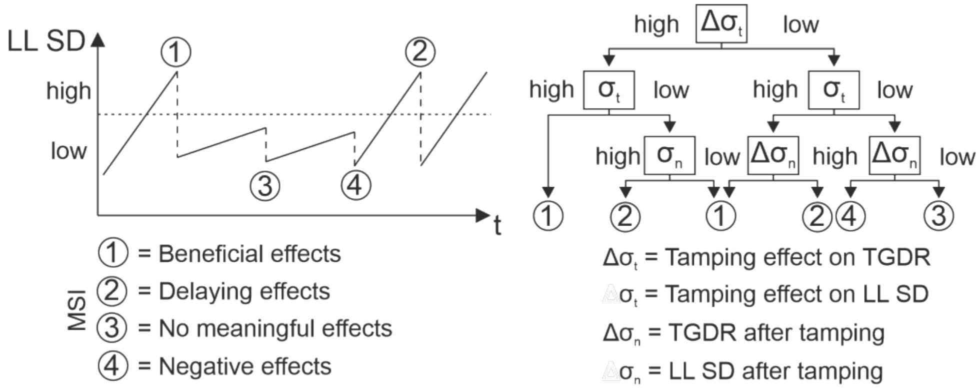


*Figure 19. Maintenance success indicator (MSI) computation (Paper I).*

The assessment of MSI begins with evaluating the tamping effect on the TGDR. A decreased TGDR after tamping (high effect) is considered a positive outcome, as it means that the deterioration has slowed down. Next, the tamping effect on the LL SD is evaluated. Again, a high value (high effect) is desirable, because it means that the tamping has made the track quality better than before. Lastly, depending on the previous two evaluations, the LL SD after tamping or the TGDR after tamping is evaluated. A low LL SD and TGDR after tamping are considered positive outcomes. The values separating what is considered a high or a low effect are a source of future research, as this type of investigation will require vast data regarding track deterioration and track maintenance from multiple track sections (Paper I).

Previous research has also investigated the modelling of maintenance effects (Audley and Andrews, 2013; Li *et al.*, 2019; Martey and Attoh-Okine, 2018; Soleimanmeigouni *et al.*, 2018). These studies have modelled the recovering effects of maintenance based on the LL SD before maintenance. Soleimanmeigouni *et al.* (2018) reported the LL SD values before tamping and the recovery in the LL SD achieved by tamping. These values vary from 0.4–2 mm for LL SD before tamping and –0.1–1.2 mm LL SD for tamping recovery. Recovery values for the LL SD of two Finnish track sections mentioned earlier, LUIMA and Rantarata were calculated to be between 0 and 1.7 mm.

However, further investigations on maintenance effects could not be reliably performed from the Finnish data, because it is not known whether past maintenance has included only tamping or also other repairs. All typical track superstructure maintenance, such as ballast cleaning or rail renewal, require tamping afterward. These types of maintenance actions have a profoundly different effect on the track geometry behaviour compared with tamping alone (Audley and Andrews, 2013). Additionally, the tamping strategy influences the maintenance effects, as tamping complete areas (also referred to as line tamping) has been found to improve the track condition more than tamping partial areas (also referred to as spot tamping) (Soleimanmeigouni *et al.*, 2018). Therefore, data on the types of maintenance actions is required before maintenance effect modelling can be reliably calculated in Finland. Fortunately, the collection of such data has already started, and it is believed that in a few years, it will be possible to make these types of analyses. Additionally, fractal analyses of past track inspections could provide information on whether some structural remediations have been made (Landgraf and Hansmann, 2019). Fractal analyses should also be considered in future research on maintenance effects on track geometry in Finland.

The track geometry deterioration modelling is also used to predict the next time maintenance is required (Paper I). However, track maintenance planning includes much more diverse information than track geometry deterioration alone, for example, the age of components, decay in components not reflected in track geometry, and available track downtime for repairs. Consequently, track geometry deterioration predictions are not maintenance plans but only initial data for creating such plans. Therefore, the implementation of deterioration predictions into practice must be controlled to prevent maintenance personnel from considering the predictions as ready-to-use maintenance plans. The implementation of track geometry deterioration modelling into practical asset management is discussed further in section 4.3.

### Synthesis

When initially selecting the modelling method, the goal was set to find the most useful track geometry deterioration model for railway asset management in Finland with the least detrimental deficiencies. The main benefits of the chosen modelling method, robust linear modelling, were as follows:

- Only a few track inspections are enough to satisfy the minimum initial data required for modelling.
- Modelling is based on linear behaviour, which has been verified to depict the track geometry deterioration behaviour well in multiple environments that have similar traffic and track geometry measurement processes.
- Modelling closely imitates the actual past deterioration path of individual locations on a track section, responding to spatial variability along a track section.
- Modelling results are intuitive for practical utilisation; the TGDR, maintenance effects, and predictions can be easily visualised.

The practical use and benefits of the modelling are further discussed in section 4.1.2.

The selected modelling process has its limitations, that is, the lack of data around the time of maintenance and the amount of data required for making predictions (Paper I). The deterioration process continues after the last periodical measurement before maintenance and continues after maintenance before the next measurement. However, no data from this period is obtained with periodical track inspections. Hence, the modelled maintenance effects are affected by the timing of maintenance in relation to track inspections. Therefore, the true maintenance effects are not captured, but rather the long-term effects of the maintenance. This limitation has little impact on the deterioration modelling itself but it means that the maintenance effects cannot be considered to be absolutely accurate. Moreover, maintenance effects should be estimated after a few measurements have been performed since the latest maintenance to capture the long-term effects of maintenance.

Furthermore, the accuracy of the predictions is dependent on the available data from the current tamping cycle. Therefore, predictions made soon after tamping are generally inaccurate. However, this should not pose issues in practice, because the goal of tamping is to reset the deterioration cycle, postponing the need to plan the next tamping. High LL SD values in the inspection after tamping are not a desired outcome, as maintenance effects are supposed to be durable. Therefore, deterioration modelling should not be the only way of monitoring rapid deterioration after maintenance. Automatic alerts should be created based on LL SD values after tamping to flag occurrences, where deterioration rates or LL SD values are exceptionally high after tamping.

To summarise, the most notable limitations of robust linear modelling were the following:

- Maintenance effect calculations are based on measurements whose timing varies relative to the maintenance.
- Prognosis accuracy is highly dependent on the amount of data since the latest maintenance cycle.

The limitations of maintenance effects cannot be overcome without using additional measurement data. The prognosis accuracy could be improved if previous tamping cycles were considered in the calculations (Andrade and Teixeira, 2014; Andrews et al., 2014). However, this is not advisable in Finland until reliable maintenance records are available, as the previous tamping may have been accompanied by other maintenance actions, which greatly influence the subsequent deterioration rate. Considering that the modelling is used for investigating the condition of track structures and long-term maintenance planning, the practical and intuitive outputs of the robust linear optimisation modelling process are thought to outweigh its limitations.

Future research on track geometry deterioration modelling should include the verification of the tamping identification algorithm, fractal analyses of past track inspections, investigations of the MSI limit values, and defining of the TGDR on an international level (Paper I). The tamping identification algorithm can be verified soon, as there are enough recorded tamping actions in maintenance data repositories. Additionally, maintenance effects can be calculated more accurately, once the timing of maintenance is known. This will also enable investigations of the effects of different maintenance actions. Fractal analyses can also bring new information regarding past unrecorded maintenance actions, as these can reveal whether structural remediations have been made. The MSI limit values should be investigated using track inspection histories from multiple different track sections and environments together with respective maintenance records. Lastly, the term TGDR should be defined and multiple different operating environments and track types should be researched for investigating their differences. This will require wide international collaboration of different infrastructure owners sharing their data.

### 3.1.2 Practical use of modelling results

Track geometry deterioration modelling results should benefit practical asset management. However, the models and the produced numerical results are not easily accessible to the asset management personnel, as it is not within their domain to be experts in mathematical modelling. Therefore, providing only the models and numerical modelling results will not make a tangible change in practical asset management. Therefore, research was focused on making the modelling results accessible for asset management personnel using data visualisation (Paper I).

The visualisations were constructed to serve different user needs and depict different modelling outputs. The three output categories were past deterioration exploration, maintenance effectiveness, and maintenance predictions. These output categories were visualised for three different use-cases: 1) cross-section, 2) track section, and 3) network-level modelling results analysis (Figure 20) (Paper I). The cross-section level analysis is intended for investigating the track geometry behaviour of individual spots on a track section, for example, bridge transitions, turnout areas, and rail joints. The track section level analysis presents information on time–location axes, which enables visualising longer (e.g., 1–10 km) sections of track in one figure. However, even these visualisations have their limitations, as subjectively examining the time–location figure of, for example, 100 km of track becomes exhaustive. Therefore, a network-level analysis is required to provide illustrations for entire track sections or even track networks. The network-level illustrations do not exhibit the specific locations of problematic or good condition sections of track but on their quantity instead, which is important to know when considering the amount of resources required in the future.

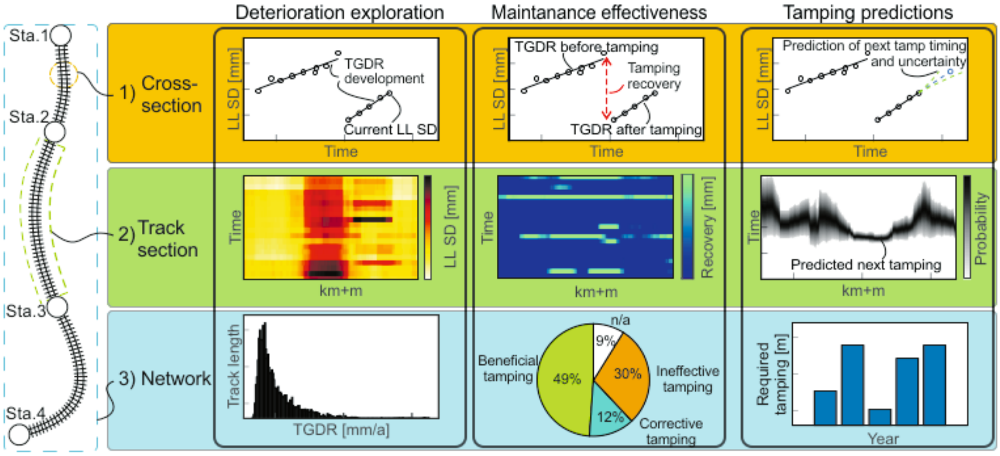


*Figure 20. Visualisations of track geometry deterioration modelling (Paper I).*

All three levels of observation have an intuitive meaning and a function in practical railway asset management (Paper I). The cross-section level caters to the needs of maintenance personnel, who require information on the history and prognosis of specific locations' conditions to make decisions on spot maintenance. The track section level is useful for long-term maintenance planning, for example, when planning line tamping or superstructure renewal areas. The network-level benefits general asset management, in which the condition of whole track sections and the entire network is to be monitored.

The practical use of track geometry deterioration modelling results is not widely discussed in past research. Some of the illustration concepts from Figure 20 are used in articles to present initial data or modelling results (Caetano and Teixeira, 2016; Famurewa *et al.*, 2016; Neuhold *et al.*, 2020; Nielsen *et al.*, 2020; Soleimanmeigouni *et al.*, 2020). Nevertheless, their implementation in practice has not been in the scope of past research. By producing easily accessible and categorised visualisations of track geometry deterioration modelling results, the gap between research and practice is narrowed and the tangible benefits of research can be obtained.

## 3.2 Investigating root causes of track geometry deterioration

Defective track structures must be identified and restored to avoid unnecessary repetitive maintenance and track downtime. Defective structures can be located using track geometry deterioration modelling presented in section 3.1. However, these models rarely reveal the reasons for deterioration, which must be known for choosing the correct maintenance methods. Hence, this section explores ways to produce knowledge on the root causes of observed track geometry problems. This is approached by mining available railway asset data.

### 3.2.1 Exploratory railway asset data mining

The goal of exploratory data mining is to explain patterns found in data (Larose and Larose, 2015). The data pattern explained in this study was the occurrence of high TGDRs. Track asset data was used for creating the explanations. The initial data included the attributes presented in Table 2. The initial railway asset data was constructed so that each row in the data depicted a one-meter-long section of track, which was described by the columns containing the different attributes (Paper II).

*Table 2. Initial data and parameters for data mining.*

| **Data source** | **Data attribute** | **Data type** |
|---|---|---|
| Track recording car measurements | TGDR | Ratio |
| GPR | Track structural layer thicknesses and moisture indices | Ratio |
| Continuous laser scanning | Embankment width and ditch depth | Ratio |
| Continuous track deflection measurement | Track deflection mean and variance | Ratio |
| Soil maps | Subsoil type and frost susceptibility | Categorical |
| Asset data warehouse | Track asset (bridge, turnout, etc.) location and type | Categorical |
| Maintenance data | Tamping history | Categorical |
| Video and visual assessment of data | Foundation type and asset data corrections | Categorical |

The data mining method selected for exploring the railway asset data was General Unary Hypotheses Automaton (GUHA). The GUHA method is a descriptive data mining method that generates hypotheses based on the initial data and user questions (Hájek *et al.*, 2010). The hypotheses are statements regarding correlations found within the initial data. The initial data used in GUHA can be versatile, as GUHA supports, for instance, text, numeric, binary, and categorical data (Berka, 2016). This is possible because GUHA is based on binary logic, in which the initial data is categorised and binary values present the categories the values belong to (Hájek and Havránek, 1978). The initial data used in GUHA can also contain partly contradictory or deficient data, as the GUHA logic allows obtaining results with varying confidence (Turunen, 2018). This is a major benefit for the railway domain, as it is common that not every attribute of every track structure is included in the initial data.

GUHA is implemented in dedicated computer software, LISp-Miner (Berka, 2016; Rauch, 2013; Turunen, 2018). The basic idea of using LISp-Miner is to form question about the initial data and search for hypotheses to those questions using GUHA. The generic process of using LISp-Miner in GUHA data mining is demonstrated in Figure 21.

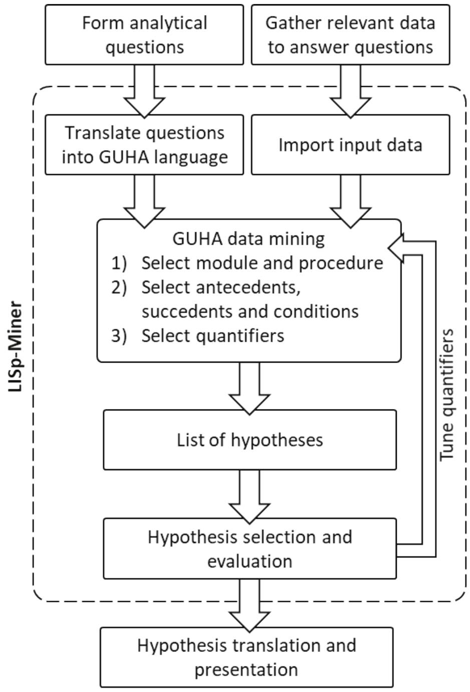


*Figure 21. GUHA data mining process using LISp-Miner.*

First, the user must compose analytical questions and then gather relevant data that can be used for answering said questions. The types of questions asked in data mining are intuitive, for example, "Are there attributes in my data set that almost always exclude the occurrence of some outcome?"

After forming the initial analytical questions and gathering relevant data, these are translated and imported to LISp-Miner, respectively. The question formation in LISp-Miner begins with choosing a module and procedure (Berka, 2016). Different modules and procedures are for solving different tasks (Berka, 2016). For example, the 4ft-Miner can make insightful descriptions of concepts found in data, whereas the MCluster-Miner can segment the data into meaningful subgroups (Berka, 2016). When the module is chosen, the questions are translated into GUHA language using antecedents ($\varphi$), succedents ($\psi$), conditions ($\gamma$), and quantifiers (Berka, 2016). The antecedent, succedent, and condition can be any attribute, category of one attribute, or group of attributes from the initial data. The occurrences of these attributes are evaluated using quantifiers that assess their correlations using a contingency table. Different LISp-Miner modules can have varying contingency tables (Rauch, 2013). The general form of the 4ft-quantifier contingency table is presented in Table 3. The contingency table parameters indicate the number of rows from the initial data $a + b + c + d = n$, whose relationship is evaluated as follows:

- $a$ is the number of rows that satisfy both $\varphi$ and $\psi$,
- $b$ is the number of rows that satisfy $\varphi$ but not $\psi$,
- $c$ is the number of rows that satisfy $\psi$ but not $\varphi$,
- $d$ is the number of rows that do not satisfy $\varphi$ nor $\psi$ (Turunen, 2018).

*Table 3. Contingency table for 4ft-quantifier.*

| $\gamma$ | $\psi$ | $\neg\psi$ | $\sum$ |
|---|---|---|---|
| $\varphi$ | $a$ | $b$ | $r$ |
| $\neg\varphi$ | $c$ | $d$ | $s$ |
| $\sum$ | $k$ | $l$ | $n$ |

Quantifiers assessing the contingency table have an intuitive meaning. For example, the above average dependence quantifier $\sim^{+}_{q,Base}$ tests a condition:

$$\frac{a}{a+b} \geq (1+q)\frac{a+c}{a+b+c+d} \wedge a \geq Base, \tag{1}$$

in which *a*, *b*, *c*, and *d* are contingency table frequencies; *q* is the user-defined dependence; and *Base* is a quantifier assessing the number of occurrences in the defined contingency table slot (Berka, 2016). This can be translated to an intuitive form: Is there a subgroup, with more than the *Base* amount of data, in which the antecedent and succedent are satisfied at least $q + 1$ times more often than the antecedent is satisfied in the whole data set? This can be used to investigate whether some track features occur much more frequently on sections with high TGDR than the whole data on average.

After selecting the module, attributes, and quantifiers, GUHA data mining is run, which means the algorithm searches for hypotheses satisfying the preconditions. The result of GUHA data mining is a list of hypotheses. The hypotheses are answers to the user's analytical questions. These hypotheses satisfy the conditions laid in the question formation. If the question formation has been too general, there will be an enormous number of hypotheses, which will be exhaustive to investigate further and most of the hypotheses will probably be trivial answers. Contrarily, if the question formation has been too strict, there will not be any hypotheses. Both outcomes should be avoided; the user should aim to obtain a reasonable number of hypotheses to investigate further. This amount can be considered between 10 and 100 in typical cases. Usually, the GUHA data mining process must be iterated by tuning the quantifiers to obtain a reasonable number of relevant hypotheses. When an adequate number of hypotheses is obtained, the user can begin to investigate the contents of the hypotheses further, selecting one hypothesis at a time. A hypothesis can be evaluated using the contingency table and the data that satisfied the preconditions. Examples of evaluating interesting hypotheses can be found in section 3.2.2 and Papers II and III.

The results of data mining should be reacted to with deliberation, as the results only denote correlation, not causality or statistical significance. These are for the users of data mining methods to investigate with other methods. Causality cannot be investigated from static data as there is no temporal dimension. Currently, the only available temporal railway data was track inspection time series data, and other data was static. Static data does not generally enable causality investigations. For example, if track geometry problems are encountered on a track section with frost insulation boards, but the data does not reveal whether the problems have occurred prior to the boards' installation, causality cannot be determined. Causality investigations using asset data would require much more rigorous asset data warehousing, which, fortunately, has been ongoing for a few years already in the

Finnish state railway asset management. However, causality investigations from asset data will require many years' data, but the important first step of collecting the data has already been taken to enable these investigations in the future. Statistical significance is also important to investigate whether the results from one or more track sections are to be generalised to the whole network. However, it must be kept in mind that statistically significant phenomena may not be practically significant (McShane *et al.*, 2019). Instead of relying only on statistical testing, the underlying phenomenon explaining some data patterns related to railway track geometry deterioration should also be approached with experimental (soil) mechanical research. Especially in investigating the root causes of track geometry deterioration, the role of data mining is to show the way forward for future experimental research, when there are too many influencing factors and combinations for testing all possible instances.

### 3.2.2 Practical use cases for railway asset data mining

#### Practical aspects of GUHA data mining

The use of data mining in practice is rather conflicting; data mining methods are typically very complex to use, but they solve pragmatic issues. Therefore, data mining should be considered a precision instrument rather than an all-purpose tool. Mining railway asset data should be focused on revealing intricate correlations within extensive and diverse data sets. Therefore, simple tasks should not be approached with data mining methods. Tasks that are too simple for data mining include investigating the correlations between only a few parameters or investigations of isolated occurrences in data. Data mining can provide answers to these too, but it is too robust a method for the task, and it would take significantly more time and effort than, for example, data visualisation.

As long as the initial data is versatile, success in data mining is achieved by asking interesting questions. Some generic examples of suitable question formats are as follows:

- Which subsets (almost) always exhibit similar behaviour?
- Which subsets behave very closely to the average behaviour?
- Which subsets behave very differently from the average behaviour?

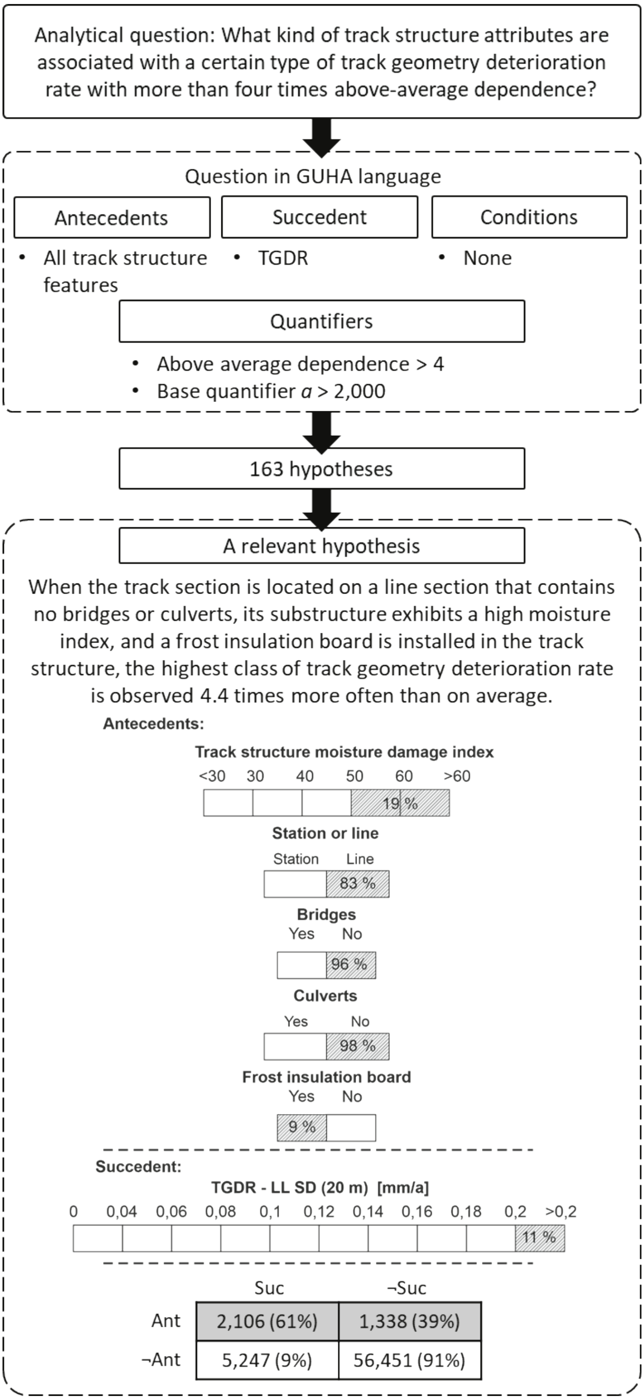
- Do some parameters have a significant correlation to some behaviour when the data is limited to a certain subset?
- Can a change in one or only a few parameters have a major effect on the behaviour of some subset?

These types of questions do not directly approach investigating a research question, like the root causes of track geometry deterioration. Instead, these questions narrow down the (combination of) attributes that are associated with certain types of outcomes. This information is then used to deduce the relevant attributes explaining the behaviour. Several different types of questions are required to approach the main research question from different angles. Paper II discusses the different modules and quantifiers used for investigating the root causes of track geometry deterioration. In summary, these were the following:

- The 4ft-Miner module with the p-implication quantifier, which can be used to characterise the typical behaviour of certain track structure types. For example, on the Kouvola–Kotka track section, a high TGDR was observed on 87% of structures built on low embankments exhibiting high track deflection mean and variance (Paper III).
- The 4ft-Miner module with the above-average quantifier, which can be used to reveal extraordinary correlations between track structure types and track geometry deterioration behaviour. For example, on the Luuamäki–Imatra track section, a high TGDR was observed 4.4 times more often than on average on track structures located on line sections without bridges or culverts, substructure moisture indices were high, and a frost insulation board had been installed in the structure (Paper II).
- The AC4ft-Miner module with the p-implication quantifier, which can be used to investigate, which attributes have a dominant effect on the behaviour of some track structure types. For example, on the Luumäki–Imatra track section, when the track moisture index is very high and the number of past tamping actions is low, a high TGDR is observed on 79% of the structures where a frost insulation board has been installed and on 14% of track sections where no frost insulation board has been installed (Paper II).

### An exemplary data mining task

One of the data mining tasks reported in Paper II is presented in Figure 22 and discussed further here to elaborate and illustrate the use of the GUHA method. The data mining task began with an analytical question: What kind of track structure attributes are associated with a certain type of track geometry deterioration rate with more than four times above-average dependence? The available initial data concerned the Luumäki–Imatra track section and contained the attributes presented in Table 2. The analytical question was translated into GUHA language using the 4ft-Miner module and above average dependence quantifier, which was elaborated in Equation 1. The Base quantifier was set above 2 000, which meant that the resulting hypotheses must be supported by at least 2 000 rows of data (or metres of track). The four-times above average quantifier meant that the resulting hypotheses' correlation between cases, in which the antecedent and succedent are satisfied, must be at least four times that of the frequency the succedent is satisfied in the whole data. The antecedents could be any class or any combination of classes from all track structure features, whereas the succedent could be any class or combination of classes of the TGDR. No conditions were applied.



Analytical question: What kind of track structure attributes are associated with a certain type of track geometry deterioration rate with more than four times above-average dependence?

When the track section is located on a line section that contains no bridges or culverts, its substructure exhibits a high moisture index, and a frost insulation board is installed in the track structure, the highest class of track geometry deterioration rate is observed 4.4 times more often than on average.

| | Suc | ¬Suc |
|---|---|---|
| Ant | 2,106 (61%) | 1,338 (39%) |
| ¬Ant | 5,247 (9%) | 56,451 (91%) |

*Figure 22. A GUHA data mining task and its results.*

The data mining task resulted in 163 hypotheses, which were presented on a list containing the antecedents, succedents, and strengths of the hypotheses. From this list, one relevant hypothesis was chosen for closer inspection. This hypothesis stated that *When the track section is located on a line section that contains no bridges or culverts, its substructure exhibits a high moisture index, and a frost insulation board is installed in the track structure, the highest class of TGDR is observed 4.4 times more often than on average in the whole data* (Paper II). Figure 22 visualises the classes selected in the hypotheses. The percentages within the shaded areas indicate the amount of data within the class, for example, 83% of the entire length of the track section is located on line sections instead of stations.

The correlation between the antecedents and the highest class of TGDR was 61%, whereas the highest TGDR class is observed on 11% of the whole track section. In GUHA terminology, this denotes a 4.4 time above average dependence. Therefore, the hypothesis' antecedents have a substantially higher correlation to the highest TGDR class than all structures overall. Of the antecedents, line sections, bridges, and culverts only slightly limit the data, as most of the track section is located on a line section, and there are few bridges and culverts. Nevertheless, the high moisture damage indices and frost insulation boards limit the data to a much greater extent; only 19% of structures exhibit high moisture damage indices and 9% of structures have frost insulation boards. Therefore, the interesting attributes in the hypothesis antecedents are the high moisture damage index and frost insulation boards. Coincidentally, frost insulation boards can disturb the GPR measurements causing increased moisture damage index values. Therefore, further data mining tasks should be conducted to investigate both the joint and separate correlations of these attributes to high TGDR values. These types of data mining tasks could be performed using the AC4ft-Miner, in which the effects of varying one attribute can be assessed. This one hypothesis cannot be used to draw conclusions about the types of structures correlated with high TGDRs. Instead, this hypothesis provided an interesting observation from the data, which can be used when designing subsequent data mining tasks. Further investigations can be found in Paper II.

Paper III reported a case study, in which the GUHA method was applied to railway asset management decision making. This study used the GUHA method to find interesting track structure types that correlated with high TGDRs. Sampling points were assigned to areas where GUHA had indicated problematic track structure types and also to reference areas. The aim was to verify whether GUHA could point out deficient track structure types. The sampling revealed moderate differences between track structure soil material quality and the track geometry deterioration behaviour; areas, where subballast material did not pass current regulations, exhibited higher deterioration rates. Additionally, the data mining implied problematic behaviour at locations where structural layer thicknesses were sufficient and yet frost insulation boards had been installed in the track structure. This peculiar observation stood out, and these sampling locations were found to contain fouled, moist ballast layers on top of the frost insulation boards. Thus, the study proved that with the available data and GUHA data mining, the sampling could be focused on these interesting areas and deteriorated structures were found. Subsequently, future research on the effects of frost insulation board installation depth and technique was initiated to investigate the mechanical reason for this behaviour. The data mining in this study was not used to make the final conclusions about the frost insulation board mechanics but to focus research on these interesting structure types. A similar approach is suggested for further research: data mining should be used to narrow the research focus to make further investigations more efficient. Therefore, the practical use case and benefit of explorative data mining in railway asset management is considered making further investigations more efficient.

### Synthesis

In the context of railway asset management, exploratory data mining is most beneficial in the early stages of designing track structure restorations. Data mining results can guide designers to investigate particular structure types and locations that have been found to correlate with poor track performance. These results can be used to assign site investigations more efficiently, focusing primarily on the problematic areas. Additionally, deficient structure types can be identified from data across multiple track sections, and these structure types can be redesigned in future restoration projects. Therefore, data mining should be performed in the very early stages of designing or even before initiating designing to understand the current structure behaviour better. Later on, once the general structure behaviour is realised, specific track locations can be investigated with data visualisation and correlation analyses, besides to more traditional methods like sampling and field testing.

Other researchers have also used exploratory data mining and data analytics in the railway domain (Liu *et al.*, 2012; Mirabadi and Sharifian, 2010; Sammouri *et al.*, 2013; Zarembski *et al.*, 2016). However, these studies have not investigated the root causes of track geometry defects, but rather the root causes of train equipment failures (Sammouri *et al.*, 2013), the impact of track geometry quality on rail defects (Zarembski *et al.*, 2016), and factors attributing to train accidents (Mirabadi and Sharifian, 2010) and derailments (Liu *et al.*, 2012). Nonetheless, all these studies have a common starting point; some interesting but harmful event has been recorded in data, and other relevant data is gathered to investigate features contributing to these events. Even the methods used to investigate correlations are rather similar: association rules and correlation analyses. Therefore, it appears that the GUHA method would be suitable for many other applications in the railway domain in addition to investigating the root causes of track geometry deterioration.

The methodological limitations of GUHA data mining were discussed in Paper II. These were the method's dependence on initial data and the amount of effort required for results analysis. First, if the initial data does not contain some important feature, data mining cannot consider that feature in the analyses. Therefore, the initial data must contain all relevant parameters depicting the track structure. This limits the use of exploratory data mining on track sections, from which little data is available. With that said, the possibility of using exploratory data mining to investigate root causes of track geometry deterioration should motivate the asset manager to collect enough of this type of data, as data mining has the potential to save money on expensive field investigations. Second, the practical limitation of using data mining in asset management is the complexity of the method and its results. A railway domain expert may not understand much about the data mining method or its initial results, but neither will a data scientist understand the meaning of the data mining results in the context of railway structures. However, it is possible to translate the hypotheses into human language and visualise their contents to ease results interpretation and communication between different domain experts. Therefore, data mining should be carried out in close cooperation with data analysts and railway asset managers making the best use of both domain experts.

## 3.3 Implementing track geometry deterioration analyses (TGDA) into asset management

Track geometry deterioration analyses (TGDA) should complement maintenance decision making processes by offering information regarding the development of the track structure condition. This information includes track geometry deterioration modelling results and data analyses, such as those described in sections 3.1 and 3.2. However, the implementation of TGDA into railway asset management is difficult because track geometry monitoring is a safety-critical process, which might be disturbed if development is rushed. In a worst-case scenario, the implementation of complex new systems might distract the personnel from noticing safety-critical track geometry faults, resulting in a train derailment. Therefore, a controlled process was studied for the implementation of TGDA into practical asset management (Paper IV). This included the application of maturity models as the basis for development, investigating the current maturity level of an asset management organisation, and creating a framework for advancing maturity in TGDA (Paper IV).

### 3.3.1 Maturity models as the basis for TGDA development

Maturity models were selected as the basis for TGDA development because they offer a step-by-step structure for controlled development. Controlled development is vital for steering the development of a safety-critical process and sustaining organisational motivation for development by offering intermediate goals. These are common goals for maturity model applications (Goncalves Filho and Waterson, 2018). The benefits of maturity models include spreading awareness of the different aspects of the analysed process, setting a frame of reference for systematic development, and ensuring quality throughout the process (Wendler, 2012).

The maturity model for this study was adapted from the UIC Railway Application Guide for implementing asset management through ISO 55001 (2016). The maturity model presented by the UIC (2016) is a general maturity model with six levels (or states), and it is already in use in railway asset management. As presented in Figure 23, the adapted maturity model consolidated the six maturity levels into four: 1) ensuring safety, 2) monitoring track quality, 3) track geometry management, and 4) optimising track geometry (Paper IV).

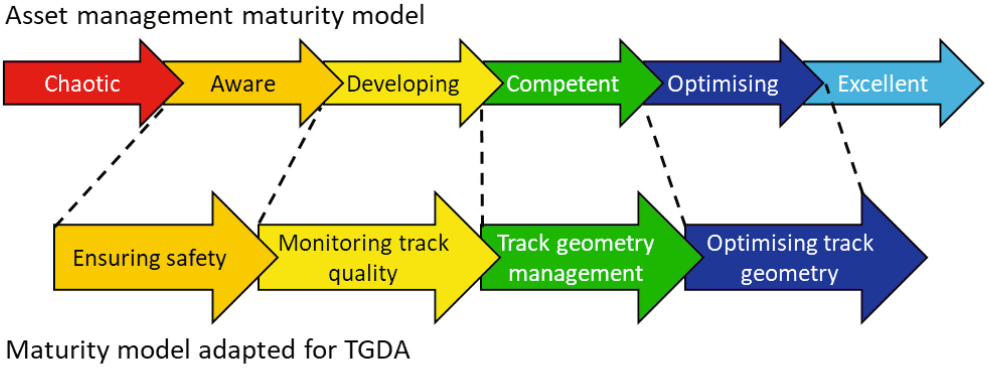


*Figure 23. General and adapted maturity models (Paper IV).*

The four levels (Figure 23) were adapted based on the different types of track geometry deterioration analysis types demonstrated in Paper IV. The initial level in the UIC guide (2016) maturity model was disregarded in the adapted model, as that level is equivalent to having no regular track inspections, which is not an option for a responsible railway owner. Therefore, the first level in the adapted model was ensuring safety, which was defined as having reactive processes, in which track faults are repaired after they are noticed in track inspections. On the second level, track geometry is monitored by forming time series data and analysing it manually. Recurrent faults and progressive deterioration can be noticed with subjective assessments, but the means of quantifying these are not yet available on the second level. The third level, track geometry management, introduces track geometry deterioration modelling, which provides the means for quantifying the deterioration behaviour and planning maintenance. Moreover, track geometry databases and modelling results are connected to other asset data, which allows for investigating the correlations between track assets and track faults. The fourth level, optimising track geometry, incorporates the optimisation and prioritisation of maintenance resources. The excellent level in the UIC guide (2016) was not regarded, because, in track geometry management, having fully optimised maintenance is the furthest goal currently.

The validity of the adapted maturity model was assessed in expert workshops where the maturity model was used as a basis for a TGDA development framework (Paper IV). Validity testing could include also, for example, quantitative surveys (Wendler, 2012). However, these were not deemed necessary, as this study aimed to only utilise the structure of the maturity model, the use of validation through expert assessment was considered sufficient (Paper IV).

Indeed, the adapted maturity model was not the primary outcome of this study but a platform, on which the framework was later built (Paper IV). The sole purpose of the framework is not to move up on the maturity levels but to give the asset management organisation a depiction of future capabilities that might not be otherwise recognised and motivate the organisation to advance development (Maier *et al.*, 2012). When an asset management organisation progresses its maturity level, the model should be revised to see if goals can be set even further than before. To conclude, the adapted maturity model here is not for academic purposes alone, but rather, the application of maturity models is seen as a way of establishing a systemic approach for planning TGDA development.

### 3.3.2 Investigating the current maturity level of TGDA

The current maturity level of TGDA in Finland was investigated using semi-structured interviews (Paper IV). The interviewees (Table 4) included 22 highly experienced professionals from track maintenance, track management, and the track owner. Most interviews were group interviews, often by the request of the interviewees to allow for colleagues to supplement each others' answers. All interviews were conducted online over Teams due to the COVID-19 pandemic. The interviews were recorded, and comprehensive memos were made based on the recordings. The participants were given the opportunity to comment on the memos after the interview. The structure of the interviews is reported in Paper IV.

*Table 4. Summary of interviewees.*

| | **Track maintenance** | **Track management** | **Track owner** |
|---:|:---:|:---:|:---:|
| Companies | 3 | 4 | 1 |
| Experts | 5 | 12 | 3 |
| Interviews | 3 | 4 | 3 |
| Average years of experience | 18 | 18 | 18 |

The complete, commented, and, in one case, revised memos were analysed using the ATLAS.ti 9 software. First, the interview memos were coded, and code groups were created. Altogether 1001 code references were created, which formed 67 unique codes in 9 code groups (Figure 24). The interview analyses were thoroughly researched for a comprehensive outlook on the practices related to track geometry measurements. The top three most common codes were *maintenance contracts* (n = 50), *reactions to measurements* (n = 49), and *maintenance planning* (n = 47), when the average number of references per code was 15. The most relevant findings included **the current main use cases of TGDA**, **future development needs**, **biggest obstacles**, and **different user types**. These are introduced in the following paragraphs.

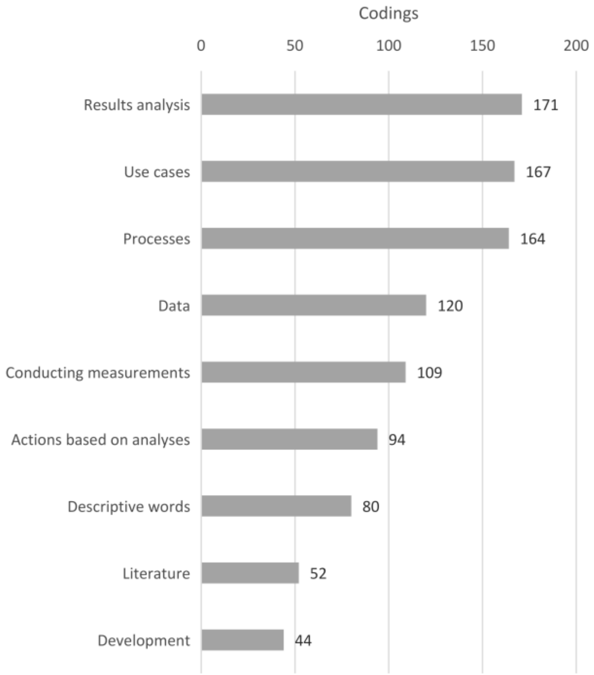


*Figure 24. Interview analysis code groups.*

According to the interviews, **the current main use cases of TGDA in asset management** are checking the current condition of track geometry, drafting tamping plans, and assessing track maintenance contract bonuses and sanctions. Further evaluation of the current condition and tamping planning is primarily done by conducting site inspections. The condition assessment and tamping planning is based on the experience and expertise of the personnel, and little to guide practical work is written in guidelines. Another frequently mentioned use case is track maintenance contract supervision. Track maintenance contracts include clauses determining acceptable track quality and response times for repairing observed track geometry faults. The track quality is tied to a bonus-sanction model, in which the contractor is awarded bonuses for achieving higher quality than required and sanctioned if the condition is poorer than required.

**The future development needs** discussed in the interviews can be divided into three categories: 1) automation, 2) system development, and 3) new in-depth information. Development needs in automation included automatic alerts, flexible reporting, tamping optimisation, and automatic tamping plans. Automatic alerts are required to inform the user about abrupt changes between the last two measurements, recurring faults, and measurement results exceeding threshold values. These alerts make results interpretation faster, as the user can prioritise focus on these interesting areas. Moreover, the system should automatically compare the latest measurement to previous measurements, which eases the human workload. Flexible reporting means that the user can freely select an area, which will be reported as a whole and for which graphs are drawn and key figures are calculated. Currently, the report content is fixed and is based on the measurement run length rather than user need. Tamping optimisation can be best described using a characterising example: If you have five kilometres of track that should be tamped and the resources to tamp only three kilometres, which three kilometres of the five should you tamp to achieve the highest overall quality with the available resources. Automatic tamping planning was often mentioned by maintenance personnel, and it refers to tamping plans being automatically created based on track recording car results. However, this development requires plenty of measuring system development as track geometry measurements are relative and tamping plans are drawn in absolute coordinates. Connecting these two measurement types will require major technical innovations. Nevertheless, this was considered the ultimate level of automation in TGDA by the maintenance personnel. System development needs mentioned in the interviews included mainly automatic data transfer between systems. Especially, track geometry faults should be transferred to maintenance databases automatically from measurements. This is currently done manually. Furthermore, asset data should be available when inspecting track geometry measurements, preferably, in geoinformation format so that assets and faults could be visualised on maps.

The new in-depth information refers to information that has not been available before in Finland. This included track geometry history, maintenance history, parameters for maintenance effectiveness, guidelines on what track components different geometry faults refer to and how to fix them, and predictions on the required maintenance and investments. Once in practical use, track geometry deterioration modelling could provide the required basic information for these needs, namely, the track geometry history, maintenance history, and maintenance effectiveness. Guidelines on which track components different geometry faults refer to require a vast amount of further research on the causality of track faults, although some faults are more easily identifiable than others. For example, track gauge widening on wooden sleeper tracks can be considered to result from poor sleeper and fastening condition. Predictions on future maintenance and investments were mentioned often in the interviews and are considered a major factor in improving maintenance effectiveness.

**The biggest obstacles** for future development were considered unavailable numeric track inspection data and relatively short five-year maintenance contracts, which deter private companies from developing their own systems. This was an interesting observation as these two seem to contradict one another. The professionals in the field want to have the data to themselves, but they want the infrastructure owner to develop the analysis systems. This can be interpreted to mean that there is a great interest to use and learn more about the measurements, but the companies do not believe in their capabilities or do not wish to use their own resources for development.

Three **different user types** were identified from the interviews: 1) track maintenance personnel, 2) asset manager, and 3) director. The first user type was track maintenance personnel. They require information on the types of faults found, the possible causes for the fault, and the type of remedy that should be appointed and when. These users are also interested in optimising maintenance to boost quality indices with the least possible maintenance efforts. The second user type was asset managers who oversee track maintenance. Their responsibility is to assess whether the decisions made by the maintenance personnel are valid and have meaningful effects. They also assess track maintenance contract bonus-sanction models, which are tied to track quality. The third user type, a director, is interested in general trends in track quality and maintenance effectiveness. These are generally area managers or infrastructure owner directors. What is noteworthy is that one person can exhibit traits from all types of users, meaning they require multiple different data analyses. However, some users require very little information, as some are satisfied with only checking overall track quality indices. Furthermore, some interviewees did not see the value of TGDA or any data analytics and instead considered site inspections more valuable. These observations suggest that the framework application must be flexible to users with different levels of demand and detail regarding TGDA. Moreover, there may be some change resistance among professionals who are content with the current processes.

In conclusion, the findings from the interviews suggested that the maturity of TGDA is on the *monitoring track quality* level in Finland. Some more advanced analyses had been attempted by individuals, but these were not a part of official processes. The information from the interviews was also utilised in creating the framework for TGDA development.

Similar studies using semi-structured interviews have been performed in the Netherlands, Sweden, and United Kingdom, where the infrastructure asset management challenges and improvements have been investigated (Al-Douri *et al.*, 2016; Dadashi *et al.*, 2014; Schraven *et al.*, 2011). In the Netherlands, 12 semi-structured interviews of public agency employees were conducted to increase the understanding of the challenges in infrastructure decision making (Schraven *et al.*, 2011). The study found challenges in, for instance, setting strategic objectives and measuring performance with subjective quality assessments (Schraven *et al.*, 2011). The Swedish study interviewed eight experts to investigate whether the Swedish asset management has access to sufficient track information for efficient maintenance decision making (Al-Douri *et al.*, 2016). The study found several technical aspects of track inspection and data analyses that needed to be improved (Al-Douri *et al.*, 2016). Moreover, a lack of a long-term maintenance strategy was seen to hinder maintenance planning (Al-Douri *et al.*, 2016). In the United Kingdom, 20 semi-structured expert interviews were used to investigate the status of intelligent infrastructure in railway management (Dadashi *et al.*, 2014). The purpose was to create a framework based on the observations. The study from the United Kingdom revealed three user types for intelligent infrastructure systems, two of which were very similar to the ones found in Finnish research: track workers/maintenance personnel and strategic analysts/directors. Contrarily to the findings from the Netherlands and Sweden, the results from the United Kingdom suggested little concern for the technological aspects of the development, as some interviewees were confident with the available technological advances. However, the implementation of intelligent infrastructure systems and their use in practice was considered to be a greater challenge.

The development needs mentioned in past research from the Netherlands (Schraven *et al.*, 2011), Sweden (Al-Douri *et al.*, 2016), and the United Kingdom (Dadashi *et al.*, 2014) concur with the observations made in Finland. Namely, the lack of data analysis systems for analysing the massive amount of asset data is a common issue. Moreover, the need to serve different user types and heuristic analyses is notable. These observations motivated research further into creating a framework for advancing TGDA.

### 3.3.3 Creating a framework for TGDA development

A framework for developing TGDA in Finnish railway asset management was created based on the previously reported adapted maturity model (Paper IV). The framework was created during three workshops with experts from track maintenance companies, consultant companies, and the infrastructure owner. The design of the workshops is presented in Figure 25.

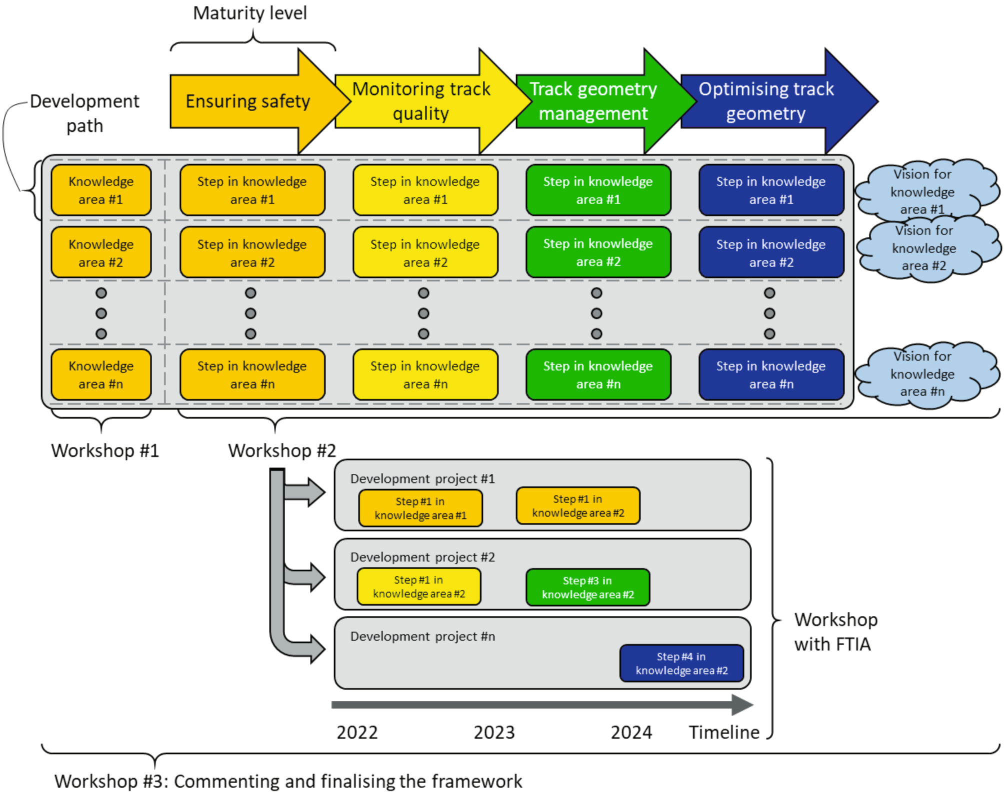


*Figure 25. Design of the workshops (Paper IV).*

The terminology in Figure 25 is defined as follows; maturity levels provide the structure for different development stages. Knowledge areas are the different topics that the framework concerns. Development paths contain the steps that must be taken to advance the knowledge area maturity toward its vision (Paper IV).

The framework was accumulated over three workshops with all stakeholders and an intermediate workshop with only the infrastructure owner, FTIA (Paper IV). The first workshop concentrated on familiarising the participants with the work ahead and forming the knowledge areas, which structure the framework. There were 23 participants who were divided into groups of three to four people, mixing people with different backgrounds into the same groups. Each group was asked to draft mind maps of the processes and topics that are affected by track inspections. After the workshop, the mind maps were analysed with ATLAS.ti 9 software. The analysis included coding the mind maps and creating code groups. The code groups were modified into the knowledge areas for the framework. Based on the findings from the first workshop, six knowledge areas were created: 1) **measurement analysis**, 2) **data systems**, 3) **maintenance**, 4) **asset renewal**, 5) **knowledge**, and 6) **contracts**. **Measurement analysis** refers to the processes for receiving, transferring, modelling, and visualising the track geometry measurement results. **Data systems** include the software required for measurement analysis and data storage. **Maintenance** refers to planning, executing, and supervising the track work defined in the current maintenance contracts, for example, tamping and spot maintenance. **Asset renewal**, conversely, refers to track work not included in the current maintenance contract, for example, superstructure renewals. The term **knowledge** includes the skills needed for utilising TGDA, as well as training and guidelines. **Contracts** cover the required contracts for all parts of the process, for instance, maintenance, asset management, and data systems.

The second workshop focused on creating the development paths for the established knowledge areas. The 17 participants were divided into six groups and each group was assigned one knowledge area and a preliminary vision for said area. Then, the participants edited the preliminary vision and filled a blank four-level maturity model with the required steps for achieving the revised vision. Once these were completed, the groups were rotated twice to enable commenting and supplementing of other groups' work. At the end of the second workshop, six preliminary development paths had been created and commented on.

Before the third and final workshop, a draft of the complete framework was presented to the infrastructure owner, FTIA, in an intermediate workshop. In this workshop, the framework draft was compared with the ongoing development projects to detect overlapping work. Fortunately, no projects overlapping the framework were identified, and the framework development could continue.

The finalised framework was presented in the third and final workshop. The 16 participants were asked to evaluate, comment, and supplement the framework. As well as giving the stakeholders (participants) a chance to comment on the framework, the final workshop was a way to present and engage the stakeholders on future development.

The preliminary output of the workshops was a conceptual version of the framework. This conceptual framework included step-by-step development paths for each of the knowledge areas. The steps were individual technologies or actions reported in the workshops, such as a track inspection database or maintenance effectiveness assessments. These steps were connected with arrows indicating the order of progress. The connected steps formed chains of events, which demonstrate the actions that must precede one another to keep progress logical. As an example of one chain of events, Figure 26 presents all the preceding steps for automatic maintenance plan drafting.

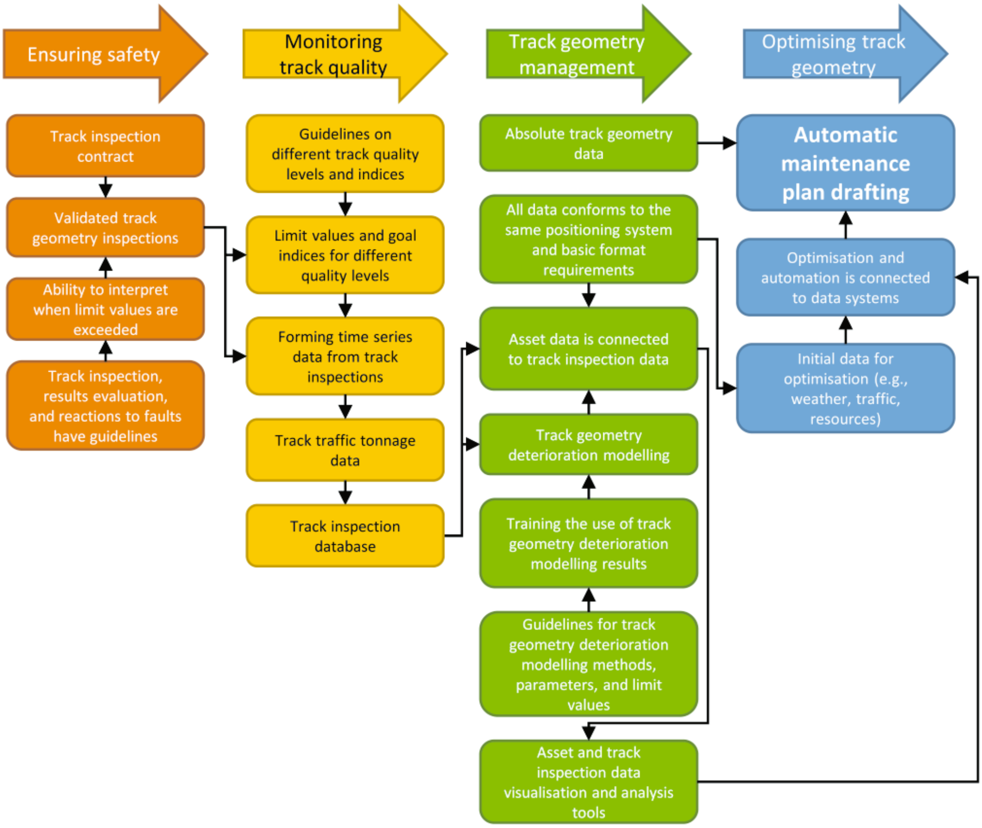
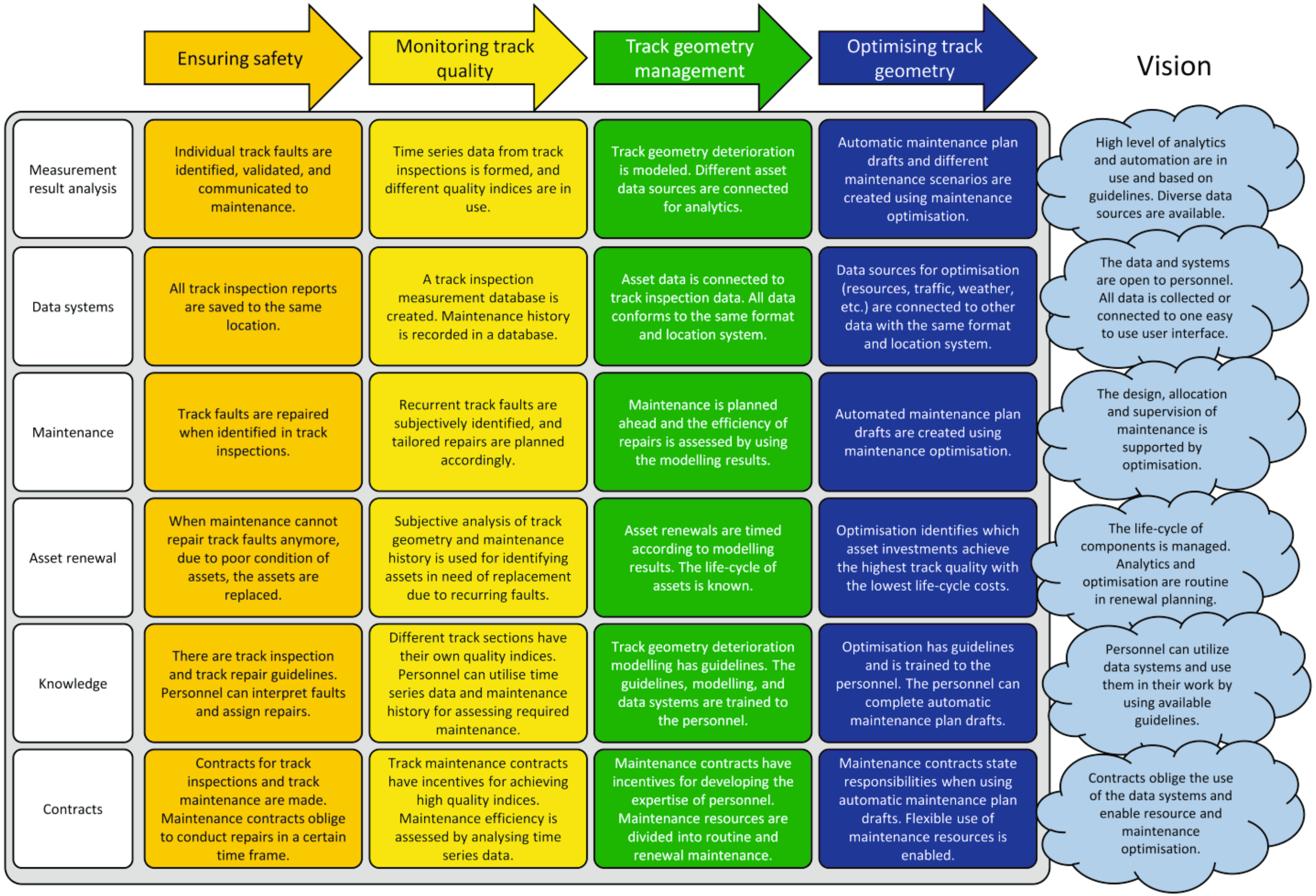


*Figure 26. Chain of events for automatic maintenance plan drafting.*

The arrows in Figure 26 indicate prerequisites for a step in the framework. For example, a prerequisite for track geometry deterioration modelling is a track inspection database, for which a prerequisite is traffic tonnage and so forth. The colour of the step denotes the maturity level in question. The division into knowledge areas has been omitted from Figure 26 to compact the image. These types of results can be used in practice by setting one step as a goal and making sure all necessary preceding steps are completed before attempting to implement the final step. To summarise all the different chains of events, all these individual steps were consolidated into a framework for advancing TGDA presented in Figure 27.



| | Ensuring safety | Monitoring track quality | Track geometry management | Optimising track geometry | Vision |
|---|---|---|---|---|---|
| Measurement result analysis | Individual track faults are identified, validated, and communicated to maintenance. | Time series data from track inspections is formed, and different quality indices are in use. | Track geometry deterioration is modeled. Different asset data sources are connected for analytics. | Automatic maintenance plan drafts and different maintenance scenarios are created using maintenance optimisation. | High level of analytics and automation are in use and based on guidelines. Diverse data sources are available. |
| Data systems | All track inspection reports are saved to the same location. | A track inspection measurement database is created. Maintenance history is recorded in a database. | Asset data is connected to track inspection data. All data conforms to the same format and location system. | Data sources for optimisation (resources, traffic, weather, etc.) are connected to other data with the same format and location system. | The data and systems are open to personnel. All data is collected or connected to one easy to use user interface. |
| Maintenance | Track faults are repaired when identified in track inspections. | Recurrent track faults are subjectively identified, and tailored repairs are planned accordingly. | Maintenance is planned ahead and the efficiency of repairs is assessed by using the modelling results. | Automated maintenance plan drafts are created using maintenance optimisation. | The design, allocation and supervision of maintenance is supported by optimisation. |
| Asset renewal | When maintenance cannot repair track faults anymore, due to poor condition of assets, the assets are replaced. | Subjective analysis of track geometry and maintenance history is used for identifying assets in need of replacement due to recurring faults. | Asset renewals are timed according to modelling results. The life-cycle of assets is known. | Optimisation identifies which asset investments achieve the highest track quality with the lowest life-cycle costs. | The life-cycle of components is managed. Analytics and optimisation are routine in renewal planning. |
| Knowledge | There are track inspection and track repair guidelines. Personnel can interpret faults and assign repairs. | Different track sections have their own quality indices. Personnel can utilise time series data and maintenance history for assessing required maintenance. | Track geometry deterioration modelling has guidelines. The guidelines, modelling, and data systems are trained to the personnel. | Optimisation has guidelines and is trained to the personnel. The personnel can complete automatic maintenance plan drafts. | Personnel can utilize data systems and use them in their work by using available guidelines. |
| Contracts | Contracts for track inspections and track maintenance are made. Maintenance contracts oblige to conduct repairs in a certain time frame. | Track maintenance contracts have incentives for achieving high quality indices. Maintenance efficiency is assessed by analysing time series data. | Maintenance contracts have incentives for developing the expertise of personnel. Maintenance resources are divided into routine and renewal maintenance. | Maintenance contracts state responsibilities when using automatic maintenance plan drafts. Flexible use of maintenance resources is enabled. | Contracts oblige the use of the data systems and enable resource and maintenance optimisation. |

*Figure 27. Framework for advancing TGDA (Paper IV).*

The framework for advancing TGDA (Figure 27) describes the end state of each maturity level and knowledge area (Paper IV). This limits the amount of detail presented in the framework as all individual technological advancements are not mentioned. However, this goal-oriented framework was created to serve as a tool for strategic planning of future development projects. In strategic planning, individual technological advancements are not in focus. Instead, presenting the big picture of where development is heading and what the goals are is much more valuable.

### Synthesis

The tangible benefit of the framework for TGDA development is the possible cost savings achieved with increased efficiency in track maintenance. The increased efficiency is achieved when more informed decisions can be made based on more robust data, which is produced by using advanced TGDA. The framework includes all relevant aspects for advancing TGDA in its six knowledge areas. The knowledge areas have distinct goals for each maturity level, dividing the progress into controllable segments. This enables creating a far-reaching vision for the future all the while overseeing current practical development.

The framework was evaluated based on analysing the outputs from the workshops, in this case, the mind maps, development paths, and comments. Other methods for evaluating workshop outputs could include, for example, recording and analysing the workshop conversations or surveying the participants after the workshop (Thoring *et al.*, 2020). The workshop results could be evaluated with these methods, or a combination of methods, referred to as triangulation (Thoring *et al.*, 2020). However, the workshop results in this study were evaluated only with output analysis, which means that for example, participant dialogues and interactions were not analysed. This is a limitation of the study but one that is justified by the aim of the workshops: to produce a framework. If the aim was altered toward the implementation of the framework or stakeholders' willingness to incorporate TGDA within their organisations, then other evaluation methods could be required.

The limitations of the framework are associated with the limitations of the workshops. In addition to the previously mentioned limitations in workshop evaluation, the limited number and homogenous background of the participants limited the input and feedback to the framework. All participants work in the Finnish railway operating environment; therefore, the applicability of the framework in different operating environments could not be validated.

Previous research has also produced frameworks, or roadmaps in other terminology, for railway technology. Dadashi *et al.* (2014) reported a data processing framework for intelligent railway infrastructure. This framework was based on four levels of understanding (*cf.* maturity levels). The outlook of their framework was similar to the conceptual framework created in the workshops. However, the framework by Dadashi *et al.* (2014) regarded a more general system architecture (intelligent infrastructure), whereas the present study focuses solely on TGDA. These two frameworks can be seen as complementing each other, one on a more general level (intelligent infrastructure) and the other on a specific level (TGDA). Other roadmaps in railway technology include the overall development of railway technology within a railway system (Blumenfeld *et al.*, 2019; Dias *et al.*, 2015). The scopes of these frameworks are much wider than that of the current study. Yet, the same ideology can be seen in these; intermediate steps are described and set as goals to segment large-scale development into a more controllable process.

Future research regarding the framework for advancing TGDA should include studying its application to different operating environments and its further implementation. The framework was created in the Finnish operating environment, and although it is constructed in a general form, its applicability to different operating environments is not yet validated. In studying the further implementation of the framework, the stakeholder attitudes and willingness should be researched to reveal possible obstacles in its practical application.

# 4 CONCLUSIONS

This research had three main purposes: 1) to model track geometry deterioration based on Finnish track inspection history (Paper I), 2) to investigate the root causes of track geometry deterioration using exploratory data mining (Papers II and III), and 3) to implement TGDA into practical asset management (Paper IV). This section concludes the main results, contemplates their significance and validity, and provides suggestions for future research.

## 4.1 Research outcomes

The research impact is discussed on two levels: scientific and practical. The scientific implications of this study include:

- the testing and development of a robust linear optimisation-based method for modelling track geometry deterioration using Finnish track inspection data (Paper I). The linear models were found suitable for the Finnish track environment, and a typical TGDR range was found to be 0–0.5 mm/MGT for the 200 m LL SD. However, diverse TGDRs were found on different track sections and different parts of the same track section. The developed models and results' visualisation techniques enable track geometry deterioration predictions, past deterioration analyses, and maintenance effectiveness quantification. The models are suitable for all track geometry parameters that follow a linear deterioration path. The results also added to the growing amount of research on track geometry deterioration modelling output values.
- the novel application of a data mining method, GUHA, for exploratory railway asset data analysis (Papers II and III). These results and corresponding field tests demonstrated that track structure asset data can indicate defective track structure types when the data is analysed with appropriate means. For example, a correlation, stronger than four times above average, between certain defective structures with frost insulation boards was obtained.
- a novel application of maturity models as the basis for a development framework (Paper IV). These results demonstrated that maturity models could be applied as the basis for a development framework for TGDA implementation. The framework presents four maturity levels that define the maturity of TGDA with respect to different aspects of asset management.

The practical impacts of this research impact different organisations. The primary beneficiary of this research is the infrastructure owner, which is the FTIA in the Finnish operating environment. Using the results, the infrastructure owner can predict and evaluate the condition of their track assets (RO #1), investigate the root causes for deterioration (RO #2), and the aforementioned can be implemented into practical asset management processes (RO #3). All these results aim at the same outcome: higher maintenance effectiveness. Higher maintenance effectiveness is achieved by allocating maintenance recourses more accurately, thus reducing unnecessary work. This is possible when asset management has access to data on track geometry deterioration behaviour and its causes. The infrastructure owner has the most to gain from this because besides saving on the maintenance costs, asset investment decisions will be based on more sound data, which enables optimising the life cycle of components.

Track maintenance can also benefit from the results, as they can plan maintenance more accurately with better TGDA. Areas requiring maintenance can be identified more effectively and the timing of future maintenance can be estimated. These help track maintenance in planning their resources to eliminate unnecessary work, such as back-and-forth machinery relocation or repetitive ineffective maintenance. This benefits track maintenance, as they can achieve the required track quality with fewer resources than before.

Finally, improving the efficiency of track maintenance impacts society as a whole. Railway maintenance requires considerable tax revenue in Finland. Therefore, minimising spending while maximising track availability is one of the most important tasks in railway ownership in Finland. The results of this study provide the means to make well-informed decisions on track maintenance, thus saving taxpayer money and track downtime. Additionally, the end-customers, who are either passengers or cargo transporters, have better access to railway traffic. The service and reliability of railways increase when the track downtime required for maintenance can be decreased with better maintenance planning and disruptions become rarer as there are fewer unexpected track faults. These make rail traffic a more attractive means of transport as a whole, which is important not only for industrial competitiveness but also for promoting environmentally sustainable transport.

## 4.2 Validity, reliability, and limitations

The validity of robust linear track geometry modelling in predicting and depicting the development in track geometry quality is well established. Many studies from different environments have found the linear modelling suitable for describing track geometry deterioration behaviour (Andrade and Teixeira, 2015; Audley and Andrews, 2013; Caetano and Teixeira, 2016; Khajehei *et al.*, 2019; Lee *et al.*, 2020; Li *et al.*, 2019; Neuhold *et al.*, 2020). However, considering the validity of using track (geometry) inspections as a way of depicting the condition of the railway track is more complex. As discussed in section 2.3, deterioration in many different parts of a track structure is shown in track geometry. However, not all deterioration is shown in track geometry, for example, internal cracking in rails, until the deterioration progresses to a dangerous level. Additionally, different faults may be observable through different track geometry indices, for example, settlements are shown in the LL, but cracked concrete sleepers may be shown in the track gauge. Therefore, track geometry deterioration modelling should not be the only way of monitoring the overall condition of railway tracks in practice. For example, visual inspections and ultrasonic rail inspections are also required. As for reliability, the robust linear track geometry deterioration modelling was tested on three different track sections. These track sections are considered rather high-class railroads with moderate to high traffic volumes. Consequently, the modelling of low-class track sections with low traffic volumes was not tested, which is a limitation of this part of the research. Other limitations include the lack of testing regarding other track geometry parameters than the LL and unsystematic maintenance records. Both these limitations are caused by the lack of available historical data, but recently implemented measurement and data storage systems will ensure that future research will be able to address these.

Using GUHA data mining in investigating the root causes of track geometry deterioration can be considered valid when the implication of the results is deliberately considered. The GUHA method can reveal interesting correlations within large data sets. The correlations are based on logic; hence there are no prediction or estimation errors involved with the method, as the correlations are based solely on the initial data. However, therein lies the major limitation of the method: dependency on initial data. If some important feature is missing from the initial data, it cannot be considered in the results. Therefore, the initial data should contain all relevant parameters depicting the researched behaviour. Furthermore, the correlations do not imply causality, as was discussed in section 3.2.2. However, if these limitations are acknowledged, the method can produce remarkable results; the

GUHA method could reveal interesting correlations between track structure features and the TGDR to investigate further with other methods.

Semi-structured interviews and workshops were used as the research methods for creating the framework for TGDA development. Semi-structured interviews have been used in past railway research and can be considered a valid research method for investigating a particular process, especially when the number of experts working on the matter is low (Al-Douri *et al.*, 2016; Dadashi *et al.*, 2014; Schraven *et al.*, 2011). Workshops as a research method are not as prevalent as interviews, even less so in railway research. However, workshops as a research method have been investigated and found beneficial in research relating to strategic planning (Phaal *et al.*, 2007; Thoring *et al.*, 2020). The reliability of the interviews and workshops is based on the representativity of the participants. All participants were from Finland and all relevant organisations participated in the research. Thus, the results can be considered reliable in the Finnish operating environment. However, the group interviews and interorganisational workshops may have prevented the participants from revealing all weaknesses in current processes, for the fear of being judged by their colleagues. Therefore, there may be some details in the framework that have not been discovered yet and future implementation should prepare to adapt the framework should these arise. Furthermore, the reliability of the results in other than the Finnish environment was not tested, and, therefore, should be a topic of future research.

## 4.3 Suggestions for future research

Suggestions for future research been have provided throughout the text in previous sections. These suggestions are composed in this section.

In future research on **track geometry deterioration modelling**, the verification of the tamping identification algorithm, calculations on the effects of different maintenance actions, and international definition and comparison of different TGDRs should be top priorities. The tamping algorithm can be validated as soon as reliable maintenance data has been gathered over a few years. Then, the tamping areas identified by the algorithm can be compared with the actual maintenance data. Following this, the effectiveness of different types of maintenance actions can be calculated, as the maintenance data will also indicate the type of performed maintenance. Relating to this, the MSI limit values should be investigated using data from multiple different track environments. Maintenance effectiveness calculations can enable quality control for past maintenance as well as making predictions on future maintenance effectiveness. Lastly, the TGDR should be defined and its range in different operating environments should be investigated internationally. International comparison of TGDRs from different environments would form the basis for standardising and creating limit values for track geometry deterioration, which would greatly benefit asset management decision making when defining whether a track section is in good condition or not.

As for **investigating the root causes of track geometry deterioration**, future research should focus on investigating the causality of the correlations found using exploratory data mining. First, the methods and available data for investigating the causality in the obtained correlations between track structure features and track geometry deterioration should be researched. This includes forming time series data on track structure features and comparing it with track geometry deterioration modelling results. This time series data should include the timing and type of past maintenance and the timing and contents of past renewals. These would enable investigating the effects of different maintenance actions and renewals on various types of track structures. This would provide more insight into what types of maintenance actions and renewals have had meaningful effects on particular track structure types.

Future research on the **framework for advancing TGDA** should include further implementation of the framework and testing the framework in different railway asset management organisations. Further implementation research on the TGDA development framework should include research on change management. The framework provides a strategy for future development, but the practical aspects of changing the way organisations utilise TGDA still require applied research. This is an exceptionally important step, as several studies have indicated that most organisations fail in implementing changes (Errida and Lotfi, 2021). Lastly, the framework should be tested in different asset management organisations to assess the contents of the framework and identify possible missing steps. If the framework is found suitable in the other railway asset management organisations, the framework should be implemented into those as well and for good reason. As this dissertation has demonstrated in many ways, railway asset management organisations can obtain tangible benefits when advancing their abilities in analysing track geometry deterioration.

# 5 REFERENCES

Agresti, A. (2015), *Foundations of Linear and Generalized Linear Models*, John Wiley & Sons, Incorporated, New York, USA.

Akagawa, S., Hori, M. and Sugawara, J. (2017), "Frost Heaving in Ballast Railway Tracks", *Procedia Engineering*, The Author(s), Vol. 189, pp. 547–553.

Al-Douri, Y.K., Tretten, P. and Karim, R. (2016), "Improvement of railway performance: a study of Swedish railway infrastructure", *Journal of Modern Transportation*, Springer Berlin Heidelberg, Vol. 24 No. 1, pp. 22–37.

Albliwi, S.A., Antony, J. and Arshed, N. (2014), "Critical literature review on maturity models for business process excellence", 2014 *IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 79–83.

Andersland, O.B. (2004), *Frozen Ground Engineering, Second Edition*, edited by Ladanyi, B., 2nd ed., book, John Wiley & Sons, Hoboken, N.J.

Andersson, C. and Dahlberg, T. (2000), "Load impacts at railway turnout crossing", *Vehicle System Dynamics*, Vol. 33, pp. 131–142.

Andrade, A.R. and Teixeira, P.F. (2010), "Uncertainty in Rail-Track Geometry Degradation: Lisbon-Oporto Line Case Study", *Journal of Transportation Engineering*, Vol. 137 No. 3, pp. 193–200.

Andrade, A.R. and Teixeira, P.F. (2012), "A Bayesian model to assess rail track geometry degradation through its life-cycle", *Research in Transportation Economics*, Elsevier, Vol. 36 No. 1, pp. 1–8.

Andrade, A.R. and Teixeira, P.F. (2014), "Unplanned-maintenance needs related to rail track geometry", *Proceedings of the Institution of Civil Engineers - Transport*, ICE Publishing, Vol. 167 No. 6, pp. 400–410.

Andrade, A.R. and Teixeira, P.F. (2015), "Statistical modelling of railway track geometry degradation using Hierarchical Bayesian models", *Reliability Engineering and System Safety*, Elsevier, Vol. 142, pp. 169–183.

Andrews, J., Prescott, D. and De Rozières, F. (2014), "A stochastic model for railway track asset management", *Reliability Engineering and System Safety*, Elsevier, Vol. 130, pp. 76–84.

Audley, M. and Andrews, J.D. (2013), "The effects of tamping on railway track geometry degradation", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 227 No. 4, pp. 376–391.

Bai, L., Liu, R., Sun, Q., Wang, F. and Wang, F. (2016), "Classification-learning-based framework for predicting railway track irregularities", *Proceedings of the Institution of*

*Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 230 No. 2, pp. 598–610.

Bai, L., Liu, R., Sun, Q., Wang, F. and Xu, P. (2015), "Markov-based model for the prediction of railway track irregularities", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 229 No. 2, pp. 150–159.

Barke, D.W. and Chiu, W.K. (2005), "A review of the effects of out-of-round wheels on track and vehicle components", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 219 No. 3, pp. 151–175.

Berggren, E.G. (2010), "Efficient Track Maintenance: Methodology for Combined Analysis of Condition Data", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 224 No. 5, pp. 353–360.

Berka, P. (2016), "Practical aspects of data mining using LISp-Miner", *Computing and Informatics*, Vol. 35 No. 3, pp. 528–554.

Biabani, M.M. and Indraratna, B. (2015), "An evaluation of the interface behaviour of rail subballast stabilised with geogrids and geomembranes", *Geotextiles and Geomembranes*, Elsevier Ltd, Vol. 43 No. 3, pp. 240–249.

Blumenfeld, M., Wemakor, W., Azzouz, L. and Roberts, C. (2019), "Developing a new technical strategy for rail infrastructure in low-income Countries in Sub-Saharan Africa and South Asia", *Sustainability (Switzerland)*, Vol. 11 No. 16, pp. 1–23.

de Bruin, T., Rosemann, M., Freeze, R. and Kulkarni, U. (2005), "Understanding the main phases of developing a maturity assessment model", *ACIS 2005 Proceedings - 16th Australasian Conference on Information Systems*.

Caetano, L.F. and Teixeira, P.F. (2016), "Predictive maintenance model for ballast tamping", *Journal of Transportation Engineering*, Vol. 142 No. 4, pp. 1–9.

Cambridge Dictionary. (2022a), "data", *Cambridge Dictionary*, retrieved 10.9.2022, available at: https://dictionary.cambridge.org/.

Cambridge Dictionary. (2022b), "framework", *Cambridge Dictionary*, retrieved 10.9.2022, available at: https://dictionary.cambridge.org/.

Chen, Q., Niu, X., Zuo, L., Zhang, T., Xiao, F., Liu, Y. and Liu, J. (2018), "A railway track geometry measuring trolley system based on aided INS", *Sensors (Switzerland)*, Vol. 18 No. 2, available at:https://doi.org/10.3390/s18020538.

CMMI Product Team. (2010), "CMMI for Development, Version 1.3", Carnegie Mellon University.

D'Angelo, G., Thom, N. and Lo Presti, D. (2016), "Bitumen stabilized ballast: A potential solution for railway track-bed", *Construction and Building Materials*, Vol. 124, pp. 118–126.

Dadashi, N., Wilson, J.R., Golightly, D. and Sharples, S. (2014), "A framework to support human factors of automation in railway intelligent infrastructure", *Ergonomics*, Taylor & Francis, Vol. 57 No. 3, pp. 387–402.

Dahlberg, T. (2001), "Some railroad settlement models - A critical review", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 215 No. 4, pp. 289–300.

Dahlberg, T. (2010), "Railway Track Stiffness Variations – Consequences and Countermeasures", *International Journal of Civil Engineerng*, Vol. 8 No. 1, pp. 1–12.

Dias, A., Da Silva, J.C., Kagueiama, H.H., Haddad, C.R., Uriona M, M., Lezana, Á.G.R. and Juliatto, D.L. (2015), "Railway industry roadmap in the State of Santa Catarina, Brazil", *Transport Problems*, Vol. 10, pp. 29–36.

Elkhoury, N., Hitihamillage, L., Moridpour, S. and Robert, D. (2018), "Degradation Prediction of Rail Tracks: A Review of the Existing Literatur ", *The Open Transportation Journal*, Vol. 12, pp. 88–104.

Errida, A. and Lotfi, B. (2021), "The determinants of organizational change management success: Literature review and case study", *International Journal of Engineering Business Management*, Vol. 13, pp. 1–15.

Esveld, C. (2001), *Modern Railway Track*, 2nd ed., MRT-Prod, Zaltbommel.

Esveld, C., Jourdain, A., Kaess, G. and Shenton, M.J. (1988), "Historic Data on Track Geometry in Relation To Maintenance", *Rail Engineering International*, Vol. 17 No. 2, pp. 16–19.

European Commission (2021), *Seventh Monitoring Report on the Development of the Rail Market under Article 15(4) of Directive 2012/34/EU of the European Parliament and of the Council*, European Commission, Brussels, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2021:0005:FIN.

Famurewa, S.M., Juntti, U., Nissen, A. and Kumar, U. (2016), "Augmented utilisation of possession time: Analysis for track geometry maintenance", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 230 No. 4, pp. 1118–1130.

Fellinger, M. (2020), *Sustainable Asset Management for Turnouts*, Doctoral thesis, TU Graz.

Ferdous, W. and Manalo, A. (2014), "Failures of mainline railway sleepers and suggested remedies - Review of current practice", *Engineering Failure Analysis*, Vol. 44, pp. 17–35.

Ferreira, L. and Murray, M.H. (1997), "Modelling rail track deterioration and maintenance: Current practices and future needs", *Transport Reviews*, Vol. 17 No. 3, pp. 207–221.

Ferreira, T.M. and Teixeira, P.F. (2012), "Rail track performance with different subballast solutions: Traffic and environmental effects on subgrade service life", *Journal of Transportation Engineering*, Vol. 138 No. 12, pp. 1541–1550.

Fonseca, J.A. and de Almeida Júnior, J.R. (2005), "CMMI RAMS Extension Based on CENELEC Railway Standard BT - Computer Safety, Reliability, and Security", in Winther, R., Gran, B.A. and Dahll, G. (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–12.

Goddard, W. and Melville, S. (2001), "Research methodology: An Introduction", Juta, Lansdowne.

Goncalves Filho, A.P. and Waterson, P. (2018), "Maturity models and safety culture: A critical review", *Safety Science*, Elsevier, Vol. 105, pp. 192–211.

Gou, H., Ran, Z., Yang, L., Bao, Y. and Pu, Q. (2019), "Mapping vertical bridge deformations to track geometry for high-speed railway", *Steel and Composite Structures*, Vol. 32 No. 4, pp. 467–478.

Grassie, S.L. (2012), "Squats and squat-type defects in rails: The understanding to date", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 226 No. 3, pp. 235–242.

Grossoni, I., Andrade, A.R., Bezin, Y. and Neves, S. (2019), "The role of track stiffness and its spatial variability on long-term track quality deterioration", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 233 No. 1, pp. 16–32.

Guler, H. (2014), "Prediction of railway track geometry deterioration using artificial neural networks: A case study for Turkish state railways", *Structure and Infrastructure Engineering*, Taylor & Francis.

Guler, H., Jovanovic, S. and Evren, G. (2011), "Modelling railway track geometry deterioration", *Proceedings of the Institution of Civil Engineers - Transport*, Vol. 164 No. 2, pp. 65–75.

Hájek, P. and Havránek, T. (1978), *Mechanizing Hypothesis Formation: Mathematical Foundations for a General Theory*, Springer Berlin Heidelberg.

Hájek, P., Holeňa, M. and Rauch, J. (2010), "The GUHA method and its meaning for data mining", *Journal of Computer and System Sciences*, Vol. 76 No. 1, pp. 34–48.

Helgesson, Y.Y.L., Höst, M. and Weyns, K. (2012), "A review of methods for evaluation of maturity models for process improvement", *Journal of Software: Evolution and Process*, John Wiley & Sons, Ltd, Vol. 24 No. 4, pp. 436–454.

Herbsleb, J.D. and Goldenson, D.R. (1995), "Systematic survey of CMM experience and results", *Proceedings - International Conference on Software Engineering*, pp. 323–330.

Higgins, C. and Liu, X. (2018), "Modeling of track geometry degradation and decisions on safety and maintenance: A literature review and possible future research directions", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 5, pp. 1385–1397.

Hirose, Y., Phaal, R., Farrukh, C., Gerdsri, N. and Lee, S. (2020), "Characterizing Maturity Levels for Organization-Wide Roadmapping Implementation", *IEEE Engineering Management Review*, Vol. 48 No. 4, pp. 133–143.

Humphrey, W.S. (1988), "Characterizing the Software Process: A Maturity Framework", *IEEE Software*, Vol. 5 No. 2, pp. 73–79.

Hyslip, J.P. (2002), "Fractal analysis of track geometry data", *Transportation Research Record*, No. 1785, pp. 50–57.

Indraratna, B., Babar Sajjad, M., Ngo, T., Gomes Correia, A. and Kelly, R. (2019), "Improved performance of ballasted tracks at transition zones: A review of experimental and modelling approaches", *Transportation Geotechnics*, Elsevier, Vol. 21, p. 100260.

Indraratna, B., Tennakoon, N., Nimbalkar, S. and Rujikiatkamjorn, C. (2013), "Behaviour of clay-fouled ballast under drained triaxial testing", *Geotechnique*, Vol. 63 No. 5, pp. 410–419.

ISO 33004. (2020), "ISO 33004 Information technology. Process assessment. Requirements for process reference, process assessment and maturity models".

Jiang, J.J., Klein, G., Hwang, H.G., Huang, J. and Hung, S.Y. (2004), "An exploration of the relationship between software development process maturity and project performance", *Information and Management*, Vol. 41 No. 3, pp. 279–288.

Jing, G., Qie, L., Markine, V. and Jia, W. (2019), "Polyurethane reinforced ballasted track: Review, innovation and challenge", *Construction and Building Materials*, Vol. 208, pp. 734–748.

Johansson, A. and Nielsen, J.C.O. (2003), "Out-of-round railway wheels—wheel-rail contact forces and track response derived from field tests and numerical simulations", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 217 No. 2, pp. 135–146.

Jovanovic, S. (2004), "Railway track quality assessment and related decision making", Conference Proceedings - *IEEE International Conference on Systems, Man and Cybernetics*, Vol. 6, pp. 5038–5043.

Kalliainen, A., Kolisoja, P. and Nurmikolu, A. (2016), "3D Finite Element Model as a Tool for Analyzing the Structural Behavior of a Railway Track", *Procedia Engineering*, Elsevier B.V., Vol. 143, pp. 820–827.

Kennedy, J., Woodward, P.K., Medero, G. and Banimahd, M. (2013), "Reducing railway track settlement using three-dimensional polyurethane polymer reinforcement of the ballast", *Construction and Building Materials*, Vol. 44, pp. 615–625.

Khajehei, H., Ahmadi, A., Soleimanmeigouni, I. and Nissen, A. (2019), "Allocation of effective maintenance limit for railway track geometry", *Structure and Infrastructure Engineering*, Taylor & Francis, Vol. 15 No. 12, pp. 1597–1612.

Khosravi, M., Soleimanmeigouni, I., Ahmadi, A. and Nissen, A. (2021), "Reducing the positional errors of railway track geometry measurements using alignment methods: A comparative case study", *Measurement*, Vol. 178, p. 109383.

Khouy, I.A., Larsson-Kråik, P.O., Nissen, A., Juntti, U. and Schunnesson, H. (2014), "Optimisation of track geometry inspection interval", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 228 No. 5, pp. 546–556.

Kostoff, R.N. and Schaller, R.R. (2001), "Science and technology roadmaps", *IEEE Transactions on Engineering Management*, Vol. 48 No. 2, pp. 132–143.

Kour, R., Karim, R. and Thaduri, A. (2019), "Cybersecurity for railways – A maturity model", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 234 No. 10, pp. 1129–1148.

Krylov, V. V, Dawson, A.R., Heelis, M.E. and Collop, A.C. (2000), "Rail movement and ground waves caused by high-speed trains approaching track-soil critical velocities", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 214 No. 2, pp. 107–116.

Kumar, S., Espling, U. and Kumar, U. (2008), "Holistic procedure for rail maintenance in Sweden", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 222 No. 4, pp. 331–344.

Kumara, J.J. and Hayano, K. (2016), "Deformation characteristics of fresh and fouled ballasts subjected to tamping maintenance", *Soils and Foundations*, Elsevier, Vol. 56 No. 4, pp. 652–663.

Landgraf, M. and Hansmann, F. (2019), "Fractal analysis as an innovative approach for evaluating the condition of railway tracks", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 233 No. 6, pp. 596–605.

Larose, D.T. and Larose, C.D. (2015), *Data Mining and Predictive Analytics*, John Wiley & Sons, Incorporated, New York, USA.

Latvala, J., Nurmikolu, A. and Luomala, H. (2016), "Problems with Railway Track Drainage in Finland", *Procedia Engineering*, Elsevier B.V., Vol. 143, pp. 1051–1058.

Lee, J.S., Hwang, S.H., Choi, I.Y. and Choi, Y. (2020), "Deterioration Prediction of Track Geometry Using Periodic Measurement Data and Incremental Support Vector Regression Model", *Journal of Transportation Engineering, Part A: Systems*, Vol. 146 No. 1, pp. 04019057–1.

Li, D. and Davis, D. (2005), "Transition of Railroad Bridge Approaches", *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 131 No. 11, pp. 1392–1398.

Li, D., Hyslip, J., Sussmann, T. and Chrismer, S. (2015), *Railway Geotechnics*, 1st ed., CRC Press, London.

Li, D. and Selig, E.T. (1995), "Evaluation of railway subgrade problems", *Transportation Research Record*, No. 1489, pp. 17–25.

Li, D. and Selig, E.T. (1996), "Cumulative Plastic Deformation for Fine-Grained Subgrade Soils", *Journal of Geotechnical Engineering*, Vol. 122 No. 12, pp. 1006–1013.

Li, S., Liu, R., Li, C. and Wang, F. (2019), "Recovery Measure of Tamping on Different Track Geometry Irregularity Indexes", *Journal of Transportation Engineering, Part A: Systems*, Vol. 145 No. 11, p. 05019005.

Li, X., Ekh, M. and Nielsen, J.C.O. (2016), "Three-dimensional modelling of differential railway track settlement using a cycle domain constitutive model", *International Journal for Numerical and Analytical Methods in Geomechanics*, John Wiley & Sons, Ltd, Vol. 40 No. 12, pp. 1758–1770.

Li, Z., Zhao, X., Dollevoet, R. and Molodova, M. (2008), "Differential wear and plastic deformation as causes of squat at track local stiffness change combined with other track short defects", *Vehicle System Dynamics*, Vol. 46 No. 1, pp. 237–246.

Lichtberger, B. (2005), *Track Compemendium*, 1 st., Eurail Press.

Liikennevirasto. (2017), *Eurokoodin Soveltamisohje - Geotekninen Suunnittelu - NCCI 7 (Eurocode Application Guide - Geotechnical Design - NCCI 7)*, available at: https://ava.vaylapilvi.fi/ava/Julkaisut/Liikennevirasto/lo_2013-35_ncci7_web.pdf.

Liikennevirasto. (2018), *Ratatekniset Ohjeet (RATO) Osa 3 Radan Rakenne (Railway Technical Guidelines (RATO) Part 3 Track Structure)*, 13/2018, Väylävirasto, available at: https://ava.vaylapilvi.fi/ava/Julkaisut/Liikennevirasto/lo_2018-13_rato3_web.pdf.

Liu, X., Saat, M.R. and Barkan, C.P.L. (2012), "Analysis of causes of major train derailment and their effect on accident rates", *Transportation Research Record*, No. 2289, pp. 154–163.

Lundqvist, A. and Dahlberg, T. (2005), "Load impact on railway track due to unsupported sleepers", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 219 No. 2, pp. 67–77.

Luomala, H. (2019), *Monitoring the Vertical Deformation Behaviour of Road and Railway Structures*, Doctoral dissertation, Tampere University, available at: https://trepo.tuni.fi//handle/10024/115341.

Luomala, H., Rantala, T., Kolisoja, P. and Mäkelä, E. (2017), "Assessement of track quality using continuous track stiffness measurements", *Georail 2017, 3rd International Symposium Railway Geotechnical Engineering*, IFSTTAR, Marne La Vallee, France, pp. 281–290.

Maier, A.M., Moultrie, J. and Clarkson, P.J. (2012), "Assessing organizational capabilities: Reviewing and guiding the development of maturity grids", *IEEE Transactions on Engineering Management*, IEEE, Vol. 59 No. 1, pp. 138–159.

Martey, E.N. and Attoh-Okine, N. (2018), "Modeling tamping recovery of track geometry using the copula-based approach", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 8, pp. 2079–2096.

McShane, B.B., Gal, D., Gelman, A., Robert, C. and Tackett, J.L. (2019), "Abandon Statistical Significance", *American Statistician*, Vol. 73 No. 1, pp. 235–245.

Mehrali, M., Esmaeili, M. and Mohammadzadeh, S. (2020), "Application of data mining techniques for the investigation of track geometry and stiffness variation", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 234 No. 5, pp. 439–453.

Meier-Hirmer, C., Senée, A., Riboulet, G., Sourget, F. and Roussignol, M. (2006), "A decision support system for track maintenance", *WIT Transactions on the Built Environment*, Vol. 88, pp. 217–226.

Mercier, S., Meier-Hirmer, C. and Roussignol, M. (2012), "Bivariate Gamma wear processes for track geometry modelling, with application to intervention scheduling", *Structure and Infrastructure Engineering*, Vol. 8 No. 4, pp. 357–366.

Meriam Webster. (2022), “data”, *Merriam-Webster.Com*, retrieved 10.9.2022, available at: https://www.merriam-webster.com.

Mermec. (2022), “ROGER 1000”, retrieved 10.9.2022, available at: https://www.mermecgroup.com/measuring-trains-br-and-systems/recording-cars/104/roger-1000.php.

Messaadi, M., Grossoni, I., Shackleton, P., Shevtsov, I., Bezin, Y. and Dollevoet, R. (2021), “Rail degradation due to thermite weld discontinuities: Field experience”, *Engineering Failure Analysis*, Elsevier Ltd, Vol. 128 February.

Mezher, S.B., Connolly, D.P., Woodward, P.K., Laghrouche, O., Pombo, J. and Costa, P.A. (2016), “Railway critical velocity - Analytical prediction and analysis”, *Transportation Geotechnics*, Elsevier Ltd, Vol. 6, pp. 84–96.

Miao, Q., Niu, F., Lin, Z., Luo, J. and Liu, M. (2020), “Comparing frost heave characteristics in cut and embankment sections along a high-speed railway in seasonally frozen ground of Northeast China”, *Cold Regions Science and Technology*, Elsevier, Vol. 170, p. 102921.

Mirabadi, A. and Sharifian, S. (2010), “Application of association rules in Iranian Railways (RAI) accident data analysis”, *Safety Science*, Elsevier Ltd, Vol. 48 No. 10, pp. 1427–1435.

Neuhold, J., Vidovic, I. and Marschnig, S. (2020), “Preparing Track Geometry Data for Automated Maintenance Planning”, *Journal of Transportation Engineering, Part A: Systems, American Society of Civil Engineers*, Vol. 146 No. 5, p. 4020032.

Niazi, M., Wilson, D. and Zowghi, D. (2005), “A maturity model for the implementation of software process improvement: An empirical study”, *Journal of Systems and Software*, Vol. 74, pp. 155–172.

Nielsen, J.C.O., Berggren, E.G., Hammar, A., Jansson, F. and Bolmsvik, R. (2020), “Degradation of railway track geometry – Correlation between track stiffness gradient and differential settlement”, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 234 No. 1, pp. 108–119.

Nurmikolu, A. and Kolisoja, P. (2005), “Extruded polystyrene (XPS) foam frost insulation boards in railway structures”, *Proceedings of the 16th International Conference on Soil Mechanics and Geotechnical Engineering*, pp. 1761–1764.

Paixão, A., Fortunato, E. and Calçada, R. (2015), “Design and construction of backfills for railway track transition zones”, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 229 No. 1, pp. 58–70.

Paixão, A., Varandas, J.N. and Fortunato, E. (2021), “Dynamic Behavior in Transition Zones and Long-Term Railway Track Performance”, *Frontiers in Built Environment*, Vol. 7 March, pp. 1–16.

Paulk, M.C., Curtis, B., Chrissis, M.B. and Weber, C. V. (1993), “Capability maturity model, version 1.1”, *IEEE Software*, Vol. 10 No. 4, pp. 18–27.

Phaal, R., Farrukh, C.J.P. and Probert, D.R. (2007), "Strategic roadmapping: A workshop-based approach for identifying and exploring strategic issues and opportunities", *EMJ - Engineering Management Journal*, Vol. 19 No. 1, pp. 3–12.

Pitterman, B. (2000), "Telcordia Technologies: the journey to high maturity", *IEEE Software*, Vol. 17 No. 4, pp. 89–96.

Plasser & Theurer. (2022a), "EM160", retrieved 10.9.2022, available at: https://www.plassertheurer.com/en/machine/track-recording-cars/em160.

Plasser & Theurer. (2022b), "Absolute track measurement", retrieved 10.9.2022, available at: https://aktuell.plassertheurer.com/en/aktuell135/absolute-track-measurement.

Plasser & Theurer. (2022c), "09-3X Dynamic E3", retrieved 10.9.2022, available at: https://www.plassertheurer.com/en/machine/track-tamping-machines/09-3x-dynamic-e3.

Poeppelbuss, J., Niehaves, B., Simons, A. and Becker, J. (2011), "Maturity Models in Information Systems Research: Literature Search and Analysis", *Communications of the Association for Information Systems*, Vol. 29, available at:https://doi.org/10.17705/1cais.02927.

Pylkkänen, K. and Nurmikolu, A. (2015), *Routa Ja Routiminen Ratarakenteessa (Frost and Frost Action in Track Structures)*, Liikennevirasto.

Quiroga, L.M. and Schnieder, E. (2012), "Monte Carlo simulation of railway track geometry deterioration and restoration", *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, Vol. 226 No. 3, pp. 274–282.

Rauch, J. (2013), *Observational Calculi and Association Rules*, 1st ed., Vol. 469, Springer Berlin Heidelberg, Berlin, Heidelberg, available at:https://doi.org/10.1007/978-3-642-11737-4.

Röglinger, M., Pöppelbuß, J. and Becker, J. (2012), "Maturity models in business process management", *Business Process Management Journal*, Vol. 18 No. 2, pp. 328–346.

Sadeghi, J. and Askarinejad, H. (2009), "An investigation into the effects of track structural conditions on railway track geometry deviations", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 223 No. 4, pp. 415–425.

Sammouri, W., Côme, E., Oukhellou, L., Aknin, P. and Fonlladosa, C.E. (2013), "Floating train data systems for preventive maintenance: A data mining approach", *Proceedings of 2013 International Conference on Industrial Engineering and Systems Management, IEEE - IESM 2013, International Institute for Innovation*, Industrial Engineering and Entrepreneurship - I4e2.

Sánchez, A., Bravo, J.L. and González, A. (2017), "Estimating the Accuracy of Track-Surveying Trolley Measurements for Railway Maintenance Planning", *Journal of Surveying Engineering*, Vol. 143 No. 1, p. 05016008.

Sañudo, R., Dell'Olio, L., Casado, J.A., Carrascal, I.A. and Diego, S. (2016), "Track transitions in railways: A review", *Construction and Building Materials*, Vol. 112, pp. 140–157.

Sato, Y. (1995), "Japanese Studies on Deterioration of Ballasted Track", *Vehicle System Dynamics*, Taylor & Francis, Vol. 24 No. 1, pp. 197–208.

Scanlan, K.M., Hendry, M.T., Martin, C.D. and Schmitt, D.R. (2018), "The spatial correlation between track roughness and ground-penetrating radar inferred ballast degradation", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 7, pp. 1917–1931.

Schraven, D., Hartmann, A. and Dewulf, G. (2011), "Effectiveness of infrastructure asset management: Challenges for public agencies", *Built Environment Project and Asset Management*, Vol. 1 No. 1, pp. 61–74.

Selig, E.T. and Li, D. (1994), "Track modulus: its meaning and factors influencing it", *Transportation Research Record*, Vol. 1 No. 1470, pp. 47–54.

SFS-EN 13848-1 + A1. (2019), "Railway applications. Track. Track geometry quality. Part 1: Characterization of track geometry", European commitee for standardization.

SFS-EN 13848-2. (2020), "Railway applications. Track. Track geometry quality. Part 2: Measuring systems. Track recording vehicles", European commitee for standardization.

Silvast, M., Nurmikolu, A., Wiljanen, B. and Levomaki, M. (2010), "An inspection of railway ballast quality using ground penetrating radar in Finland", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 224 No. 5, pp. 345–351.

Silvast, M., Nurmikolu, A., Wiljanen, B. and Levomäki, M. (2013), "Identifying frost-susceptible areas on Finnish railways using the ground penetrating radar technique", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 227 No. 1, pp. 3–9.

Soleimanmeigouni, I., Ahmadi, A., Arasteh Khouy, I. and Letot, C. (2018), "Evaluation of the effect of tamping on the track geometry condition: A case study", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 2, pp. 408–420.

Soleimanmeigouni, I., Ahmadi, A. and Kumar, U. (2018), "Track geometry degradation and maintenance modelling: A review", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 1, pp. 73–102.

Soleimanmeigouni, I., Ahmadi, A., Nissen, A. and Xiao, X. (2020), "Prediction of railway track geometry defects: a case study", *Structure and Infrastructure Engineering*, Taylor & Francis, Vol. 16 No. 7, pp. 987–1001.

Suiker, A.S.J. and de Borst, R. (2003), "A numerical model for the cyclic deterioration of railway tracks", *International Journal for Numerical Methods in Engineering*, Vol. 57 No. 4, pp. 441–470.

Sussmann, T.R., Ebersöhn, W. and Selig, E.T. (2001), "Fundamental nonlinear track load-deflection behavior for condition evaluation", *Transportation Research Record*, Vol. 3 No. 1742, pp. 61–67.

Sussmann, T.R., Selig, E.T. and Hyslip, J.P. (2003), "Railway track condition indicators from ground penetrating radar", *NDT & E International*, Vol. 36 No. 3, pp. 157–167.

Suzuki, T., Ishida, M., Abe, K. and Koro, K. (2005), "Measurement on dynamic behaviour of track near rail joints and prediction of track settlement", *Quarterly Report of RTRI (Railway Technical Research Institute)*, Vol. 46 No. 2, pp. 124–129.

Szmagliński, J., Wilk, A., Koc, W., Karwowski, K., Chrostowski, P., Skibicki, J., Grulkowski, S., et al. (2022), "Verification of Satellite Railway Track Position Measurements Making Use of Standard Coordinate Determination Techniques", *Remote Sensing*, 14, 1855.

Tanaka, H., Yamamoto, S., Oshima, T. and Miwa, M. (2018), "Methods for detecting and predicting localized rapid deterioration of track irregularity based on data measured with high frequency", *Quarterly Report of RTRI (Railway Technical Research Institute)*, Vol. 59 No. 3, pp. 169–175.

Thoring, K., Mueller, R.M. and Badke-Schaub, P. (2020), "Workshops as a research method: Guidelines for designing and evaluating artifacts through workshops", *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 5036–5045.

Turunen, E. (2018), "Paraconsistent Many-Valued Logic in GUHA Framework", Acta Informatica Pragensia, *Acta Informatica Pragensia*, Vol. 7 No. 1, pp. 104–111.

UIC. (2002), *Rail Defects UIC Code 712 R.*

UIC. (2008), *Best Practice Guide for Optimum Track Geometry Durability.*

UIC. (2016), *UIC Railway Application Guide: Practical Implementation of Asset Management through ISO55001*, UIC Railway Technical Publications, available at: http://www.nvfnorden.org/library/Files/Utskott-och-tema/Drift-och-underhall/Møter-og-protokoller/UIC ISO 55000 Guidelines_final.pdf.

Vale, C. and M. Lurdes, S. (2013), "Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line", *Reliability Engineering and System Safety*, Vol. 116, pp. 91–98.

Väylävirasto. (2020), *Rataverkon Tavoiteltava Kehityskuva Vuoteen 2050 (Railway Network Vision of Development into the Year 2050)*, available at: https://urn.fi/URN:ISBN:978-952-317-798-7.

Wang, Y., Wang, P., Wang, X. and Liu, X. (2018), "Position synchronization for track geometry inspection data via big-data fusion and incremental learning", *Transportation Research Part C: Emerging Technologies*, Elsevier, Vol. 93, pp. 544–565.

Wendler, R. (2012), "The maturity of maturity model research: A systematic mapping study", *Information and Software Technology*, Elsevier B.V., Vol. 54 No. 12, pp. 1317–1339.

Williams, B.A., Holder, D., Edwards, J.R., Dersch, M.S. and Barkan, C.P.L. (2016), "Quantification of the lateral forces in concrete sleeper fastening systems", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 230 No. 7, pp. 1714–1721.

Xu, P., Liu, R., Sun, Q. and Jiang, L. (2015), "Dynamic-Time-Warping-Based Measurement Data Alignment Model for Condition-Based Railroad Track Maintenance", *IEEE Transactions on Intelligent Transportation Systems*, IEEE, Vol. 16 No. 2, pp. 799–812.

Yurlov, D., Zarembski, A.M., Attoh-Okine, N., Palese, J.W. and Thompson, H. (2019), "Probabilistic Approach for Development of Track Geometry Defects as a Function of Ground Penetrating Radar Measurements", *Transportation Infrastructure Geotechnology, Transportation Infrastructure Geotechnology*, Vol. 6 No. 1, pp. 1–20.

Zarembski, A.M., Einbinder, D. and Attoh-Okine, N. (2016), "Using multiple adaptive regression to address the impact of track geometry on development of rail defects", *Construction and Building Materials*, Vol. 127, pp. 546–555.

Zhang, Y.-J., Murray, M. and Ferreira, L. (2000), "Modelling rail track performance: an integrated approach", *Proceedings of the Institution of Civil Engineers - Transport*, Vol. 141 No. 4, pp. 187–194.

# PUBLICATION
# I

## Advancing railway asset management using track geometry deterioration modelling visualization

Sauni, M., Luomala, H., Kolisoja, P. and Nummi, T.

Journal of Transportation Engineering, Part A: Systems, Vol. 148 No. 2: 04021106.
10.1061/JTEPBS.0000626

# Advancing Railway Asset Management Using Track Geometry Deterioration Modeling Visualization

Mikko Sauni[1]; Heikki Luomala[2]; Pauli Kolisoja[3];
and Tapio Nummi[4]

---

**Abstract:** Railway tracks need to be monitored to ensure safe operations and cost-effective maintenance. The monitoring is commonly conducted using a track recording car that describes deviations from an ideal track geometry. Over time, the measurements provide time series data that can be used to model the observed track geometry deterioration process. However, without simplification, the modeling results are generally too complex to be utilized to their full extent in track asset management. Therefore, this study aimed to implement visualization techniques for track geometry deterioration modeling results analysis which benefit track asset management. The best practices on track geometry deterioration modeling were studied and applied to the track geometry history of a track section located in Finland. After testing the establishing modeling principles, proposals were made regarding the use of the results in practice. This paper presents visualization techniques that use the modeling results of individual cross-sections to generate information about longer sections of track and even whole rail networks. These visualizations digest the massive amount of information from the modeling and present it in an informative way for practitioners to utilize and benefit from. Thus, this study fills the gap between research and practice in railway track geometry deterioration modeling. **DOI: 10.1061/JTEPBS.0000626.**

---

## Introduction

Multiple studies have demonstrated that track geometry deterioration is not a random process, but one that can be idealized and modelled [see Higgins and Liu (2018) and Soleimanmeigouni et al. (2018) for literature reviews]. Fig. 1 presents the idealized track geometry deterioration behavior of one cross section, which is always tamped at the exact same longitudinal level (LL) deviation value, and tamping always corrects the geometry to the original level. The figure demonstrates the theoretical diminishing effect of tamping due to fouling of ballast (Shenton 1985; Dahlberg 2001; Lichtberger 2011). In Phase 1 of Fig. 1, initial settlements increase deviations at an exponential pace. This is sometimes referred to as ballast memory. Following the initial settlement, Phase 2 describes a linear deterioration path that generally ends when a tamping intervention limit is reached. This is followed by a theoretical failing phase (3), which describes the track end-of-life, but in practice, this phase is avoided either by conducting maintenance or ceasing traffic.

Neuhold et al. (2020) provided a foundation for modeling track geometry deterioration based on actual track geometry car measurements. This approach was based on modeling the behavior of one cross section, thus providing results for a localized point on a track section, as longer sections of track were out of scope in that research. The current study extends the work of Neuhold et al. (2020) and demonstrates how track geometry deterioration modeling of cross sections can be utilized for investigating the behavior of not only cross sections, but also of longer sections of track and even the whole rail network.

Modeling track geometry deterioration based on track geometry car measurements provides highly practical information about the development of the condition of railway tracks. However, the real-world benefits of track geometry deterioration modeling can be obtained only if the modeling results are made accessible and understandable to practitioners in asset management. For this purpose, the results need to be generalized into key figures and representative visualizations that serve the heuristic nature of decision making in track asset management. Otherwise, the results achieved in academia will not have an impact in practice. The, this study also investigated indicators and visualizations created from track geometry deterioration modeling that would be beneficial to practical track asset management. The purpose was to ease the interpretation of the modeling results by providing summarized information for decision making, thus filling the gap between research and practice on track geometry deterioration modeling.

Two research questions were formed based on these research gaps:

1. How to use cross-section-based track geometry deterioration modeling for longer sections of track and for the whole rail network?
2. How to present track geometry deterioration modeling results in a way that practitioners can easily interpret and benefit from?

The scope of this paper is limited to visualizing stochastic long-term modeling of longitudinal deviations measured periodically using a track geometry measurement car. The purpose of this paper is to adapt the best current practices of track geometry deterioration modeling and bring them closer to practical application using

---

[1] Doctoral Researcher, Research Centre Terra, Tampere Univ., Korkeakoulunkatu 7, Tampere FI-33720, Finland (corresponding author). ORCID: https://orcid.org/0000-0002-7034-5835. Email: mikko.sauni@tuni.fi

[2] Project Manager, Research Centre Terra, Tampere Univ., Korkeakoulunkatu 7, Tampere FI-33720, Finland. ORCID: https://orcid.org/0000-0002-7113-3527. Email: heikki.luomala@tuni.fi

[3] Professor, Research Centre Terra, Tampere Univ., Korkeakoulunkatu 7, Tampere FI-33720, Finland. Email: pauli.kolisoja@tuni.fi

[4] University Lecturer, Faculty of Information Technology and Communication Sciences, Tampere Univ., Kalevantie 4, Tampere FI-33100, Finland. Email: tapio.nummi@tuni.fi

Note. This manuscript was submitted on June 28, 2021; approved on September 30, 2021; published online on November 24, 2021. Discussion period open until April 24, 2022; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering, Part A: Systems***, © ASCE, ISSN 2473-2907.

J. Transp. Eng., Part A: Systems, 2022, 148(2): 04021106

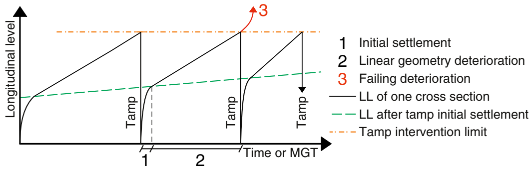


**Fig. 1.** Theoretical behavior of track geometry.

visualization techniques, not to create completely new modeling techniques or to validate current models.

## Track Geometry Deterioration Modeling

The best practices for preparing data for track geometry deterioration modeling were elaborated by Neuhold et al. (2020). This study follows these established practices as closely as possible, but some adjustments needed to be made to suit the data measured in Finland. The purpose was not to advance the methods presented by Neuhold et al. (2020), but rather to alter the methods for the available data. Table 1 summarizes the slight differences between the track geometry deterioration modeling described by Neuhold et al. (2020) and the methods used in this study. The following subsections further elaborate these differences.

### *Initial Data*

The data used for the demonstrations in this paper consists of a ten-year semi-annual track geometry car measurement history from a track section, Luumäki–Imatra, in Finland. The examined section is a 53 km long mixed traffic single line track section with a maximum speed of 140 km/h for passenger trains. The yearly gross tonnage of freight traffic is around 12 megatons. The measurements were preformed using an EM120 track recording car (Plasser & Theurer, Linz, Austria), which uses chord measurements from three bogies spaced 5 and 7 m apart. The data is recorded every 0.25 m. No major renewals were reported during this time period; only routine maintenance. Neuhold et al. (2020) used similar data, but their initial data was much more extensive, albeit measured more sparsely. Nevertheless, these differences do not matter, as the purpose of this study is to visualize the results, which does not require such a vast initial data set that statistical analyses require.

### *Track Geometry Parameter and Index*

For long-term track geometry deterioration modeling, the LL standard deviation (SD) is commonly the chosen parameter, as most of the gradual displacements occur in the vertical direction (UIC 2008; Vale and Calçada 2014). SD provides a smooth depiction of the original deviation signal, and is defined as follows:

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

where $x_i$ = the current value of a signal; $\bar{x}$ = the mean value of a signal; and $N$ = the number of values in a sample (Eurocode EN13848-6 2014). Neuhold et al. (2020) opted for a modified SD, in which both the left and right rail were considered simultaneously, and the result was multiplied by 1.35 to make the result comparable with a conventional SD. This approach could not be adopted for this study, because reliable data repositories were available only for the left rail. This does not affect the end results of this study, as the applied visualization techniques are applicable to examining each rail individually or simultaneously.

SD can be calculated in fixed (also referred to as segmented) or rolling (also referred to as moving or continuous) windows, using any distance. Fixed SD calculation windows tend to be easier to communicate, but they misrepresent information when deviations occur in the edges of windows or if there are only local irregularities in the middle of otherwise stable track, as demonstrated by Neuhold et al. (2020). The use of rolling windows was found suitable for the data in Neuhold et al. (2020), as well as for this study.

Adjusting the length of the rolling window influences the sharpness with which the SD follows the original signal (Fig. 2). The appropriate rolling SD calculation window length can be considered to be roughly between 10 and 200 m, based on the lengths used in previous research (Andrade and Teixeira 2011; Tanaka et al. 2018; Neuhold et al. 2020; Audley and Andrews 2013). A shorter window SD more sharply follows the original signal, but too short a window might result in the same problems as those encountered when using only the original signal, namely, instability in alignment and a fluctuating signal. Too long a window may cause similar problems to those faced when using fixed windows, where some irregularities may be hidden due to adjacent smooth track.

This study opted for a 200 m SD window, as there were some alignment issues between sequential measurements, as presented in Fig. 2. Neuhold et al. (2020) experienced similar problems and opted for a 100 m SD window. Furthermore, this study preferred the 200 m over the 100 m SD, because the 200 m SD is recognized by the European Standard 13848-6 (2014), which gives a good basis for standardizing the modeling principles and results.

### *Core Modeling Methods for Track Geometry Deterioration*

Modeling the track geometry deterioration rate (TGDR) on a large scale requires a general depiction of past behavior, for example, when analyzing the decade-long behavior of a track section. Therefore, tamping intervals are usually adopted as the minimum interval length for a deterioration period. This leads to a simplification of the TGDR by using some mathematical idealization. The core mathematical approaches to LL SD deterioration modeling are

**Table 1.** Differences in track geometry deterioration modeling between Neuhold et al. (2020) and this study

| Modeling principles | Neuhold et al. (2020) | This study |
|---|---|---|
| Initial data | 16-year track geometry car history from 4,400 km in 5 m increments | 10-year track geometry car history from 60 km in 0.25 m increments |
| Geometry index | LL 100 m continuous modified SD | LL 200 m continuous SD |
| Alignment correction | None | None |
| Modeling method | Linear regression | Robust linear regression |
| Outlier handling | Outlier detection algorithm MAD | Robust linear regression |
| Tamping activity identification | Tamping records and negative TGDR | Negative TGDR and tamping area identification algorithm |
| Prognosis accuracy measure | Prediction and real end quality comparison | Prediction interval |

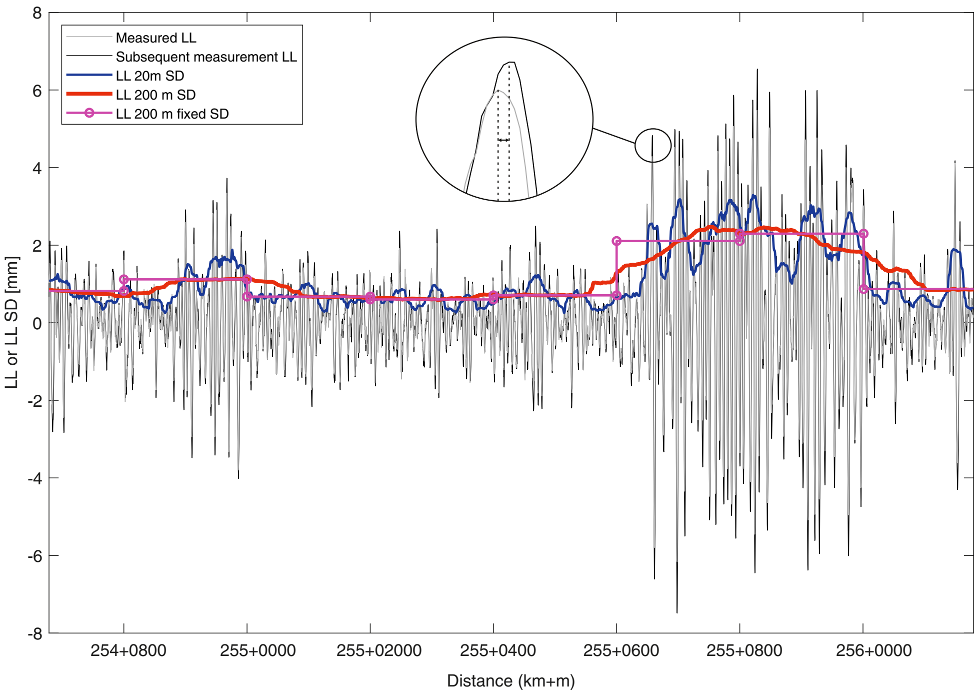


**Fig. 2.** Original measurement signal alignment and different indices.

linear and exponential modeling. Linear modeling of track geometry deterioration is the most popular modeling approach, based on the literature (Caetano and Teixeira 2015; Khajehei et al. 2019; Lee et al. 2018; Li et al. 2019; Neuhold et al. 2020; Nielsen et al. 2020; Soleimanmeigouni et al. 2020). In addition, it provides the best fit for the available data in Finland. However, it has been argued that exponential models suit some data sets better than linear models (Quiroga and Schnieder 2012; Famurewa et al. 2016). The use of exponential models can be justified by their ability to take the initial settlement into account. Generally, these models are suited better for track sections with high traffic volume and frequent (e.g., daily) track geometry measurements (Tanaka et al. 2018).

Real world measurements often provide outliers that need to be accounted for in modeling. Neuhold et al. (2020) used the mean absolute deviation (MAD) to identify and erase outliers. In this study, outliers could not be removed as in Finnish conditions they might indicate frost heave or other abrupt occurrences, which should be presented to asset managers. Therefore, outlier handling was considered when choosing the linear modeling method, not as a separate operation.

The linear regression modeling of geometry deterioration can be conducted using either simple or robust techniques. Simple techniques include algorithms such as least squares. These algorithms provide fast calculation times but tend to be influenced by outliers. However, outlier detection and removal algorithms can be run before using simple algorithms to improve the results. Robust algorithms, such as linear programming, generally produce numerous possible estimations and choose the best-fitting one. A plot of the differences between the algorithms (Fig. 3) depicts how in the second tamping interval (measurements between 2011 and 2017) the robust algorithm better describes the long-term behavior than least squares, when the initial settlement is not regarded.

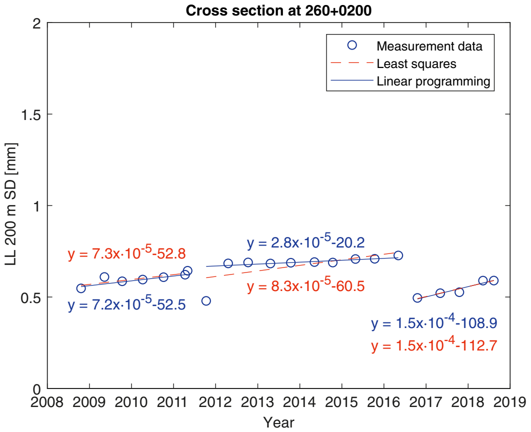


**Fig. 3.** Difference between a simple and robust linear optimization.

The major critique of all linear models is that they fail to capture the initial settlement behavior after tamping. The initial settlement lasts only days on a track section with heavy traffic, whereas the track geometry measurements are conducted every 2–6 months. Therefore, the probability of capturing the initial settlement caused by tamping in a track geometry measurement is low. This is not to say that it does not exist, but the recorded cases are few, and the effect of initial settlement on increasing deterioration is negligible when modeling on a large scale.

Fig. 3 presents a noticeable initial settlement captured by the track geometry measurements in 2011. However, the effects of the initial settlement on the linear regression can be eliminated by using a robust linear programming algorithm instead of simple techniques. Removing the effect of the initial settlement on linear regression modeling can be justified by arguing that capturing the long-term trend of the track geometry deterioration is more valuable than portraying the initial settlement. Furthermore, the initial settlement does not provide useful information about the deterioration, as the realized effect of tamping is the level of deviation after the initial settlement, as depicted in Fig. 1. Robust linear modeling ignores the initial settlement in the model, as is intended, but unlike outlier removal techniques, the outlier is left visible in the data, which is important, as these outliers can provide useful information to asset managers about the successfulness of tamping.

## Tamping Identification

There are two main causes for a decreasing TGDR, namely, measurement inaccuracy and tamping (with or without other maintenance actions). The measurement inaccuracy usually accounts for only the slightest deterioration decreases, which are most likely to occur when the TGDR is close to zero. In these cases, the fluctuation in the measurement results is mostly attributed to the measurement technique rather than an improvement of the physical track geometry. These deterioration rates should be considered as zero in deterioration modeling. The decreasing deterioration due to tamping is usually clearly noticeable, especially in cases where the TGDR is not close to zero. However, modeling the tamping, i.e., the decreasing deterioration rate due to tamping, is very difficult due to multiple and generally unknown variables related to the timing and reason for tamping.

Solving the problem of what constitutes as tamping instead of measurement noise in the track geometry data is important, as tamping intervals form the basis for track geometry deterioration modeling. This problem can be overcome by systematically recording tamping data, but when the records are be incomplete or missing altogether, such as is the case in Finland, the solution must be applied in track geometry deterioration modeling.

The simplest solution for revealing tamping in track geometry measurement data is to set a threshold for the decrease in the track geometry that will be considered tamping, as proposed by Neuhold et al. (2020). The threshold can be a fixed value or be dependent on time or deterioration level. A threshold value is adequate for detecting tamping that has had a significant effect on the irregularity level, which is suitable especially for segmented track geometry indices. However, when the LL SD is low before the tamping or the tamping has little effect on the LL SD, a threshold value will not consider these cases as tamping. This modeling case is common when modeling track geometry using a rolling SD. This issue is best demonstrated when the effect of tamping has roughly the same value as the limit for detecting tamping (Fig. 4). In these cases, on some of the cross sections in the tamped area a tamping is noticed, but on others, it is not. This leads to undesirable results, which are apparent when the sum of tamping times per cross section and the cross section LL SD histories are plotted (Fig. 4). These results will not be altered even with the use of a threshold dependent on the time or the irregularity level. All fixed thresholds for determining tamping will fail to separate tamping from measurement fluctuation in the cases of low irregularity levels, as the tamping effect on the irregularity is about the same amount as the measurement fluctuation. For that reason, an algorithm is required to determine these cases.

The algorithm created in this study searches the data for small individual areas with corrections in track geometry and disregards them as tamping if they are not associated with an adjacent tamped area. The generic form of the code is displayed in Fig. 5. The algorithm does not consider areas to be tamped if the area is less than a minimum length, which was set at 20 m in this study. Shorter areas with negative TGDRs were considered to have been caused by measurement inaccuracy. The accuracy of the algorithm could not be numerically verified, as there are no systematic historical tamping records in Finland. Verifying and improving the tamping detection algorithm is a source of further research.

## Tamping Effect Quantification

The decrease in the level of irregularity calculated from track geometry car measurements before and after tamping indicates whether

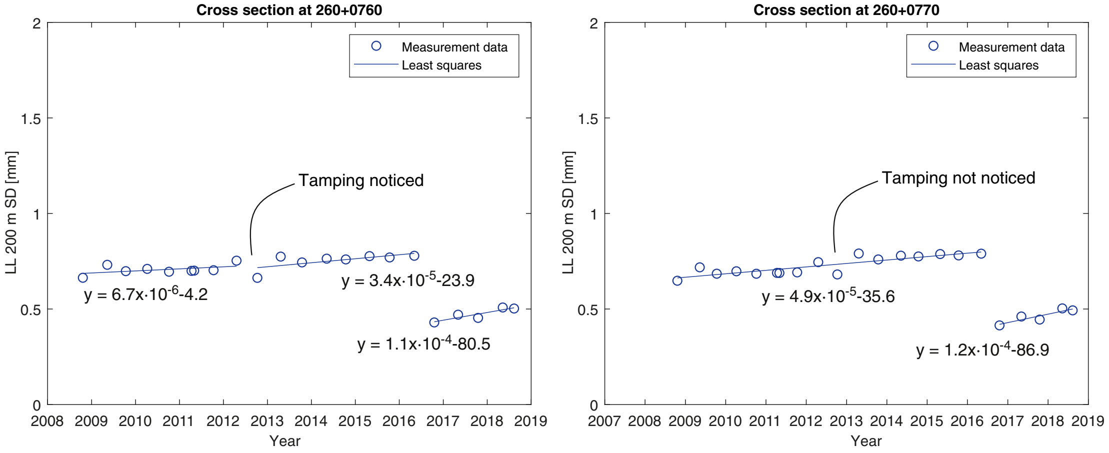


**Fig. 4.** Two close-by cross-sections with different interpretations of tamping history.

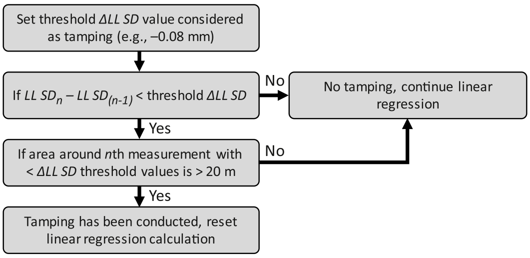


**Fig. 5.** Generic algorithm for detecting tamped areas.

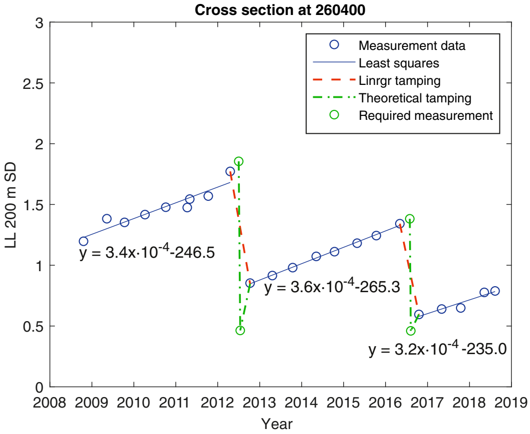


**Fig. 6.** Theoretical and modelled tamping effects.

the tamping has successfully provided a stable ballast layer for the trains to pass, or whether the initial settlement cancelled much of the smoothness provided in the tamping. For example, in Fig. 6, tamping has had a meaningful effect on the level of irregularities, but in Fig. 3, irregularity has returned to the original state after one measurement in the second tamping interval, albeit the irregularities have stayed at a low level throughout the history.

The effects of tamping, without considering the initial settlement after tamping, can be calculated using the information obtained from a robust linear regression model. A robust model is not prone to outliers, which means that if the initial settlement is captured in the first measurement after tamping, it will not be regarded in the results. Thus, the tamping effect describes the effect the tamping has had when regarding the long-term behavior.

The effect is calculated by extrapolating the regression lines before and after tamping to the time of tamping and calculating the difference in their irregularity level (Fig. 7). By doing so, the effect of tamping on the level of deterioration $\sigma_t$ is

$$\sigma_t = \sigma_{tb} - \sigma_{ta} \tag{2}$$

where $\sigma_{tb}$ denotes the level of deterioration calculated from the line equation before tamping at the time of tamping $\Delta\sigma_{n-1}(t_t)$; and $\sigma_{ta}$ denotes the level of deterioration calculated from the line equation after tamping at the time of tamping $\Delta\sigma_n(t_t)$. If the exact time of tamping $t_t$ is not known, it can be assumed that reasonable results can be obtained by

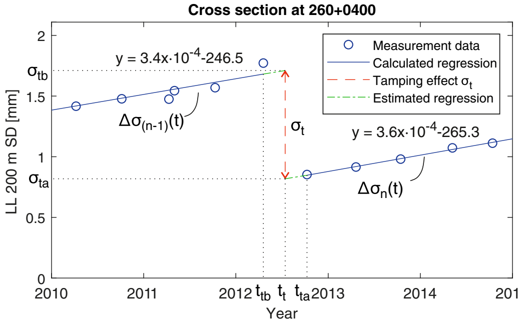


**Fig. 7.** Tamping effect calculation principle.

$$t_t = t_{tb} + (t_{ta} - t_{tb})/2 \tag{3}$$

where $t_a$ denotes the time of the first measurement after tamping; and $t_b$ denotes the time of the last measurement before tamping.

The TGDR after tamping is another important measure for determining the effectiveness of tamping. Tamping effectiveness can be determined by examining the deterioration rates before and after tamping. Higher deterioration rates after tamping indicate that problems continue after tamping, and thus, other maintenance actions together with tamping might be required in the future. Approximately equal deterioration rates before and after tamping indicate that tamping has reset the deterioration trend to a lower level, but the deterioration continues to develop as before. Deterioration rates that are lower after tamping than before tamping indicate that the tamping has had a remedial effect on the track structure.

The change in the TGDR ($\Delta\sigma_t$) can be assessed by comparing the slopes of the linear track geometry deterioration models before and after tamping. There are three approaches to comparing the deterioration rates:

Absolute comparison

$$\Delta\sigma_{t\text{Absolute}} = \Delta\sigma_{n-1} - \Delta\sigma_n \tag{4}$$

Relative comparison

$$\Delta\sigma_{t\text{Relative}} = \Delta\sigma_{n-1}/\Delta\sigma_n \tag{5}$$

Normalized absolute comparison

$$\Delta\sigma_{t\text{Normalized}} = (\Delta\sigma_{n-1} - \Delta\sigma_n)/\Delta\sigma_{n-1} \tag{6}$$

where $\Delta\sigma_n$ denotes the deterioration rate of tamping interval $n$; and $\Delta\sigma_{n-1}$ denotes the deterioration in the previous tamping interval. Because deterioration rates before and after tamping can be close to zero, relative or normalized absolute comparisons can result in very large or small values, due to the divisor or the dividend being close to zero, respectively. As an alternative, the absolute comparison does not suffer from this mathematical nuisance. However, because the absolute comparison does not normalize the deterioration rate of the previous tamping interval, the results should always be presented with knowledge of the level of the deterioration rate. Otherwise, the tamping effect might seem optimistic, if the deterioration rate before tamping has been very high. If information about the level of deterioration is incorporated, the absolute comparison is the most practical comparison to use. Otherwise, relative or normalized absolute comparison yields more informative results.

J. Transp. Eng., Part A: Systems, 2022, 148(2): 04021106

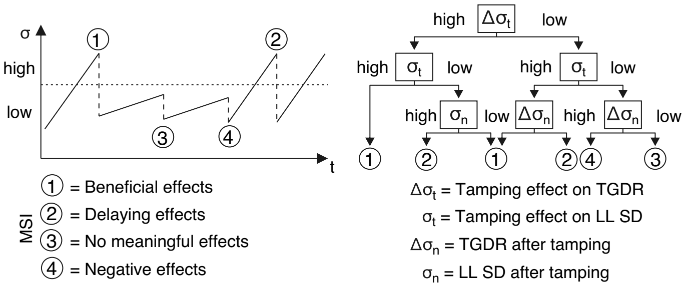


**Fig. 8.** Maintenance success indicator (MSI) logic.

Evaluating past tamping effectiveness requires examining several figures simultaneously: the values of LL SD and TGDR after tamping, and the change in LL SD and TGDR due to tamping. Because simultaneous examination of four figures is not practical, an ensemble parameter representing their possible combinations is required. The ensemble parameter, named maintenance success indicator (MSI), can be assigned to represent four different outcomes of a tamping (Fig. 8): (1) beneficial, (2) delaying, (3) not meaningful, or (4) negative.

The logic behind MSI is presented in Fig. 8. The evaluation begins by assessing the effect of tamping on the TGDR ($\Delta\sigma_t$). If the effect on the TGDR ($\Delta\sigma_t$) is high, it means that the deterioration rate has slowed down, and vice versa. Next, the tamping effect on the deterioration level ($\sigma_t$) is evaluated. A high effect is the desirable outcome. Finally, the tamping interval after tamping is evaluated by assessing either the level of deterioration after tamping ($\sigma_n$) or the TGDR after tamping ($\Delta\sigma_n$), where low values are the desirable outcome. The limit values separating high and low $\Delta\sigma_t$, $\sigma_t$, $\sigma_n$, and $\Delta\sigma_n$ can be assessed using the allowable LL SD and limit values for TGDR. However, defining the limit values for these is out of scope and a source of future research, as the initial data for assessing the limit values should be much larger than the one available for this study.

A desirable outcome for the MSI is to have as much Class 1 (beneficial effects) tamping and as little of other classes. Areas with Class 3 or 4 effects should be closely investigated. The MSI Class 1 denotes that tamping has slowed down the TGDR and significantly reduced the LL SD, thus making the track behavior better than before tamping. MSI Class 1 can also be achieved, if the effect of tamping to TGDR or LL SD has been only slight if the TGDR and LL SD after tamping have been low. This implies that these values were low before tamping, therefore, they could not have been improved any further by tamping. Class 2 MSI implies that while the tamping has restored the LL SD, the TGDR is still high, meaning the tamping has been successful, but the remediation will not last. MSI Class 2 highlights areas where tamping is successful, but other maintenance actions are also required to obtain a lasting remediation. MSI Class 3 denotes that tamping has not made a significant difference, and perhaps the planned tamping areas should be revised, if possible. MSI Class 4 suggests that errors have been made in the tamping or that the track has suffered damage after tamping, and the area should be further investigated. Special cases, where a section is tamped before and after a measurement, must be considered separately as not applicable areas (n/a), as a TGDR cannot be calculated from a single track geometry measurement. The practical use of the MSI is presented in a later section, Visualizing Track Geometry Deterioration Modeling Results.

### *Prognosis and Prognosis Accuracy Measures*

Predicting future LL SD values using linear regression models is simple. Extrapolating the linear regression model of the newest available tamping interval usually provides reasonably good results. However, if an area has been recently tamped, the insufficient number of measurements after tamping may not accommodate linear regression. In these cases, the linear regression model can be based on the previous tamping interval.

However, while the predictions of future LL SD values based on linear regression are simple to make, the predicted values themselves are not informative enough. The reliability of the predictions must be described as linear estimations have varying degrees of uncertainty. Neuhold et al. (2020) assessed prognosis accuracy by comparing the predicted and real end quality. However, this study could not adopt such a method, as the prognosis accuracy needs to be expressed for a future prediction.

The best way to describe the reliability of future predictions is by using the prediction interval (PI). The PI offers a simple way to describe where future observations (LL SD values) produced by the model will occur, with a specified confidence. The PI can be described in a generic form as

$$\mathrm{PI} = \hat{\mu} \pm t_{\alpha/2,n-p}\, s\sqrt{1 + x_0'(\mathbf{X^tX})^{-1}\mathbf{x_0}} \qquad (7)$$

where $\hat{\mu}$ denotes the predicted timing of reaching a LL SD limit value; $t$ is the quantile of t-distribution having $df = n\text{–}p$; $\alpha$ is the specified confidence level (for example 90%); $s$ is the square root of the residual sum of squares; $(\mathbf{X^tX})^{-1}$ is the covariance matrix of parameters; and $\mathbf{x_0}$ is a column vector of the particular values of interest (LL SD limit value), at which the prediction is calculated (Agresti 2015).

The PI does not provide a probability for a future observation but instead describes the level of confidence of the model predictions. For example, a 90% PI provides the range where approximately 90% of future observations produced by the model should occur. The PI can be communicated easily by plotting the range until a maintenance limit value is met. For example, in Fig. 9, the 90% PI indicates that the set maintenance limit of 1 mm LL SD is met between 2020 and 2022, and the most likely timing for reaching the limit is in 2021.

## Visualizing Track Geometry Deterioration Modeling Results

Asset management requires simple-to-use information that provides a good overall representation of the track conditions. The

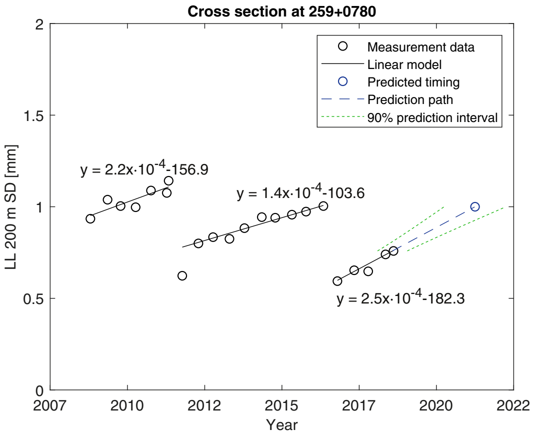


**Fig. 9.** Prediction of future LL SD and its PI.

information from track geometry deterioration modeling should be reported for asset management in three categories: past deterioration exploration, past maintenance effectiveness evaluation, and future tamping need predictions. The scope of observation for each category should consider at least three levels:

1. Cross section level
2. Track section level,
3. Network level.

The following sections elaborate what the different scopes and categories are used for and what novel information they produce.

### Cross Section Level

The cross section level scope provides the necessary tools for examining the track geometry deterioration of localized areas on a track section. The cross section here represents the 200 m LL SD values around that area. The analyses provide specific information about short problematic areas, for example, transition zones, which are more typical than long sections of poor condition track. The behavior of cross sections can be visualized in a time–LL SD perspective (Fig. 10). From these illustrations, the deterioration history can be observed, which includes the past changes in the TGDR and the past tamping times and their effects, and also, the timing of the next tamping can be predicted.

With these illustrations and results, the asset manager can answer, for instance, the following questions:

- Is the cross section problematic (high TGDR)?
- When had the problematic behavior begun?
- Are there seasonal differences in the TGDR?
- Has tamping been effective in maintaining good track geometry?
- When will the area around the cross section require further maintenance?

By examining the past track deterioration behavior, maintenance effectiveness, and the predicted next maintenance intervention timing, the asset manager can guide maintenance by choosing the correct approach for remediation and time the remediation. In practice, the illustrations show whether the location has been tamped multiple times with no lasting improvement to the track geometry. In these cases, the asset manager can assign further investigations to determine appropriate spot repair to remedy the problem instead of tamping the area once again with futile effects. Furthermore, the asset manager can investigate the origins and development of problematic behavior, like seasonal differences or sudden increases in the deterioration rates caused by track work or extreme weather conditions. The asset manager also attains information on when the next maintenance action should be taken at that location.

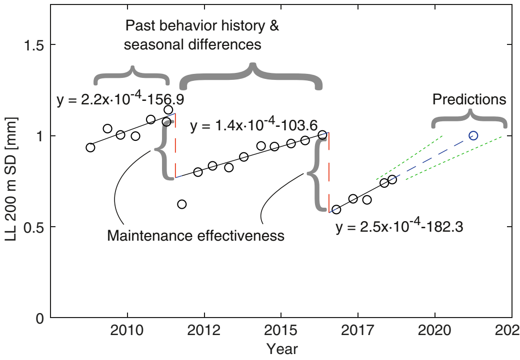


**Fig. 10.** Track geometry deterioration modeling results illustrated in cross-section level.

### Track Section Level

The track section-level scope of observation provides information on how the deterioration behavior differs within a longer section of track. The longer section of track is commonly between 1 and 10 km long, because even longer sections become difficult to assimilate. The goal is to detect problematic zones, so that the analysis can focus on those areas. The heatmaps of LL SDs [Figs. 11(a and c)], and tamping effects [Fig. 11(b and d)] can provide useful ways for analyzing the history of longer sections of track. The tamping effects here refer to the decrease in the LL 200 m SD.

These figures should be examined together, as examining them separately does not provide sufficient information for reliable analysis. For example, in Fig. 11, the LL SD in the area around 265 + 0300 and 265 + 0500 remained at a rather high level until tamping in 2016, after which the LL SD has been moderately high, but tamping has not been required. This would lead to the conclusion that this area poses no concerns currently. Conversely, even though the LL SD in the area around 265 + 0600 and 265 + 0800 is moderate, the frequent tamping suggests that some problems do exist, but the tamping interval is so short that the track geometry measurements do not capture the actual behavior. This area should be closely monitored in future measurements.

In addition to assessing past deterioration behavior, predictions of the future tamping areas and their timing should be kept updated for asset management. The predicted tamping areas can be plotted with the principle presented in Fig. 12. The *x*-axis indicates the location, whereas the *y*-axis indicates the timing. The plot is color-coded to represent the PI of the predicted timing of the next tamping; a darker color indicates a more likely timing of the next tamping.

In track maintenance, the tamping timing prediction illustrations, like the one shown Fig. 12, can be utilized to plan future tamping areas. The graph also provides a general idea about the reliability of the predictions. For example, in Fig. 12, the LL SD in the area between 254 + 0450 and 254 + 0650 reaches a maintenance limit value at around 2022, and the PI is quite narrow. This would be a good indicator to plan tamping at that area for that

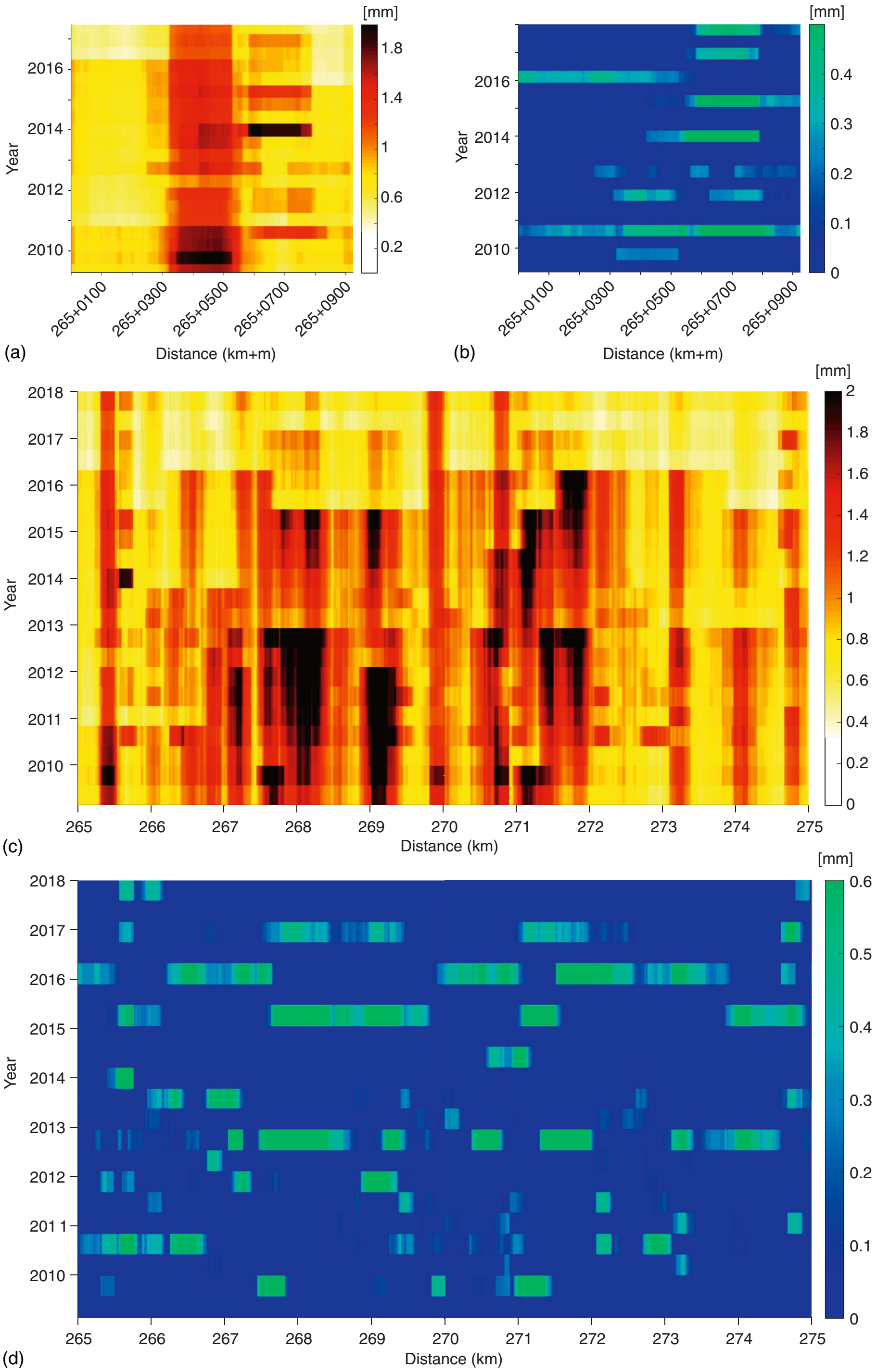


**Fig. 11.** Heatmaps of LL SD and tamping effects: (a) heatmap of LL 200 m SD (1 km); (b) heatmap of tamping effects (1 km); (c) heatmap of LL 200 m SD (10 km); and (d) heatmap of tamping effects (10 km).

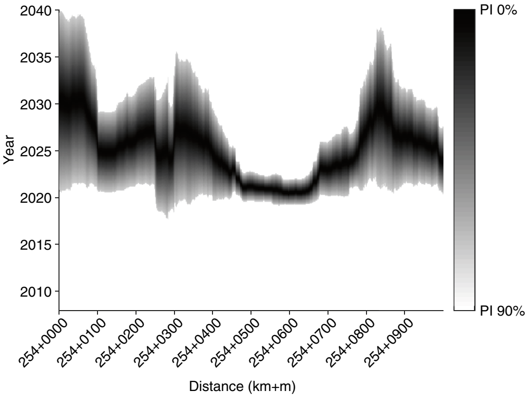


**Fig. 12.** Prediction of the next tamping timing.

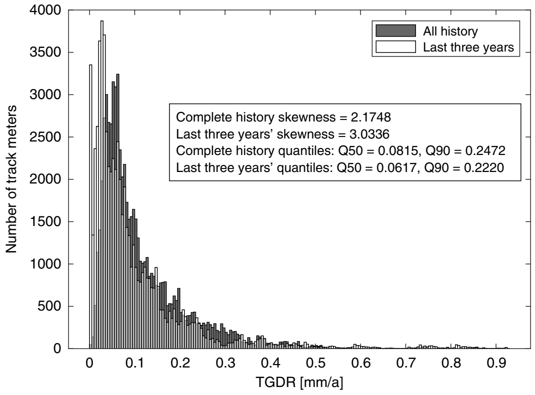


**Fig. 13.** Complete and last three years' geometry history of a track section.

year. As an example of another type of case, the area around 254 + 0000 and 254 + 0100 has a very wide PI, and it is predicted that tamping will be required around the year 2030. This area should not require tamping in the near future, but because the predictions are still ambiguous, the area should be focused on when new track geometry measurements are performed, and the predictions are updated.

With the information provided from the track section level analysis, the asset manager can answer the following questions:

- Where are the problematic areas located on a track section, and how severe are they?
- For how long has the problematic behavior been observed, and is it seasonal?
- How has past maintenance affected the track geometry deterioration on the track section?
- When will different parts of the track section require tamping or other maintenance?

The information obtained from illustrations following the principles shown in Figs. 11 and 12 can be used to guide maintenance by assessing the systemization of past maintenance, drafting tamping plans, and conducting the same analyses as for the cross section level, only now for a longer segment of track (e.g., 1–10 km). Past maintenance performance can be evaluated by visualizing the past tamping areas. Tamping areas disconnected from each other and frequently tamped areas indicate that tamping planning should be revised. Future tamping plans can be drafted using the illustrations by connecting areas with similar timing estimations for the next tamping. In addition, all the analyses mentioned for cross section level are valid for the track section level as well.

### *Network Level*

Network-level track geometry deterioration assessment needs to represent complete track sections in generalizing figures and illustrations. The network-level analysis is intended to apply to tens or even hundreds of kilometers of track. For this purpose, the analysis turns to the statistics of the selected network. The past track geometry deterioration behavior of a network can be evaluated using a histogram displaying the number of track meters where a certain mean TGDR has been observed (Fig. 13).

The histograms of the TGDRs observed on the track sections can be compared by plotting them in the same figure and then examining them. In addition to the visual examination of the histograms, key figures from the histograms can be produced to enhance the evaluation. The suggested approach is to report the median (50%) and 90% quantile of the distribution. The metrics on skewness should also be reported, as the histograms are often skewed to the right due to problematic areas exhibiting significantly larger deterioration rates compared to the median.

As an example of an analysis, Fig. 13 demonstrates two TGDR histograms from the same 60 km section of track: one containing the modeling of the complete 10 years' measurements, and the other containing the modeling of the last three years' measurements. At first glance, the last three years seem to be more to the left than the complete history, indicating that the conditions have improved on this section of track. This is supported by the lower median (50% quantile) and 90% quantile of the last three years compared with the complete history. However, the higher skewness on the history of the last three years suggests that the parts of the track enduring very poorly continue to be a problem.

The network-level modeling result analysis should also review the past maintenance effectiveness by presenting the amount of effective tamping on the network. For this purpose, the MSI of tamping should be summarized by calculating the total amount and percentages of different MSIs on the network. Fig. 14 demonstrates how the tamping history of the MSI assessment is used for network-level analysis. The time–location view shows when and where tamping has occurred, and the color represents the MSI of the tamping. The pie chart shows a summary of the proportion of different MSIs on the observed section. If the MSI Categories 3 or 4 are overrepresented in the network summary pie chart, the asset manager can examine the areas where tamping has not been effective and further investigate those areas. This principle can be used for the network level, but for readability, the example contains only a two-kilometer long section of track. The information from the histograms and MSI summaries can be used to guide asset management in deciding when the next major renewal should be carried out, instead of performing routine maintenance. If the TGDR histogram shows a major part of the track section having a high TGDR, and MSI summaries imply that routine maintenance has had little effect on retaining sufficient track geometry, the asset manager should start preparing for track renewals.

The network-level analysis also requires assessments of future tamping needs. These are best presented by a bar chart depicting the

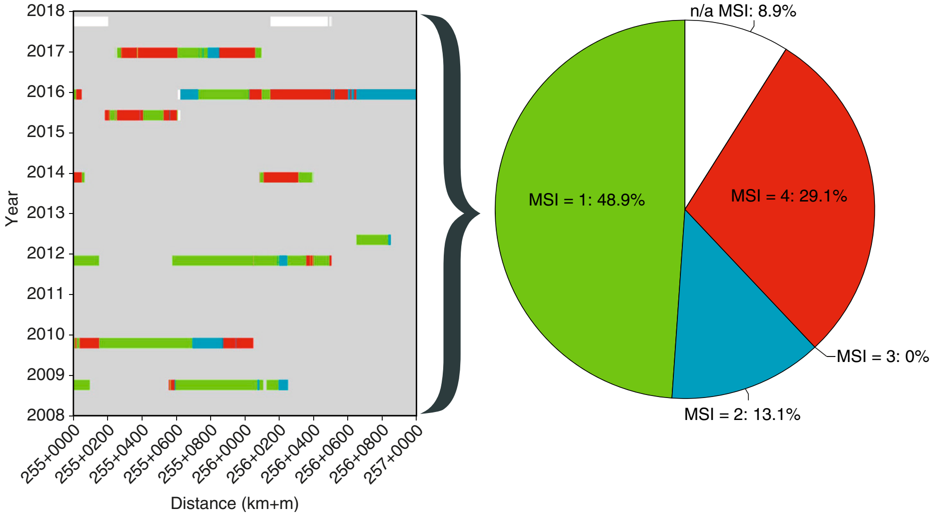


**Fig. 14.** Illustrations of MSI for a two kilometer section of track.

number of meters of track to be tamped in future years (Fig. 15). From illustrations like the one presented by Fig. 15, the asset manager can assess the need for maintenance funding and resources for the years to come. For example, if maintenance contracts are to be tendered, the asset manager can inform bidders about the expected amount of tamping required so that the amount of maintenance work and the number of tamping machines can be estimated. Furthermore, if the asset manager wants to represent the uncertainty in the estimations of future tamping needs, uncertainty can be calculated using, for example, a Monte Carlo simulation, as the linear models and PIs contain the necessary input data for simulating future observations; however the practical application for this is left for future research.

By combining the network level illustrations, the asset manager can answer the following questions:

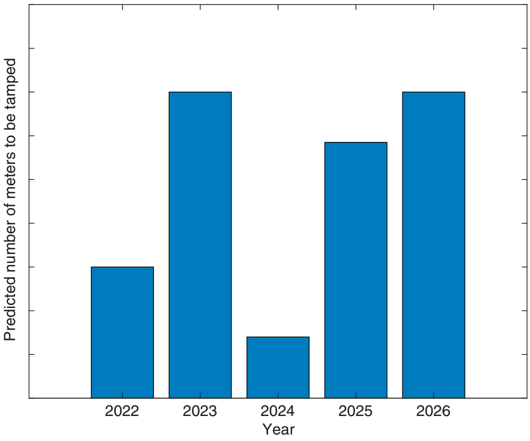


**Fig. 15.** Demonstration of an illustration of future tamping needs.

- What are the mean and extreme TGDR values on the network?
- How has the TGDR evolved on a network level, and how do different networks compare?
- Has tamping been effective on the network?
- What is the expected amount of tamping on the network for the years to come?

## Conclusion

This paper presented visualization techniques that help bring track geometry deterioration modeling from research into practice. In everyday asset management, the problem has been that track geometry deterioration modeling results are generally too complex and difficult to handle in daily operations. Therefore, this paper provided suitable visualization techniques, with which track geometry deterioration modeling results can be utilized by practitioners in asset management. This paper also demonstrated how modeling based on cross section data can be used for examining the track geometry behavior of longer sections of track and even for the whole rail network.

This study adapted the principles of track geometry deterioration modeling from the work of Neuhold et al. (2020), by slightly altering some aspects to better suit the data measured in Finland. The modeling approach was a robust linear model of a rolling 200 m LL SD. Future observation prediction accuracy was estimated using the PI, and past maintenance success was measured using the MSI. The proposed modeling methods are best suited for the LL SD, as it is generally observed to exhibit linear deterioration behavior. Other indices could work just as well, providing they follow linear deterioration behavior.

The main innovation of this study, i.e., how to connect suitable track geometry deterioration visualization techniques to different practical situations in asset management, is summarized in Fig. 16. The visualizations provide the following benefits:

- The cross-section level analysis helps to analyze isolated defects, their history and future development.
- The track section-level analysis provides a way to analyze longer sections of track to identify problematic areas and explore maintenance history, along with future maintenance timing.

J. Transp. Eng., Part A: Systems, 2022, 148(2): 04021106

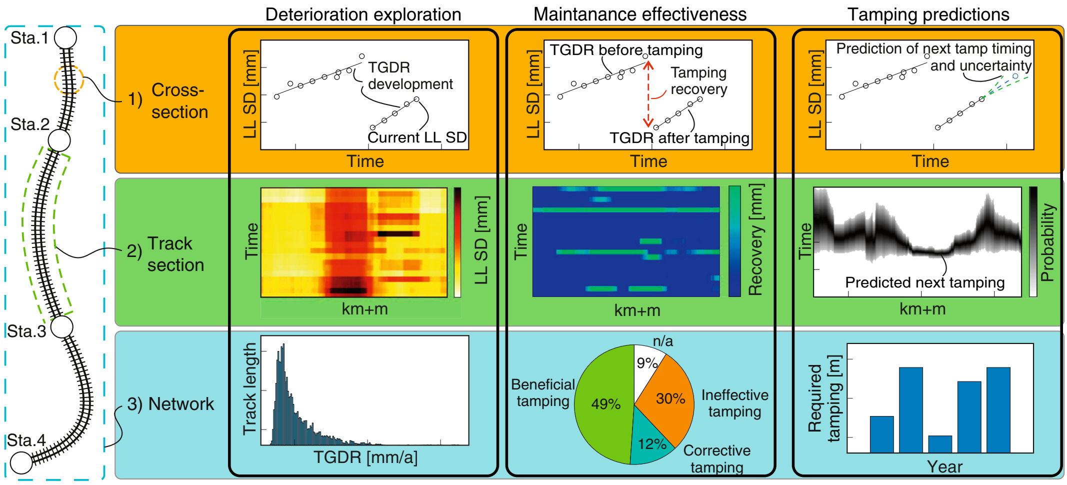


**Fig. 16.** Visualizations generated from track geometry deterioration modeling for asset management.

- The network level analysis summarizes the condition, past maintenance effectiveness, and future maintenance needs of complete track sections or even rail networks, into simple illustrations that help to make strategic decisions and allocate resources.

Lastly, several needs for future research were identified:

- A reliable method for detecting tamping areas in the track geometry history without the use of tamping records should be established to enhance the accuracy of linear regression modeling.
- The initial settlements after tamping should be further investigated to determine their duration in different circumstances and effect on relative and absolute track geometry.
- Region-specific research on TGDRs and suitable limit values for them should be conducted to provide comparable TGDRs from different environments.
- Uncertainty measures for the estimated amount of required future tamping should be defined.

## Data Availability Statement

Some or all data, models, or code generated or used during the study are proprietary or confidential in nature and may only be provided with restrictions.

## Acknowledgments

The authors want to thank the Tampere University Foundation Sr and the Finnish Transport Infrastructure Agency (Väylävirasto) for their valued co-operation.

## References

Agresti, A. 2015. *Foundations of linear and generalized linear models*. 1st ed., 95–99. New York: Wiley.

Andrade, A. R., and P. F. Teixeira. 2011. “Uncertainty in rail-track geometry degradation: Lisbon-Oporto line case study.” *J. Transp. Eng.* 137 (3): 193–200. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000206.

Audley, M., and J. D. Andrews. 2013. “The effects of tamping on railway track geometry degradation.” *Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit* 227 (4): 376–391. https://doi.org/10.1177/0954409713480439.

Caetano, L. F., and P. F. Teixeira. 2015. “Optimisation model to schedule railway track renewal operations: A life-cycle cost approach.” *Struct. Infrastruct. Eng.* 11 (11): 1524–1536. https://doi.org/10.1080/15732479.2014.982133.

CEN (European Committee for Standardization). 2014. *Railway applications. Track. Track geometry quality. Part 6: Characterisation of track geometry quality.* Eurocode. EN13848-6. Brussels, Belgium: CEN.

Dahlberg, T. 2001. “Some railroad settlement models—A critical review.” *Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit* 215 (4): 289–300. https://doi.org/10.1243/0954409011531585.

Famurewa, S. M., U. Juntti, A. Nissen, and U. Kumar. 2016. “Augmented utilisation of possession time: Analysis for track geometry maintenance.” *Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit* 230 (4): 1118–1130. https://doi.org/10.1177/0954409715583890.

Higgins, C., and X. Liu. 2018. “Modeling of track geometry degradation and decisions on safety and maintenance: A literature review and possible future research directions.” *Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit* 232 (5): 1385–1397. https://doi.org/10.1177/0954409717721870.

Khajehei, H., A. Ahmadi, I. Soleimanmeigouni, and A. Nissen. 2019. “Allocation of effective maintenance limit for railway track geometry.” *Struct. Infrastruct. Eng.* 15 (12): 1597–1612. https://doi.org/10.1080/15732479.2019.1629464.

Lee, J. S., I. Y. Choi, I. K. Kim, and S. H. Hwang. 2018. “Tamping and renewal optimization of ballasted track using track measurement data and genetic algorithm.” *J. Transp. Eng. Part A Syst.* 144 (3): 04017081. https://doi.org/10.1061/JTEPBS.0000120.

Li, S., R. Liu, C. Li, and F. Wang. 2019. “Recovery measure of tamping on different track geometry irregularity indexes.” *J. Transp. Eng. Part A Syst.* 145 (11): 05019005. https://doi.org/10.1061/JTEPBS.0000274.

Lichtberger, B. 2011. *Track compendium*. 2nd ed., 393–396. Hamburg, Germany: DVV Media Group GmbH I Eurailpress.

Neuhold, J., I. Vidovic, and S. Marschnig. 2020. “Preparing track geometry data for automated maintenance planning.” *J. Transp. Eng. Part A Syst.* 146 (5): 04020032. https://doi.org/10.1061/JTEPBS.0000349.

Nielsen, J. C., E. G. Berggren, A. Hammar, F. Jansson, and R. Bolmsvik. 2020. “Degradation of railway track geometry—Correlation between track stiffness gradient and differential settlement.” *Proc. Inst. Mech.*

*Eng., Part F: J. Rail Rapid Transit* 234 (1): 108–119. https://doi.org/10.1177/0954409718819581.

Quiroga, L. M., and E. Schnieder. 2012. “Monte Carlo simulation of railway track geometry deterioration and restoration.” *Proc. Inst. Mech. Eng., Part O: J. Risk Reliab.* 226 (3): 274–282. https://doi.org/10.1177/1748006X11418422.

Shenton, M. J. 1985. “Ballast deformation and track deterioration.” In *Track technology*, 253–265. London: Thomas Telford. https://doi.org/10.1680/tt.02289.0026.

Soleimanmeigouni, I., A. Ahmadi, and U. Kumar. 2018. “Track geometry degradation and maintenance modeling: A review.” *Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit* 232 (1): 73–102. https://doi.org/10.1177/0954409716657849.

Soleimanmeigouni, I., A. Ahmadi, A. Nissen, and X. Xiao. 2020. “Prediction of railway track geometry defects: A case study.” *Struct. Infrastruct. Eng.* 16 (7): 987–1001. https://doi.org/10.1080/15732479.2019.1679193.

Tanaka, H., S. Yamamoto, T. Oshima, and M. Miwa. 2018. “Methods for detecting and predicting localized rapid deterioration of track irregularity based on data measured with high frequency.” *Q. Rep. RTRI* 59 (3): 169–175. https://doi.org/10.2219/rtriqr.59.3_169.

UIC (Union of Railways). 2008. *Best practice guide for optimum track geometry durability. Railway technical publications*. Paris, France: UIC.

Vale, C., and R. Calçada. 2014. “A dynamic vehicle-track interaction model for predicting the track degradation process.” *J. Infrastruct. Syst.* 20 (3): 04014016. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000190.

# PUBLICATION
# II

## Investigating Root Causes of Railway Track Geometry Deterioration – A Data Mining Approach

Sauni, M., Luomala, H., Kolisoja, P. and Turunen, E.

Frontiers in Built Environment, Vol. 6: 122.
10.3389/fbuil.2020.00122

frontiers
in Built Environment

METHODS
published: 03 August 2020
doi: 10.3389/fbuil.2020.00122

# Investigating Root Causes of Railway Track Geometry Deterioration – A Data Mining Approach

*Mikko Sauni*[1]*, *Heikki Luomala*[1], *Pauli Kolisoja*[1] *and Esko Turunen*[2]

[1] *Built Environment, Tampere University, Tampere, Finland,* [2] *Information Technology and Communication Sciences, Tampere University, Tampere, Finland*

Railway track geometry deterioration indicates degradation in the underlying track structures. Monitoring and predicting this behavior are important as is investigating the root causes contributing to the deterioration. Without knowing the causes, assigned remediation might not result in a long-lasting correction. However, there is little research regarding the pragmatic aspects of investigating the root causes of track geometry deterioration utilizing real-world data sources. For this purpose, a new method was explored. After reviewing methodologies, the chosen approach was an association rule data mining method: General Unary Hypotheses Automaton (GUHA). The initial data used in data mining comprise data from asset management and multiple measurement systems, including a track geometry measurement vehicle, a track stiffness measurement device, ground penetrating radar, and lidar. The results of the GUHA data mining are hypotheses based on the initial data and can be used to indicate the most common and uncommon types of structures regarding their track geometry deterioration behavior and the attributes governing the behavior of a certain structure type. Therefore, the GUHA method was found to be a suitable method for investigating the root causes of track geometry deterioration from comprehensive railway track structure data.

**Keywords: association rules, condition monitoring, data mining, railway track, track geometry deterioration**

***Edited by:***
*Sakdirat Kaewunruen,*
*University of Birmingham,*
*United Kingdom*

***Reviewed by:***
*Ruilin You,*
*China Academy of Railway Sciences,*
*China*
*Mohd Haniff Bin Osman,*
*Universiti Kebangsaan Malaysia,*
*Malaysia*

****Correspondence:***
*Mikko Sauni*
*mikko.sauni@tuni.fi*

***Specialty section:***
*This article was submitted to*
*Transportation and Transit Systems,*
*a section of the journal*
*Frontiers in Built Environment*

***Received:*** *26 May 2020*
***Accepted:*** *07 July 2020*
***Published:*** *03 August 2020*

***Citation:***
*Sauni M, Luomala H, Kolisoja P*
*and Turunen E (2020) Investigating*
*Root Causes of Railway Track*
*Geometry Deterioration – A Data*
*Mining Approach.*
*Front. Built Environ. 6:122.*
*doi: 10.3389/fbuil.2020.00122*

## INTRODUCTION

Railway track structures endure harsh conditions and countless damaging loading cycles in their life cycle. During this life cycle, which usually lasts many decades, the structures degrade and require intermediate maintenance. However, the need for maintenance is not generally homogeneous along the length of a track section: Some areas require much more frequent maintenance than others. If the heterogeneous nature of degradation is not accounted for, dangerous conditions regarding train safety can occur. Furthermore, if the uniform maintenance for the whole track section is assigned according to the needs of the weakest parts of the track, plenty of unnecessary maintenance will be conducted, and money will be wasted. Therefore, the condition of the whole track section needs to be monitored.

The condition monitoring of track structures is widely conducted using track geometry measurement vehicles that measure deviations of track geometry using onboard measurement systems (Esveld, 2001). The deviations indicated by the measurement systems indicate wear or movement in the track structures. These measurement systems continue to be developed, and new technologies are applied, for example, conducting measurements from in-service vehicles (Weston et al., 2015).

The track geometry measurements require detailed analyses to ensure train safety, and they are traditionally done by comparing the measurement results to limit values set by the track owner. Many different techniques can be used in this type of results analysis as described by Berawi et al. (2010).

Analyzing the condition of track geometry using more sophisticated methods has been a popular branch of science as is evident from the number of different approaches for track deterioration modeling (Higgins and Liu, 2018). Especially, track geometry deterioration modeling has been popular. Track geometry deterioration is the process of uneven settling of track structures, which is observed by obtaining increasing deviations in track geometry, when new measurements are conducted and time progresses. If this process is modeled with great detail, either with a deterministic or a stochastic model, the required maintenance can be planned in advance, which leads to better use of the track availability and reduced maintenance costs.

Track geometry deterioration modeling is a worthwhile exercise as it has been proven to reduce costs in asset management (Andrews et al., 2014). However, deterioration modeling is solving only half the problem. Another important aspect to consider is investigating the root causes of track geometry deterioration. These root causes are here defined as the track structure features associated with increased track geometry deterioration rates, for example, insufficient drainage or subgrade deformation issues. Fixing these types of problems for the long term might require maintenance activities that are different from routine maintenance.

The most common maintenance activity for correcting track geometry deviations is tamping. Tamping is the process of lifting the rails and ties while compacting the ballast under the ties being lifted. Tamping can level the track geometry to provide a smooth running surface for trains. However, the effects of tamping are not permanent (Audley and Andrews, 2013). Furthermore, tamping does not increase the resilience of structures *per se*, but only provides temporary correction of geometry. Deteriorated or defective track structures continue to cause the track geometry to rapidly deteriorate to the state before tamping. Therefore, to attain a more lasting effect, the root causes for track geometry deterioration must be investigated to assign suitable remediation.

This aspect of investigating root causes for track geometry deterioration has been researched far less than track geometry deterioration modeling. Guler et al. (2011) used neural networks to predict track geometry deterioration based on certain track asset data. Sadeghi and Askarinejad (2009, 2012) have provided stochastic approaches to analyzing the effects of track structure conditions and track components to track geometry.

Although these studies have modeled the effects of different components and conditions, they do not strictly assess the root causes of track geometry deviations. For example, the severity of some features is assessed in these studies, but the commonness of a problem type is not. To advance the investigation of the root causes of track geometry deterioration, new methods have to be tested and applied. For this purpose, a method is explored: first, by searching a promising method by type, and second, by testing the chosen method using actual railway track structure data.

Choosing a method for investigating the root causes of track geometry deterioration can be taken in steps. First, it must be decided whether to create a deterministic model or use a stochastic approach. Using a deterministic model requires many experimental values and knowledge of the chain of events leading to deteriorated track geometry. Although many track settlement models are available (Dahlberg, 2001), their use for this purpose may not be suited as these models rely greatly on detailed descriptions of different loading and support conditions. This information is practically impossible to provide for all the different types of structures on a track section.

Stochastic models, on the other hand, can utilize already available data, and inarguably, there is a great volume of data recorded from track structures that can be utilized. This data, in the case of Finland, includes the track geometry measurement history; ground penetrating radar (GPR) measurements that can provide a continuous thickness and moisture index for different structure layers; laser scanning (lidar) results to indicate embankment shape, from which drainage depth can be assessed; track asset data, such as bridges, turnouts, and culverts; and continuous track deflection measurements conducted as demonstrated by Luomala et al. (2017).

Therefore, the next step should be to select one stochastic approach, from which there are many to choose. Considering the complexity of the multivariate heterogeneous initial data, the search should be pointed to data mining methods that can digest this type of data.

Data mining can be understood in many ways and terms. Terms, such as machine learning and deep learning, are associated with the subject and are sometimes used interchangeably. Even though there is no single conclusive definition of data mining, one well-established way to define it is to use the terminology provided by Fayyad et al. (1996). In this terminology, data mining is a step in a larger process that is knowledge discovery from data (KDD). KDD begins with raw data, and after many steps in preprocessing the data and applying data mining methods and expert judgment, knowledge can be retrieved as the result. In this process, data mining is the step in which data analysis and discovery algorithms are applied to produce patterns or models from the data (Fayyad et al., 1996).

Data mining is in itself a whole branch of science, from which there are many methods to choose. As previously mentioned, the terminology in the field is not irrefutable, but some generalizations can be made. Data mining can be divided into two categories with different primary goals: predictive or descriptive methods (Fayyad et al., 1996). The predictive or supervised methods, in other terminology (Tsui et al., 2006), focus on learning past behavior and predicting future observations based on a given input. Descriptive or unsupervised methods, in other terminology (Tsui et al., 2006), find patterns or relationships within the provided data, thus giving new insight about the data that could not be observed with human effort. Most methods do not belong to one category absolutely but generally exhibit stronger ties to one than the other (Fayyad et al., 1996).

Of these two methodologies, descriptive data mining is the more fitting choice because finding root causes of track geometry deterioration is closely related to finding novel patterns and relationships from data and presenting them to the end user. Descriptive data mining methods can be classified to include clustering, summarization, association rules, and sequence discovery (Dunham, 2003) of which clustering and association rules provide the best descriptions of the relationships between different data sources, whereas summarization and sequence discovery are more useful in cases such as text mining or customer purchase tracking, respectively.

Clustering is organizing the data into groups that represent data points that are more similar to data inside the cluster than outside it (Jain et al., 1999). Association rules provide insight on which data sources are most associated with other data sources with a specified confidence, often using Boolean logic (Agrawal et al., 1993). Of these two tasks, association rules better fit the purpose of this research.

To enhance current practices regarding track geometry deterioration analysis, the ability to investigate the root causes of track geometry deterioration using the association rule data mining algorithm General Unary Hypotheses Automaton (GUHA) was tested. The choice of the method was based on the reviewed methodologies and tasks. GUHA provides a way to assess the relationship of different input data attributes. In practical terms, using GUHA, associations between available railway track structure data sources and developments in track geometry can be investigated.

# MATERIALS AND METHODS

## Initial Data

The initial data used in the data mining presented in section "Results from Applying GUHA to Railway Track Structure Data" concern the Luumäki–Imatra track section located in Eastern Finland. The track section was initially built in the 1960s and was renewed at the beginning of the 21st century. The track section is a 65-km-long electrified single-track line, which has both passenger and cargo traffic. A major renewal is being planned on the track section in question because faster and heavier trains are required to increase the line's efficiency. The condition of the track section varies: Some sections of the track exhibit problematic structures, whereas others have required little maintenance during their life cycle.

The initial data available from the structures of this track section were conformed into a single matrix (CSV spreadsheet), in which a row of data depicts a 1-m-long section of track that is described by the columns representing the features of the track structure. The initial data matrix contained 65,142 rows and 25 columns. Of the 25 columns, 24 contained attributes used in data mining, and one column contained location information in the form of track meters. This was used only for locating interesting occurrences, not for data mining. **Figure 1** presents a snapshot of the initial data, and **Table 1** elaborates the attributes of the data.

The initial data were essentially either ratio or nominal data depending on the data origin. Ratio data, in this context, refers to data having a true zero, order, and quantifiable differences between data points. Nominal data, in this context, refers to categorical or binary data in which no ordering, direction, or distances for the data points are present.

The attribute for track geometry deterioration rate is further elaborated in section "Track Geometry Deterioration Rate". Track deflection was measured using a continuous track deflection measurement device presented by Luomala et al. (2017). Two attributes were created from the track deflection measurements: deflection level (mean) and variations (variance) in deflection. Furthermore, track deflection measurements provided geometry cant data, which were used to identify track geometry elements such as curves and straights.

GPR measurements provided the structural layer moisture indices and layer boundaries, using which layer thicknesses were calculated. The structural layer thicknesses were calculated for ballast, subballast, and embankment. Furthermore, an attribute for the whole structure thickness, a combination of the aforementioned, was provided. GPR measurements also revealed bedrock depths in places where the bedrock level was shallow.

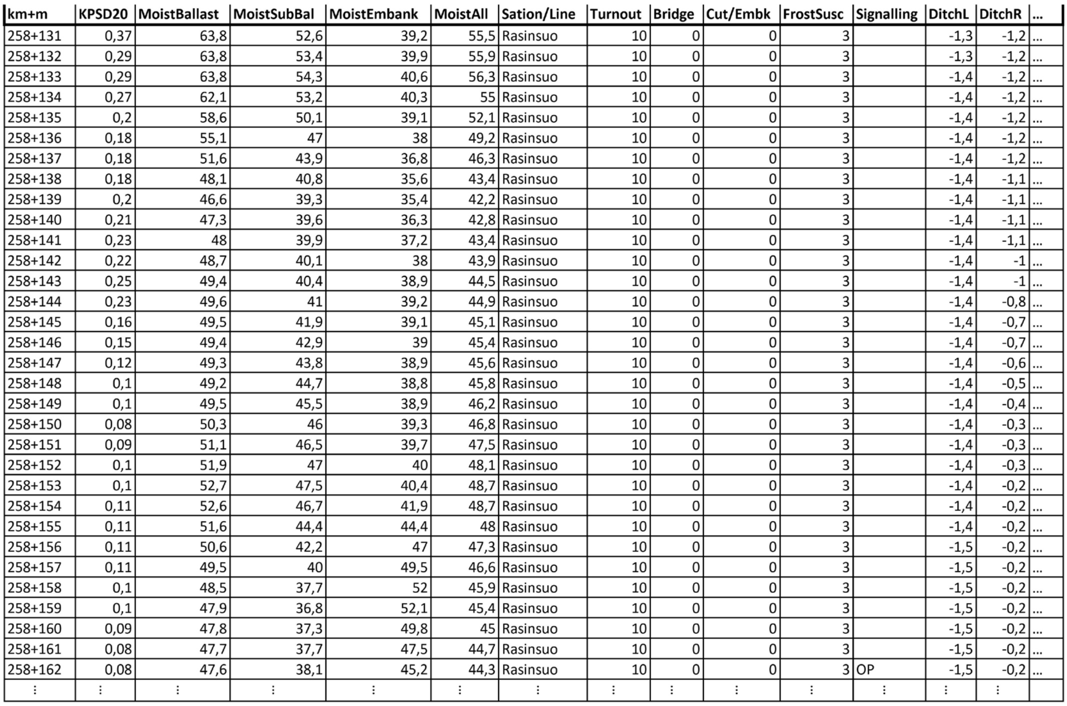
As a peculiarity, Finnish track structures are relatively thick compared with structures in warmer regions. The lowest allowable new track structure thickness using frost-resistant materials varies between 2.0 and 2.6 m, depending on the region. If the required track structure thickness is not met or if frost heave problems are observed on old track sections, frost insulation boards can be installed in the track substructure to reduce frost penetration. These frost insulation boards are extruded polystyrene boards that can withstand high pressure. Before the 2000s, some expanded polystyrene (EPS) boards were installed in track structures, but these did not endure well, and the use of EPS boards in track structures has since been banned.

As presented in **Table 1**, ditch depth was calculated from the laser scanning point clouds. Soil maps and historical data were used to assess the frost susceptibility of the subgrade. Asset data included binary and categorical attributes for frost insulation boards, stations, level crossings, bridges, culverts, turnouts, cuttings, and wayside signaling equipment. Some of the asset data were retrieved from the railway asset management data warehouse, and some of the data were created using the video feed of the track section combined with the GPR interpretations and laser point clouds. Accordingly, track assets could be accurately located.

The used initial data exhibited missing values. However, due to the GUHA method's ability to handle them and their small quantity, the missing values were left in the data. Some missing data were intentionally left blank and was handled in the software as an attribute category. For example, an empty value for a bridge implies the non-existence of a bridge. The actual missing values included ballast thickness on bridges without a ballast layer and ballast moisture in some turnouts where GPR measurements were distorted by the frog.

## Track Geometry Deterioration Rate

The process for calculating the track geometry deterioration rate is not unambiguously defined throughout literature. Therefore,



| km+m | KPSD20 | MoistBallast | MoistSubBal | MoistEmbank | MoistAll | Sation/Line | Turnout | Bridge | Cut/Embk | FrostSusc | Signalling | DitchL | DitchR | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 258+131 | 0,37 | 63,8 | 52,6 | 39,2 | 55,5 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,3 | -1,2 | ... |
| 258+132 | 0,29 | 63,8 | 53,4 | 39,9 | 55,9 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,3 | -1,2 | ... |
| 258+133 | 0,29 | 63,8 | 54,3 | 40,6 | 56,3 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,2 | ... |
| 258+134 | 0,27 | 62,1 | 53,2 | 40,3 | 55 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,2 | ... |
| 258+135 | 0,2 | 58,6 | 50,1 | 39,1 | 52,1 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,2 | ... |
| 258+136 | 0,18 | 55,1 | 47 | 38 | 49,2 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,2 | ... |
| 258+137 | 0,18 | 51,6 | 43,9 | 36,8 | 46,3 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,2 | ... |
| 258+138 | 0,18 | 48,1 | 40,8 | 35,6 | 43,4 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,1 | ... |
| 258+139 | 0,2 | 46,6 | 39,3 | 35,4 | 42,2 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,1 | ... |
| 258+140 | 0,21 | 47,3 | 39,6 | 36,3 | 42,8 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,1 | ... |
| 258+141 | 0,23 | 48 | 39,9 | 37,2 | 43,4 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1,1 | ... |
| 258+142 | 0,22 | 48,7 | 40,1 | 38 | 43,9 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1 | ... |
| 258+143 | 0,25 | 49,4 | 40,4 | 38,9 | 44,5 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -1 | ... |
| 258+144 | 0,23 | 49,6 | 41 | 39,2 | 44,9 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,8 | ... |
| 258+145 | 0,16 | 49,5 | 41,9 | 39,1 | 45,1 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,7 | ... |
| 258+146 | 0,15 | 49,4 | 42,9 | 39 | 45,4 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,7 | ... |
| 258+147 | 0,12 | 49,3 | 43,8 | 38,9 | 45,6 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,6 | ... |
| 258+148 | 0,1 | 49,2 | 44,7 | 38,8 | 45,8 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,5 | ... |
| 258+149 | 0,1 | 49,5 | 45,5 | 38,9 | 46,2 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,4 | ... |
| 258+150 | 0,08 | 50,3 | 46 | 39,3 | 46,8 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,3 | ... |
| 258+151 | 0,09 | 51,1 | 46,5 | 39,7 | 47,5 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,3 | ... |
| 258+152 | 0,1 | 51,9 | 47 | 40 | 48,1 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,3 | ... |
| 258+153 | 0,1 | 52,7 | 47,5 | 40,4 | 48,7 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,2 | ... |
| 258+154 | 0,11 | 52,6 | 46,7 | 41,9 | 48,7 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,2 | ... |
| 258+155 | 0,11 | 51,6 | 44,4 | 44,4 | 48 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,4 | -0,2 | ... |
| 258+156 | 0,11 | 50,6 | 42,2 | 47 | 47,3 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,5 | -0,2 | ... |
| 258+157 | 0,11 | 49,5 | 40 | 49,5 | 46,6 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,5 | -0,2 | ... |
| 258+158 | 0,1 | 48,5 | 37,7 | 52 | 45,9 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,5 | -0,2 | ... |
| 258+159 | 0,1 | 47,9 | 36,8 | 52,1 | 45,4 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,5 | -0,2 | ... |
| 258+160 | 0,09 | 47,8 | 37,3 | 49,8 | 45 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,5 | -0,2 | ... |
| 258+161 | 0,08 | 47,7 | 37,7 | 47,5 | 44,7 | Rasinsuo | 10 | 0 | 0 | 3 |  | -1,5 | -0,2 | ... |
| 258+162 | 0,08 | 47,6 | 38,1 | 45,2 | 44,3 | Rasinsuo | 10 | 0 | 0 | 3 | OP | -1,5 | -0,2 | ... |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |  |

**FIGURE 1 |** Snapshot of the initial data.

**TABLE 1 |** Attributes used in data mining.

| Data origin | Data attribute | Data type | Data preprocessing |
|---|---|---|---|
| Track geometry car | Track geometry deterioration rate | Ratio | Annual 20 m SD growth |
| Continuous track deflection measurement | Track deflection mean | Ratio | 20-m mean |
| Continuous track deflection measurement | Track deflection variance | Ratio | 20-m variance |
| Continuous track deflection measurement | Geometry elements straight and curve | Binary | Calculation of cant indicating a curve |
| GPR | Structural layer moisture indices | Ratio | Signal attenuation calculations |
| GPR | Structural layer thicknesses | Ratio | Signal rebound calculations |
| Continuous laser scanning point cloud | Ditch depth | Ratio | Minimum value from 4 to 8 m perpendicular to the track centerline in a 20-m distance |
| Soil maps | Subgrade frost susceptibility assessment | Categorical | Subjective classification |
| Photos and visual assessment of data | Foundation type | Categorical | Subjective classification |
| Asset data and visual assessment of data | Asset data | Categorical and binary | Subjective classification |
| Tamping records | Tamping history | Categorical | Subjective classification |

it is pertinent to fully elaborate how the calculations have been conducted, especially as the track geometry deterioration rate is used as the predominant measure of durability.

The track geometry measurement data were produced using a track recording vehicle, Plasser and Theurer EM 120 (Ttr1 51), which uses relative measurements from three bogies to determine track geometry deviations. The measurement data contained biannual measurements from 2008 to 2018. Longitudinal geometry deviations were used in calculating the deterioration rate because the longitudinal geometry is mostly affected by the movements in the structures below the track rather than only by the rails or sleepers themselves.

Different chord lengths and parameters were tested calculating the track geometry deterioration rate. A 20-m running standard deviation (SD) calculated from longitudinal deviations (LD) was chosen as it best described the original longitudinal geometry deviation signal. The SD values obtained from the consecutive measurements were used to calculate the annual increase or decrease in track geometry deterioration. The mean of the increased annual values was used to describe the track geometry deterioration rate. If the SD values significantly decreased from 1 year to another, the reduction was ignored in the track geometry deterioration rate because a large reduction in the deviation implied tamping or other maintenance and repair actions. The track geometry deterioration rate was calculated for each point in the track section in 1-m intervals to be in conformity with the other initial data.

The average deterioration rate for the Luumäki–Imatra track section was 0.103 mm/a. Track geometry deterioration rate was lower than average on 70% of the track section, meaning that problematic areas were not as common as non-problematic areas but exhibited much higher deterioration rates than the non-problematic areas. This result was expected because problematic areas are not generally long sections of the track.

**Figure 2** presents an example of the track geometry deterioration rate of two cross-sections in which the $y$-axis represents 20 m SD values of LD. The deterioration behavior of the two cross-sections is very different. The cross-section at track kilometer 260 + 390 is at the edge of a section having frost insulation boards. The cross-section at track kilometer 260 + 360 is approximately 20 m away from the section having frost insulation boards.

The track geometry deterioration rate for cross-section 260 + 390 was 0.35 mm/a, whereas the corresponding value for cross-section 260 + 360 was 0.05 mm/a. Tamping can be observed to have taken place before the 2012 and 2016 winter measurements. Surprisingly, the 2012 tamping has increased deviations at cross-section 260 + 360, which might be due to uneven ballast settlement after tamping. However, the effect is nearly negligible because the deviations at cross-section 260 + 360 do not tend to grow, and the 2016 tamping has restored the deviations to their original level. In the spring of 2011, the track geometry was measured both in April and May. These measurements produced different results at cross-section 260 + 390. Winter of 2010–2011 was especially cold in Finland, and the measurements indicate the time before frost thaw and after frost thaw as deviations have significantly increased between the two measurements.

The calculated track geometry deterioration rate was visualized and compared with other available data. The deterioration indicates the condition of a track structure. Known problem areas, such as bridge transitions (Li and Davis, 2005) and stiffness variations (Dahlberg, 2010), could be detected based on the deterioration rate. In addition, tamping and frost heave problems could be observed from the track geometry history as large reductions or fluctuation in the deviations. The track geometry deterioration rate was generally used as the succedent attribute in GUHA data mining.

## GUHA Method

The GUHA method was initially developed in the 1960s and 1970s, and its background was elaborated by Hájek and Havránek (1978). An up-to-date and comprehensive presentation of the method can be found in Jan Rauch's Observational Calculi and Association Rules (2013). The GUHA method is considered a descriptive data mining method. Hence, it is not used to make deductions or predictions, but to describe and present input data in new ways to users by producing hypotheses.

The GUHA method is based on logic formalism: the statements about data are either true (data support a statement) or false (data do not support the statement). The user provides general questions about the data. Typical data can produce millions of statements, among which only a few are true and interesting to the user. True statements, referred to as hypotheses, are considered to be answers to the user's questions.

Data mining was conducted using the LISp-Miner program, an application of the GUHA method (Rauch, 2013). The practical aspects of using LISp-Miner have been elaborated by Berka (2016). The GUHA method and its application, LISp-Miner, have considerably evolved since their discovery and are still being further developed (Novák et al., 2008; Hájek et al., 2010; Piché et al., 2014).

**Figure 3** presents the generic process for using the GUHA method and the LISp-Miner program. This process begins with collecting and formatting data into an initial data matrix that is suitable for data mining. In the initial data, rows contain observations, and columns contain attributes (also called predicates), meaning the properties the observations have. In GUHA data mining, the key is to set relevant questions, called analytical questions, related to the data. These questions can be translated into the GUHA language. Then, GUHA data mining produces various hypotheses based on the input data. The hypotheses are automatically generated according to boundary conditions that are selected by the user. The hypotheses can vary from trivial to interesting in a single data mining task. The user can choose the meaningful ones and further explore them by assessing their contingency tables and associated predicates. After analyzing the results, the user can subjectively translate the numeric results into comprehensible human language.

The boundary conditions of the predicates assigned by the user include antecedents, succedents, conditions, and quantifiers, which adjust the preconditions and consequences of data mining. Adjusting these boundary conditions influences the types and number of results produced. The user should intend to achieve a limited number of results to reveal the strongest correlations within the data.

Antecedents, succedents, and conditions are attributes from the initial data. Any attribute can be set as an antecedent, succedent, or a condition, and any number or combination of attributes can be chosen. Furthermore, the assessment of attribute categories can be adjusted by choosing the coefficient

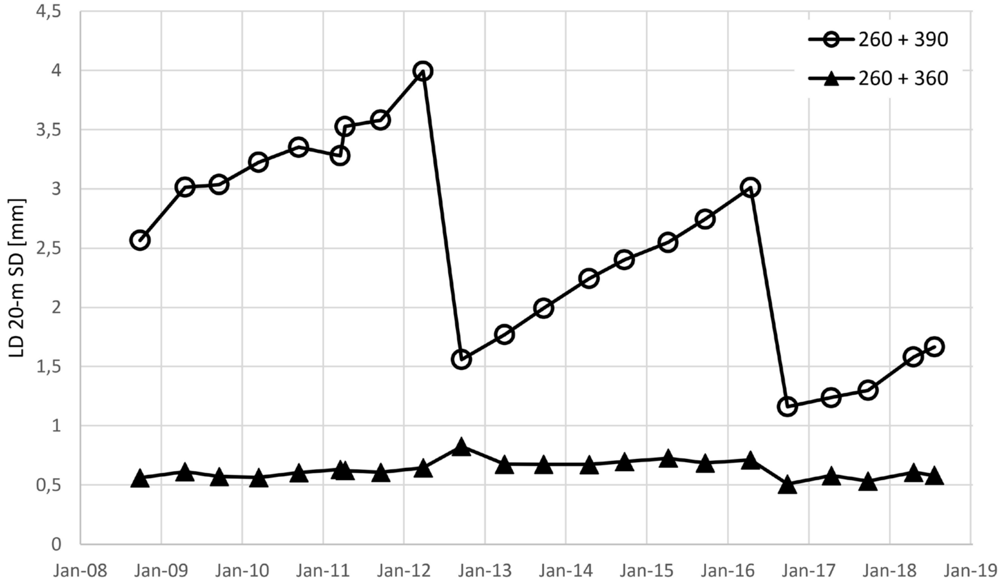
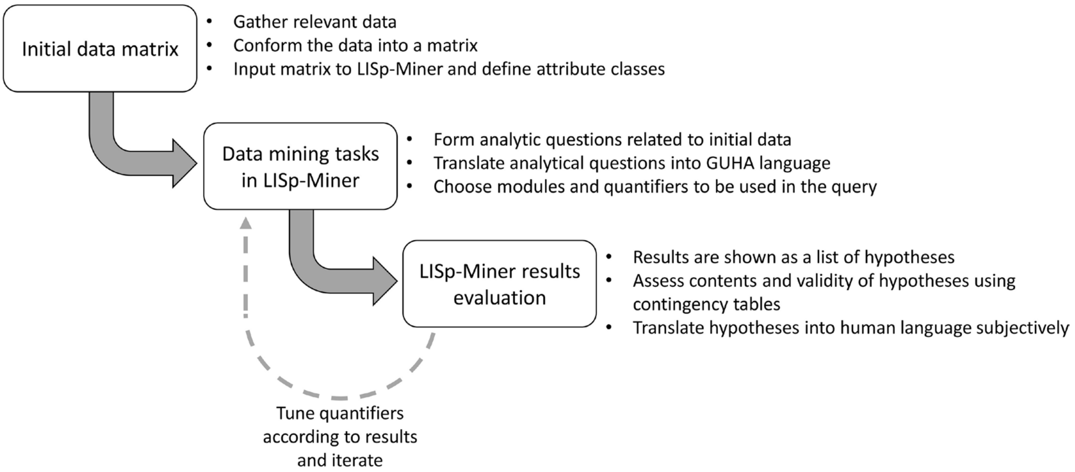


**FIGURE 2 |** LD 20-m SD of two cross-sections.



- Gather relevant data
- Conform the data into a matrix
- Input matrix to LISp-Miner and define attribute classes

- Form analytic questions related to initial data
- Translate analytical questions into GUHA language
- Choose modules and quantifiers to be used in the query

- Results are shown as a list of hypotheses
- Assess contents and validity of hypotheses using contingency tables
- Translate hypotheses into human language subjectively

**FIGURE 3 |** Principle of the data mining process using LISp-Miner.

type and length. This process adjusts how many attribute categories are regarded in one category and how the combined categories are comprised.

The results (hypotheses) in the LISp-Miner program are presented to the user as contingency tables (**Table 2**). Based on the contents of the contingency tables and hypotheses, the user can assess the meaning and importance of the hypotheses and also subjectively examine the initial data and determine the rows from the data that support a hypothesis and those that oppose a hypothesis.

In **Table 2**, $n$ is the number of initial data matrix rows regarded in a contingency table ($n = a + b + c + d$), when a is the number of objects satisfying both $\varphi$ and $\psi$; $b$ is the number of objects satisfying $\varphi$, but not $\psi$; $c$ is the number of objects not satisfying $\varphi$, but satisfying $\psi$; and $d$ is the number of objects not satisfying $\varphi$ nor $\psi$ (Turunen, 2018).

Ten data mining modules have been implemented into the LISp-Miner software, two of which were applied in the investigation of root causes for track geometry deterioration. In the 4ft-Miner module, several quantifiers can be used to evaluate the contingency table antecedent's (φ) relationship to its succedents (ψ) when condition (γ) is satisfied. In the AC4ft-Miner (action miner) module, two contingency tables are assessed and compared when some attributes remain stable and others change (called flexible attributes) between the tables (Berka, 2016). Pairs of rules specific for each quantifier are available to test the contingency tables' data.

GUHA quantifiers have an intuitive meaning, for example, "often implies," "almost equivalent," and "above average." Association rules based on quantifiers founded implication (also called p-implication or PIM) and above-average dependence were applied in both modules used. Founded implication assesses the commonness of the relationship *p* between contingency table parameters *a* and *b*. This can be expressed by,

$$\frac{a}{a+b} \geq p \text{ and } a \geq Base \text{ (Rauch, 2013).} \quad (1)$$

By adjusting *p*, the user can choose to inquire hypotheses for which the antecedents and succedents are fulfilled in $0 < p \leq 1$ of cases (Rauch, 2013). For example, the query may involve asking in which cases the antecedent φ and succedent ψ are simultaneously fulfilled in more than 90% of cases. In other words, the association between φ and ψ is supported by the data if at least 90% of the cases in which φ is satisfied also ψ is satisfied.

The association rule based on the above-average quantifier tests how much more common succedent ψ is among the antecedents φ in relation to all the instances of ψ in the whole data set. This is defined more explicitly by

$$\frac{a}{a+b} \geq (1+p)\frac{a+c}{a+b+c+d} \text{ and } a \geq Base \text{ (Rauch, 2013).} \quad (2)$$

when $p > 0$. Now, by adjusting *p*, the user can choose how many more times above-average dependence must appear for the hypothesis to be accepted. For example, by choosing $p = 1$, the software will search the cases in which ψ appears two times more often in relation to φ than ψ appears in the whole data.

**TABLE 2 |** Contingency table satisfying condition γ (Berka, 2016; Turunen, 2018).

| γ | ψ | ¬ψ | Σ |
|---|---|---|---|
| φ | $a$ | $b$ | $a + b = r$ |
| ¬φ | $c$ | $d$ | $c + d = s$ |
| Σ | $a + c = k$ | $b + d = l$ | $n$ |

Frequencies related to quantifiers are also implemented into the modules. These regulate the *Base* value: the number of occurrences in different contingency table slots. For example, a quantifier for the contingency table parameter $a \geq Base = 1000$ can be given. Then, LISp-Miner will not present any hypotheses for which fewer than 1000 cases have fulfilled the antecedents, succedents, and conditions regardless of other chosen quantifiers.

Examples of how analytical questions are formed to GUHA questions and how hypotheses found by LISp-Miner procedures are interpreted into comprehensible language can be found in the next section.

## RESULTS FROM APPLYING GUHA TO RAILWAY TRACK STRUCTURE DATA

In this section, the application of the GUHA method to railway data is demonstrated by conducting three different exemplary GUHA data mining tasks. The demonstrations show how the software is used and the types of results that can be obtained. This section only presents the data mining queries and their results. The results' domain knowledge interpretations and the possible broader implications to railway domain applications are presented in the discussion.

In the demonstrations, analytical questions about the development of the track structure condition are formed and translated to GUHA language in LISp-Miner, and answers (hypotheses) to the questions are presented. The analytical questions were inquired using the data concerning the Luumäki–Imatra track section. The technical information concerning the queries and their results is composed into **Table 3**.

**Analytical Question 1:** What kind of track structure attributes are associated with a certain type of track geometry deterioration rate with more than four times above-average dependence?

The first query was conducted using 4ft-Miner module. *Base* parameter for contingency table parameter $a \geq 2000$ and quantifier over four times above-average dependence were applied. All attributes except for the track geometry deterioration rate could be chosen for antecedents, but the program was limited to choose 2–5 attributes. The only succedent was the track geometry deterioration rate, for which 1–4 sequential classes could be chosen by the program. No conditions were applied.

The query concerning analytical question 1 resulted in 112,059,584 verifications (contingency tables), of which 163 were in accordance with the preconditions (antecedents, succedents, conditions, and quantifiers). These hypotheses were displayed

**TABLE 3 |** Technical information of queries.

| Query | Module | Quantifier | *Base* quantifier | Verifications | Number of hypotheses |
|---|---|---|---|---|---|
| 1 | 4ft | AAD ≥ 4 | $a \geq 2000$ | 112,059,584 | 163 |
| 2 | 4ft | PIM ≥ 0.9 | $a \geq 5000$ | 111,967 | 50 |
| 3 | AC4ft | State before PIM ≥ 0.7<br>State after PIM ≤ 0.4 | State before $a \geq 1000$<br>State after $a \geq 300$ | 2070 | 40 |

**TABLE 4 |** Contingency table for analytical question 1 and 2 hypotheses.

| | Hypothesis 1 | | Hypothesis 2 | |
|---|---|---|---|---|
| | **Succedent** | **¬Succedent** | **Succedent** | **¬Succedent** |
| Antecedent | 2106 | 1338 | 5350 | 415 |
| ¬Antecedent | 5247 | 56,451 | 28,948 | 9,441 |

**TABLE 5 |** Attributes for analytical question 1 hypothesis.

| Antecedent | Class |
|---|---|
| Liner or station | Line |
| Culvert | No |
| Bridge | No |
| Substructure moisture index | >50 (%) |
| Frost insulation board | Yes |
| **Succedent** | **Class** |
| Track geometry deterioration rate | >0.20 mm/a |

to the user. One of the 163 hypotheses is presented below. Its contingency table is presented in **Table 4**, and attributes are presented in **Table 5**.

Hypothesis that is one answer to analytical question 1 (statement supported by the data): When the track section is located on a line section that contains no bridges or culverts, its substructure exhibits a high moisture index, and a frost insulation board is installed in the track structure, the highest class of track geometry deterioration rate is observed 4.4 times more often than on average.

No conditions were set for the analytical question 1 query, so the whole track section, composed of 65,142 (=2106 + 1338 + 5247 + 56,451) rows of data, is presented in the hypothesis and contingency table.

**Analytical Question 2:** What kind of track structure attributes have the highest correlation to some types of track geometry deterioration rate on a line section without track structure discontinuity or frost insulation boards?

Analytical question 2 query was also conducted using the 4ft-Miner module. *Base* parameter $a \geq 5000$ and founded quantifier PIM must be over 90% ($p \geq 0.9$) were used. Antecedents included all track structure attributes aside from the track geometry deterioration rate, discontinuity attributes, stations, and frost insulation boards. The succedent included the track geometry deterioration rate, from which the program could choose 1–4 sequential classes. Sections with signaling equipment, stations, culverts, bridges, level crossings, turnouts, and frost insulation boards were excluded using conditions.

Analytical question 2 query resulted in 111,967 verifications, of which 50 were in accordance with the preconditions. One of the 50 hypotheses is presented below. Its contingency table is presented in **Table 4**, and attributes are presented in **Table 6**.

Analytical question 2 hypothesis (statement supported by the data): A lower than average track geometry deterioration rate is observed on 93% of the track structures that are founded on an embankment, exhibit 300- to 500-mm-thick ballast layers, exhibit a low structure moisture index, are located on straights, and have low track deflection variance.

Because conditions were used to exclude certain types of track, only 44,154 (=5350 + 415 + 28,948 + 9441) rows are now presented in the contingency table, meaning that 20,988 rows contained discontinuities, stations, or frost insulation boards and were not included in the data mining task.

**Analytical Question 3:** If some track structure attributes are stable, how does a change in the attribute for frost insulation boards affect a certain type of track geometry deterioration rate on a line section without track structure discontinuities?

The third analytical question was conducted using the 4ft-Action Miner (AC4ft). *Base* parameter $a \geq 1000$ for the before state and $a \geq 300$ for the after state were used. Founded implication $p \geq 0.7$ for the before state and $p \leq 0.4$ for the after state were applied. Antecedents' stable part included all track structure attributes except for frost insulation boards, track geometry deterioration rate, stations, and discontinuities. Antecedent attribute part included frost insulation boards. The succedent stable part was the track geometry deterioration rate from which the program could choose 2–4 sequential classes. In the conditions, signaling equipment, stations, culverts, bridges, level crossings, and turnouts were excluded.

Analytical question 3 query resulted in 2070 verifications, which led to 40 results. One of the 40 results is presented below. Its two adjacent contingency tables are presented in **Table 7**, and attributes are presented in **Table 8**. There were 47,881 rows of data that met the conditions and were examined in the hypothesis.

Analytical question 3 hypothesis (statement supported by the data): When the track moisture index is very high and the number of tamping times is low, a high track geometry deterioration rate is observed on 79% of the structures where a frost insulation

**TABLE 6 |** Attributes for analytical question 2 hypothesis.

| Antecedent | Class |
|---|---|
| Foundation type | Embankment |
| Ballast thickness | 300–500 mm |
| Structure moisture index | 10–40 (%) |
| Straight or curve | Straight |
| Track deflection variance | <0.01 mm |
| **Succedent** | **Class** |
| Track geometry deterioration rate | <0.10 mm/a |
| **Condition** | **Class** |
| Signaling equipment | No |
| Straight or curve | Straight |
| Foundation type | Embankment |
| Culvert | No |
| Bridge | No |
| Level crossing | No |
| Turnout | No |
| Frost insulation board | No |

**TABLE 7 |** Contingency tables for analytical question 3 hypothesis.

| | Frost insulation board | | No frost insulation board | |
|---|---|---|---|---|
| | **Succedent** | **¬Succedent** | **Succedent** | **¬ Succedent** |
| Antecedent | 2014 | 533 | 371 | 2216 |
| ¬Antecedent | 4996 | 40,338 | 6639 | 38,655 |

**TABLE 8 |** Attributes for analytical question 3 hypothesis.

| **Antecedent** | **Class** |
|---|---|
| Structure moisture index | >50 (%) |
| Number of tampings | 1–2 |
| Frost insulation board | Flexible attribute |
| **Succedent** | **Class** |
| Track geometry deterioration rate | >0.14 mm/a |
| **Condition** | **Class** |
| Signaling equipment | No |
| Straight or curve | Straight |
| Foundation type | Embankment |
| Culvert | No |
| Bridge | No |
| Level crossing | No |
| Turnout | No |

board has been installed and on 14% of track sections where no frost insulation board has been installed.

# DISCUSSION

## Case Track Section Data Mining

The hypothesis for analytical question 1 presented the combination of parameters that were more commonly associated with high track geometry deterioration rates, meaning that the track section is abnormal as regards track geometry deterioration, and the hypothesis attributes should be investigated further.

The attributes of the hypothesis include common attributes, such as line sections instead of stations and the exclusion of bridges and culverts. These do not create a distinct attribute combination as the vast majority of the track section shares these attribute types. The other two antecedents are far more infrequent in the data: high substructure moisture index and frost insulation boards. However, these two attribute values are connected due to the GPR measurement technique. Frost insulation boards increase the GPR moisture index of the substructure layer because they cause the GPR signal to deflect and give high readings that would normally indicate the appearance of moisture. Therefore, it is reasonable to deduct that the frost insulation boards are playing a major role in this hypothesis. Based on this information, the areas located on line sections in which frost insulation has been installed should be further investigated. Such investigations have been reported in Sauni et al. (2020).

The analytical question 1 hypothesis has good confidence as more than 2 km of track support the statement, and about 1.3 km of track oppose it. If the same hypothesis were to be created for the rest of the track section, only around 5.2 km of track would support it, and more than 56 km oppose it. Considering these lengths, the behavior of the track section in accordance with the hypothesis antecedents is unusual to say the least.

The hypothesis for analytical question 2 demonstrated the highest correlation to a particular type of track geometry deterioration rate. The result implied that almost all cases (93%) of track sections in accordance with the antecedents exhibit only low track geometry deterioration rates. This correlation does not deviate from the average correlation (75%) of the rest of the track section as much as the correlations in hypothesis for analytical question 1. Nevertheless, this hypothesis showed that the correlation is particularly strong as more than 5 km of track satisfying the antecedents behaves almost uniformly.

The antecedents of the hypothesis for analytical question 2 exhibit properties traditionally associated with good structures such as low moisture and low deflection variance. The results are intuitive and demonstrate that the presumptions regarding the properties presented in the antecedents are justified. Furthermore, when all the hypotheses for analytical question 2 were examined, it was apparent that all hypotheses' succedents were related to low track geometry deterioration rates. This may be the result of opting out track discontinuities and frost insulation boards from the antecedents.

A difference could be observed between the types of hypotheses obtained from analytical questions 1 and 2. Analytical question 1 produced results concerning abnormal behavior of track structures, whereas analytical question 2 produced results concerning typical behavior.

The third analytical question provided a comparison of two populations that differed by one antecedent class: frost insulation boards. According to one produced hypothesis, the existence of a frost insulation board divides track sections consisting of track built on embankment without discontinuities. On these structures with frost insulation boards, high track geometry deterioration rates are observed on 79% of structures. When only the attribute for frost insulation boards is changed to *no frost insulation board*, the commonness of high track geometry is practically converse at 14%. This result highlights the major effect of frost insulation boards on the track geometry deterioration rate.

## Prospective of GUHA in Railway Track Structure Condition Monitoring

In this section, the use of the tested LISp-Miner GUHA data mining modules and quantifiers is discussed in a broader context regarding railway track structure condition monitoring.

Stochastic analysis of railway track structures inherently leads to handling heterogeneous data that originate from multiple sources. The requirement for an analysis method and software to handle this type of data is met using LISp-Miner, as text, numerals, binary, and categorical data can all be used as they are. Furthermore, missing data and outliers can be handled within the LISp-Miner software when creating attribute categories. Thus, the GUHA method and LISp-Miner software provide an adequate basis for track structure data analysis.

From the heterogeneous track structure data, the GUHA method could be used to ask questions related to correlations between variables and their combinations. Three different types of questions were asked, for which different module-quantifier combinations were used.

The **4ft-Miner module with the PIM quantifier** can be used to inquire about the most common types of attribute combinations. For the investigation of the causes of track geometry deterioration, these questions help in understanding the most common types of track structure behavior. This helps in identifying structures, i.e., the combination of attributes that generally exhibit only a certain type of behavior.

The **4ft-Miner module with the above-average quantifier** can be used practically for the contrary of purpose as 4ft-Miner with PIM. The above-average quantifier provides extraordinary correlations between variable combinations when compared with other variables' correlations. In the context of investigating the causes of track geometry deterioration, this approach can be used to detect abnormal behavior of some structure types. This information is of value in detecting the peculiar structure types that exhibit problematic behavior.

The **AC4ft-Miner module with the PIM quantifier** approach investigates the effects of changing one or some of the attribute classes in a hypothesis. In practice, this method can reveal which attributes have the dominant effect on a certain type of structure's behavior. This feature can be used to individually detect the attributes contributing to geometry deterioration rate.

The encountered limitations of the GUHA method were the dependence on initial data and the amount of effort required for result analysis. The dependence on initial data stems from the descriptive nature of the method. If the input data do not entail the features affecting the behavior of the structure, the method cannot produce results that exhibit such features. The initial data available for the case track section were vast. However, such data sets are not readily available for all track sections. To ensure reliable and interesting results, the method should be used only if extensive data are available.

The other encountered limitation was the difficulty to communicate the results to people not familiar with GUHA. The contingency tables and attributes can be subjectively translated into comprehensible language, which aids communication. However, some of the translated hypotheses can be difficult to fully comprehend as they might contain many variables and details. To counter the difficulties, visualizing the results should be further researched.

## CONCLUSION

Successful condition monitoring of track geometry requires not only measurements and maintenance responses to deviations but also investigations into the root causes for its deterioration. For the investigations, an approach with flexible data handling and good generalization ability is required. Thus, stochastic models were examined instead of deterministic models as the latter requires much too specific input information, which is not usually available in asset management.

From the stochastic models, an association rule data mining method, GUHA, was selected to be tested. The method is a descriptive data mining method, meaning that it describes the input data and presents it to the user in an informative way. The GUHA method is applied in software, LISp-Miner, which can handle multivariate heterogeneous data and produces hypotheses that are statements generated from the input data.

The use of the GUHA method was tested on actual track structure data from the Finnish state rail network. Three GUHA module-and-quantifier combinations were examined. The results from the data mining were used to generalize the types of domain information that can be investigated using the GUHA method. Three following applications for approaches were identified:

- 4ft-Miner and PIM quantifier identifies the structure types (attribute combinations) that correlate strongly to a certain track geometry deterioration rate.
- 4ft-Miner and above-average quantifier identifies the structure types that exhibit behavior, which differs from the typical behavior of structures.
- AC4ft-Miner module and PIM quantifier identify the structure attributes affecting the behavior of structures when changed.

Using the information obtained from these approaches, the causes of track geometry deterioration can be investigated from asset data. The method points out the structure types correlating to certain behavior and identifies the attributes governing the behavior. The main limitation of the method is the dependence to the input data. If a feature is not depicted in the initial data, it cannot be present in the results either. The GUHA method and LISp-Miner contain many more approaches in addition to the three tested ones. Exploring the applicability of these in the future would be valuable.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

MS was responsible for gathering and processing the initial data and conducting the data mining. HL and PK were responsible for organizing the research and participated in the analysis of the data mining results. ET supervised the data mining process. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was funded by the Finnish Transport Infrastructure Agency (Väylä) and the Tampere University Foundation sr.

## ACKNOWLEDGMENTS

The authors would like to thank our funder, the Finnish Transport Infrastructure Agency (Väylä), and the valued cooperation between Tampere University and Väylä.

## REFERENCES

Agrawal, R., Imielinski, T., and Swami, A. (1993). "Mining association rules between sets of items in large databases. SIGMOD '93," in *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, New York, NY: ACM, 207–2016.

Andrews, J. D., Prescott, D., and De Rozieres, F. (2014). A stochastic model for railway track asset management. *Reliab. Eng. Syst. Saf.* 130, 76–84. doi: 10.1016/j.ress.2014.04.021

Audley, M., and Andrews, J. D. (2013). The effects of tamping on railway track geometry degradation. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit.* 227, 376–391. doi: 10.1177/0954409713480439

Berawi, A. R. B., Delgado, R., Calçada, R., and Vale, C. (2010). Evaluating track geometrical quality through different methodologies. *IJTech* 1, 38–47. doi: 10.14716/ijtech.v1i1.35

Berka, P. (2016). Practical aspects of data mining using LISp-miner. *Comput. Inform.* 35, 528–554.

Dahlberg, T. (2001). Some railroad settlement models–a critical review. *Proc. IMechE. Part FJ Rail and Rapid Transit.* 215, 289–300. doi: 10.1243/0954409011531585

Dahlberg, T. (2010). Railway track stiffness variations – consequences and countermeasures. *Int. J. Civil Eng.* 8, 1–12.

Dunham, M. (2003). *Basic Data Mining Tasks. Data Mining Introductory and Advanced Topics*. New Jersey: Prentice Hall, 5–9.

Esveld, C. (2001). *Modern Railway Track*, 2nd Edn. Zaltbommel: MRT-Prod.

Fayyad, U., Piatetsky-Shapiro, G., and Padhraic, S. (1996). From data mining to knowledge discovery in databases. *AI Magazine* 17, 37–54.

Guler, H., Evren, G., and Jovanovic, S. (2011). Modelling railway track geometry deterioration. *Proc. Inst. Civ. Eng. Transp.* 162, 65–75. doi: 10.1680/tran.2011.164.2.65

Hájek, P., and Havránek, T. (1978). *Mechanizing Hypothesis Formation: Mathematical Foundations for a General Theory*, 1st Edn. Berlin: Springer.

Hájek, P., Holeňa, M., and Rauch, J. (2010). The GUHA method and its meaning for data mining. *J. Comput. Syst. Sci.* 76, 34–48. doi: 10.1016/j.jcss.2009.05.004

Higgins, C., and Liu, X. (2018). Modeling of track geometry degradation and decisions on safety and maintenance: a literature review and possible future research directions. *Proc. IMechE. Part FJ Rail Rapid Transit.* 232, 1385–1397. doi: 10.1177/0954409717721870

Jain, A., Murty, M., and Flynn, P. (1999). Data clustering: a review. *ACM Comput. Surv.* 31, 264–323.

Li, D., and Davis, D. (2005). Transition of railroad bridge approaches. *J. Geotech. Geoenviron. Eng.* 131, 1392–1398. doi: 10.1061/(ASCE)1090-02412005131:111392

Luomala, H., Rantala, T., Kolisoja, P., and Mäkelä, E. (2017). "Assessment of track quality using continuous track stiffness measurements," in *Georail 2017, 3rd International Symposium Railway Geotechnical Engineering*, (Marne La Vallée: IFSTTAR), 281–290.

Novák, N., Perfilieva, I., Dvořák, A., Chen, G., Wei, Q., and Yan, P. (2008). Mining pure linguistic associations from numerical data. *Int. J. Approximate Reasoning.* 48, 4–22.

Piché, R., Järvenpää, M., Turunen, E., and Šimůnek, M. (2014). Bayesian analysis of GUHA hypotheses. *J. Intell. Inf. Syst.* 1, 42–47. doi: 10.1007/s10844-013-0255-6

Rauch, J. (2013). "Observational calculi and association rules," in *Studies in Computational Intelligence*, 1st Edn, Vol. 469, (Berlin: Springer).

Sadeghi, J., and Askarinejad, H. (2009). An investigation into the effects of track structural conditions on railway track geometry deviations. *Proc. IMechE. Part F J Rail Rapid Transit.* 223, 415–425. doi: 10.1243/09544097JRRT266

Sadeghi, J., and Askarinejad, H. (2012). Application of neural networks in evaluation of railway track quality condition. *J. Mech. Sci. Technol.* 26, 113–122. doi: 10.1007/s12206-011-1016-5

Sauni, M., Luomala, H., Kolisoja, P., and Turunen, E. (2020). "Determining sampling points using railway track structure data analysis," in *Proceedings of the 3rd International Conference (ICITG)*. Guimarães: Springer, 841–856. doi: 10.1007/978-3-030-32029-4_71

Tsui, K. L., Chen, V., Jiang, W., and Alp Aslandogan, Y. (2006). "Data Mining Methods and Applications," in *Springer Handbook of Engineering Statistics*, ed. H. Pham (London, UK: Springer-Verlag), 651–669.

Turunen, E. (2018). Paraconsistent many–valued logic in GUHA framework. *Acta Informatica Pragensia.* 7, 104–111. doi: 10.18267/j.aip.116

Weston, P., Roberts, C., Yeo, G., and Stewart, E. (2015). Perspectives on railway track geometry condition monitoring from in-service railway vehicles. *Vehicle Syst. Dyn.* 53, 1063–1091. doi: 10.1080/00423114.2015.1034730

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# PUBLICATION
# III

## Determining Sampling Points Using Railway Track Structure Data Analysis

Sauni, M., Luomala, H., Kolisoja, P. and Turunen, E.

Information Technology in Geo-Engineering. ICITG 2019. Springer Series in Geomechanics and Geoengineering, pp. 841–856.
10.1007/978-3-030-32029-4_71

# Determining Sampling Points Using Railway Track Structure Data Analysis

Mikko Sauni[1[0000-0002-7034-5835]], Heikki Luomala[1[0000-0002-7113-3527]], Pauli Kolisoja[1[0000-0001-7709-180X]] and Esko Turunen[1[0000-0002-1194-6497]]

[1] Tampere University, Tampere 33720, Finland
`mikko.sauni@tuni.fi`
`heikki.luomala@tuni.fi`
`pauli.kolisoja@tuni.fi`
`esko.turunen@tuni.fi`

**Abstract.** In railway track asset management, limited funding is available to ensure safe and punctual train traffic on an aging rail network. Assessing railway structure problems, their severity and extent is a difficult and laborious task to which different methods have been applied. Further, determining problematic areas and identifying their rehabilitation needs are two separate operations.

In this research, data mining and data analysis of railway track structure data was used to identify different types of track behavior and corresponding substructure conditions. A descriptive data mining method, Generalized Unary Hypothesis Automata (GUHA), was adopted. Soil samples were taken and tested on the basis of the conducted analyses. The purpose was to see whether deductions made from the data, concerning the condition of the track substructure, could be confirmed with soil sampling and related soil sample laboratory tests.

The research was carried out in three parts. First, multiple data sources were used to comprise an initial data matrix, which was used in data mining and data analyses. After the analyses, fifty subballast and ten ballast sampling points were chosen according to the findings from data mining and data analysis, and samples were taken and tested. The last part of the research was to see how the laboratory test results corresponded with the analyses made from the data.

The research showed that GUHA data mining and data analysis can be used to detect sections of track with problematic substructures, but further research is required to improve the initial data.

**Keywords:** Data Analysis, Data Mining, GUHA Method, Railway, Sampling.

## 1 Introduction

Designing remediation for aged railway structures can be challenging, especially when using limited resources to do so. Design should begin with finding problematic areas, identifying different types of problems, and assessing remediation to those specific problems. Track geometry deterioration due to track structure degradation is observed at different rates in different sections of track. However, there is often uncertainty in explaining why some sections of track are more problematic than others.

Some problems of track structures may be more easily explained than others. Problems related to discontinuity in the rail or the track structure can be obvious, for example, problems with stiffness variations in bridge transitions. On the other hand, substructure problems can be difficult to detect because the problems might be invisible and dependent on seasons for example. For example, treating subgrade problems with just tamping will not fix the long-term problems of track geometry [1]. In this paper, track substructure and its components are defined according to Li et al. [2]; subballast is defined as the granular layer below the ballast and above the subgrade.

Plenty of data can be produced during the life cycle of a railway line including track geometry history, ground penetrating radar (GPR) measurements, track deflection measurements, and laser scanning point clouds, for example. However, the information about the track structure can be overwhelming to address manually.

Data mining and data analyses can be used to reveal interesting patterns in vast amounts of data. Data mining and data analyses have been used in the railway sector, for example to predict track geometry deterioration, optimize track maintenance, and model track asset management [3–6].

The Generalized Unary Hypothesis Automata (GUHA) data mining method can be used to describe an available data set. The method produces hypotheses based on the data, which are statements which the data either supports or does not support. The GUHA method can be used on historical data of the track structure to determine, for example, what has been the observed condition of the track structure on certain types of track structures.

In this research, the railway track structure condition of an old railway line in Finland was assessed using GUHA data mining and data analysis. Substructure material sampling points were determined according to data mining and data analyses results. The taken soil samples were tested for grain size distribution and capillary rise. The goal was to assess the applicability of the GUHA method to railway data and the validity of data currently available concerning railway substructures.

## 2 Background

### 2.1 Track Structure Degradation

In an ideal situation, newly constructed railway structures would settle uniformly, and the track geometry would remain undisturbed apart from descending. This situation cannot be achieved because the track structure nor the dynamic traffic loads on practically any given track section are homogenous and differences in settlements occur. Differences in track settlements cause the deterioration of track geometry on track sections and even in individual cross sections [7].

Shenton [8] has provided a comprehensive explanation for track geometry deterioration with six interacting factors: 1) dynamic forces 2) rail shape 3) sleeper spacing 4) sleeper support 5) ballast settlement 6) substructure differential settlement, in addition to all of the combinations and interactions of the aforementioned. Shenton has also provided causation, in which factors 2–6 increase dynamic loading if a train is moving at a reasonable speed. Shenton's explanation is encompassing, but it does not explicitly elaborate what causes differential settlements in the substructure.

Li & Selig [9] have provided their explanation on the problems of earth structures, which consist of four factors and their interactions: 1) excessive loading (self-weight and repeated dynamic loading) 2) fine graded soils 3) high moisture 4) freezing and thawing.

Later, Li et al. [2] have identified the track substructure as the single most crucial element related to track performance. According to Li et al., a good track substructure has a strong resistance to plastic deformation and provides uniform elastic deformations, which can be achieved by having a well-drained layer consisting of angular and durable particles that resist abrasion. A poor substructure is described by a high moisture and high fines content [2].

According to these explanations, track structure degradation and its manifestation as track geometry deterioration can result from several reasons. However, if discontinuity in the rail or track structure and faults in the superstructure can be excluded with available information, problematic substructures should be detectable when fine graded soil materials or high moisture contents are observed.

## 2.2 GUHA Data Mining

Data mining, in general, is a collection of computer aided data analysis techniques that focus on discovering and modeling meaningful information in big data masses.

GUHA, an acronym for Generalized Unary Hypothesis Automata, is a logic-based method of a descriptive data mining procedure implemented to LISp-Miner software [10–12]. Given a large data matrix, in this research of 32 columns and 52,907 rows, acquiring answers to questions related to the data is possible. An encompassing description of the initial data used in this research is provided in section 3.2 and Table 2. Briefly summarized here, the columns in the data matrix represent properties (called predicates or attributes) the examined objects have, and each row characterizes an object; it either has the properties, does not have the properties, or the cell is empty. Examples of columns (predicates) in this research are structural thickness (12 categories with 0.2 m intervals), structural moisture index (7 categories), and data on assets such as bridges (3 categories) and culverts (binary categories).

Analytic questions are presented by using generalized quantifiers such as usually implies, above average, almost equivalent, etc. There are dozens of generalized quantifiers implemented in the LISp-Miner software. The software goes through the data where applicable and outputs dependencies that the data supports. Some of the answers may be already familiar to the user, but there may also be some interesting new ones. Next, three different LISp-Miner procedures used in this research are presented, as well as some interesting results found in the initial data.

Founded implication quantifiers are used to find dependencies: 'The presence of $A$ implies the presence of $B$ with confidence $p$ and support $n$'. One of the questions presented in data mining was what kind of track structures exhibit a high track geometry deterioration rate with over 80% confidence when at least 700 rows of data support the statement. After 10,703,034 verifications, 238 hypotheses were supported by the data, e.g., a hypothesis: a high track geometry deterioration rate is observed on 86% of structures that exhibit high moisture indices and high track deflection variance.

Above average quantifiers are used to find situations where among cases satisfying *A*, there are at least 100**p*% more objects satisfying *B* than there are cases satisfying *B* in the whole data. This approach was used to get an answer to inquire, for example, which track structure properties or their combinations have an above average correlation to some type of track geometry deterioration rate. After 15,762,789 verifications, 278 hypothesis were supported by the data, e.g., a hypothesis: the observed track geometry deterioration rate is high on 96% of track structures that are less than 1.4 m thick, no frost insulation board are installed, and the track deflection mean and variance are high. The corresponding correlation is 17% on other structures.

In the Action mining approach, the idea is to find dynamic features in data; some predicates are considered stable attributes and some others are flexible attributes. One of the analytical questions examined in this research is the following: how do changes in the ballast moisture index influence the observed track geometry deterioration rate when other parameters are stable. After 715,806 verifications, 437 hypothesis were supported, e.g., when the track substructure thickness is 1.0–1.6 m and ballast thickness is < 500 mm, a low track geometry deterioration rate is observed on 75% of structures with a low ballast moisture and on 15 % of structures with a high ballast moisture.

The hypotheses are translated into comprehensible language from the contingency tables that the method and program produce. In practice, LISp-Miner tests 2x2-contingency tables obtained from the data matrix. For example, for the above average quantifier, to accept a dependence '*A* among those cases that satisfy *B* is 4 times more frequent than *A* in the whole data', a condition

$$a/(a+b) > (1+p)((a+c)/(a+b+c+d) \qquad (1)$$

must be satisfied by the data in Table 1. For *A* 'Track structures thickness < 1.4 m, no frost insulation board, track deflection mean is high, deflection variance is high' and *B* 'High track geometry deterioration rate', the corresponding 2x2-contingency table is presented in Table 1; direct calculation shows that the data supports this dependence. Of course, every quantifier has its own condition for acceptance. The quantifiers and their formulas are presented in detail by Rauch [11].

**Table 1.** An example of a contingency table.

| | B | not B |
|---|---|---|
| A | a = 501 | b = 19 |
| not A | c = 8,821 | d = 43,466 |

# 3 Research Process

## 3.1 Case Track Section Kouvola–Kotka

The case study track section in this research, Kouvola–Kotka, is located in the south-eastern coastal area of Finland. The track section is a double track from Kouvola to

Juurikorpi and a single track from Juurikorpi to Kotka, with respective lengths of 35 km and 18 km.

The case study track section was originally completed as a single line track in 1890, and the double track sections have been built in the 1950s and 1990s. The age of the track section implies that the structures may not be fully compliant with today's standards. For example, the required minimum thickness of non-frost susceptible materials in a track structure on this area is 2 m, whereas the mean thickness of all structures is 1.96 m. This means that there are plenty of undersized structures along the track section. Further, the materials used may not be compliant with today's requirements.

Frost insulation boards are often installed into old track structures in Finland to reduce frost penetration into the track structure. On the case study track section, frost insulation boards have been installed on over 16 km of the 54 km track section. Frost insulation boards should be installed in the subballast 300 mm below the ballast layer, according to Finnish guidelines [13]. However, if frost insulation boards are installed when the ballast is undercut and cleaned, frost insulation boards are installed directly below the ballast layer. Unfortunately, documented information about the installation of frost insulation boards is rarely available.

## 3.2 Initial Data

The initial data included track geometry car measurements, GPR interpretations, track deflection measurements, laser scannings, and asset data (Table 2). The initial data was comprised into a single matrix, where the rows represented one-meter long sections of the track, which are described by their properties presented in the columns.

**Table 2.** Initial data sources, processing, and usage.

| Data origin | Pre-processing | Data used for | Data type |
|---|---|---|---|
| Track geometry car measurements | Annual growth of running 20 m chord standard deviation | Track geometry deterioration rate | Ratio<br>1 variable |
| GPR | Signal rebound calculations | Structural layer thicknesses | Ratio<br>4 variables |
| GPR | Signal attenuation calculations | Moisture damage index MDI | Ratio<br>7 variables |
| Continuous laser scanning | Minimum elevation 2–6 m perpendicular from track center line | Ditch depth (both sides individually) | Ratio<br>2 variables |
| Continuous track deflection measurement | Running 20 m chord mean value | Track deflection | Ratio<br>1 variable |
| Continuous track deflection measurement | Running 20 m chord variance | Track deflection variation | Ratio<br>1 variable |
| Soil maps | Interpretation | Subsoil classification | Categorical<br>1 variable |
| Video and asset data warehouse | Visual inspection | Track assets (bridges, turnouts, culverts, etc.) | Categorical and binary<br>7 variables |

The initial data was gathered in the same way over the entire length of the track section. On the double track section, the measurements and analyses concern the western track only. Continuous measurements included track geometry, GPR, laser scanning, and deflection measurements. Soil maps and asset data were homogenous too over the entire length of the track section. Video of the track was also provided, but it was used only to confirm and validate other data.

The initial data contained little missing values. The missing values were imputed using other available data. For example, an empty value for a bridge was set to mean that there was no bridge and video of the track was used to validate the imputation.

The input data for track geometry deterioration rate calculations was the track geometry car (Ttr1 51) semiannual measurements from 2008 to 2018. The track geometry deterioration rate was calculated using the annual growth of a 20 m chord standard deviation of the measurements. The mean of the track geometry deterioration rate for the whole track section was 0.14 mm/a. Over 75% of the track section displayed a track geometry deterioration rate less than the track section's average. This indicates that there are few sections of track where the track geometry deterioration rate is high, but on those sections, the rate is very high.

GPR measurements provided information on the substructural layer thicknesses and perceived moisture of the substructural layers. Layer thicknesses were used both independently and as a sum to indicate combined ballast, subballast, and embankment thickness. A moisture index was calculated separately for the ballast, subballast, and subgrade. Also, a combined value of the aforementioned, as described by Arnold et al. [14], representing the moisture content of the whole substructure and moisture damage index (MDI), was provided.

A laser scanning point cloud was used to calculate the ditch depth, in order to provide information about drainage conditions. The minimum depth 2–6 m perpendicular to the track center line was calculated. Further, a 20 m chord minimum of the aforementioned was calculated to reduce error due to foliage and wayside equipment. This value depicted the ditch depth and was calculated individually for both sides of the track.

A continuous track deflection measurement car, as presented in more detail by Luomala et al. [15], was used to measure track deflection. A 20 m chord mean and variance values were both used: the former to indicate the level of deflection and the latter to indicate changes in the deflection.

Other sources, such as soil maps, asset management data warehouses, and a video check were used to identify track assets that influence the performance of track structures. The track asset data included bridges, culverts, turnouts, level crossings, frost insulation boards, stations, cuttings, and subsoil assessments. This data was either binary or conformed to classes using dummy values such as 0 for embankment, 1 for rock cutting, and 2 for soil cutting.

### 3.3 Applying GUHA Data Mining

The initial data was used as an input in the LISp-Miner program, which is an application of the GUHA method. Questions about the correlations between track structure properties and the track geometry deterioration rate were inquired. Both outcomes, high and low track geometry deterioration rates, were inquired from several viewpoints. Quantifiers p-implication (PIM) and above average dependence were used in modules 4ft-Miner, SD4ft-Miner, and Ac4ft-Miner, but only relevant queries were reported. The data backing the hypotheses was visualized, and video from the areas was checked to verify that no other explaining features outside the hypotheses' data could be observed.

All of the queries used track structure variables as antecedents and the track geometry deterioration rate as the succedent. Conditions were applied in query 5 and 6 to eliminate unwanted variables from data mining. Several other queries were also conducted and many hypotheses per query were generated, but only the relevant queries and relevant hypotheses to this research are presented. Altogether, 69 queries were made and thousands of hypotheses were generated.

The following queries about the data were conducted and reported. Detailed information about formation and outcomes the queries is provided in Table 3.

1. What kind of combination of track structure variables is associated with a certain type of track geometry deterioration rate with more than 90% confidence?
2. What kind of combination of track structure variables is associated with a higher than average track geometry deterioration rate with more than 80% confidence?
3. How does a change in the variable for structure moisture affect a certain type of track geometry deterioration rate when other track structure variables are stable?
4. How does a change in the variable for overall structure thickness affect the most common type of track geometry deterioration rate when other track structure variables are stable?
5. How does a change in the variable for overall structure thickness affect a certain type of track geometry deterioration rate when other track structure variables are stable, and structures founded only on embankments or soil cuttings are examined?
6. How does having a frost insulation board affect a certain type of track geometry deterioration rate when only structures that are 1.6 m to 2.4 m thick are examined?

The relevant non-trivial hypotheses are presented below. The hypothesis number corresponds to the query number, for example, hypothesis number one is a result of query number one.

1. Track geometry deterioration rate is low on 92% of structures that are built on embankments, over 2.8 m thick, and have over 650 mm thick ballast layers, low moisture indices, and little track deflection variance.
2. Track geometry deterioration rate is high on 87% of structures that are built on embankments, have a low embankment thickness (< 0.5 m), and the track deflection mean and variance are high.
3. Track geometry deterioration rate is low on 86% of structures with low moisture indices, whereas the corresponding percentage is 38% on structures with high moisture indices, when in both cases structures are built on embankments and do not have a frost insulation board in their structure.

4. A high track geometry deterioration rate is as common (about 84% of structures) on structures less than 1.4 m thick as a low track geometry deterioration rate is on structures 1.6–2.0 m thick, when the track is built on an embankment, there is no frost insulation board in the track structure, and the ballast layer is less than 450 mm thick.
5. Track geometry deterioration rate is low on 74% of structures less than 1.4 m thick, whereas the corresponding percentage is 36% on structures 1.8–2.4 m thick, when in both cases structures have a frost insulation board in their structure, and the track deflection mean and variance are low, and only structures founded on soil cuttings or embankments are regarded.
6. On structures that have deep ditches, track deflection is low, and a less than 550 mm thick ballast layer, the track geometry deterioration rate is low on 79% of structures that do not have a frost insulation board, whereas the corresponding correlation is 16% on equivalent structures that have a frost insulation board, when structures only 1.6–2.4 m thick are regarded.

**Table 3.** List of the reported LISp-Miner data mining tasks.

| Query/ hypothesis | Module | Statistical quantifier | Frequencies quantifier | Verifications | Number of hypotheses |
|---|---|---|---|---|---|
| #1 | 4ft | PIM > 0.9 | a > 1,500 | 3,934,998 | 45 |
| #2 | 4ft | PIM > 0.8 | a > 700 | 10,703,034 | 238 |
| #3 | Ac4ft | PIM before > 0.7 &<br>PIM difference > 0.4 | a (before) > 2,000 | 134,344 | 48 |
| #4 | Ac4ft | PIM before > 0.8 &<br>PIM after > 0.8 | a (before and after) > 400 | 69,071,100 | 2 |
| #5 | Ac4ft | PIM difference > 0.3 | a (before) > 1,000 | 45,736 | 25 |
| #6 | Ac4ft | PIM difference > 0.4 | a (before) > 500 | 19,926 | 233 |

The hypotheses indicated that the moisture content, deflection variance, the thickness of track structure, and frost insulation boards affect the perceived performance of the track structure. The higher moisture content the structure displayed, the more track deflection the structure exhibited, and the thinner the structure was, the higher the track geometry deterioration rate was. These results are in line with the literature review in section 2.1.

Frost insulation boards produced somewhat conflicting results. Structures less than 1.4 m thick have performed well when a frost insulation board was detected in the structure. However, the data indicated poor performance on structures which were 1.6–2.4 m thick and had a frost insulation board in the structure. These structures would not generally require a frost insulation board if the materials used to build the subballast are not frost susceptible.

### 3.4 Data Analysis

The initial data was visualized in the Rail Doctor® program so that interesting structures could be located and sampling points planned. Using the information gained from literature (section 2.1) and data mining (section 3.3), structures with above average track geometry deterioration rates, structures with high moisture indices, thick structures with frost insulation boards, and structures with high deflection variation were targeted. The sampling points were manually selected using the above-mentioned criteria and avoiding discontinuity areas.

An example of data visualization is presented in Fig. 1. The x-axis represents track kilometers, and track structure data is represented on the y-axis. Starting from the top, the y-axis contains structural layer boundary depths, asset data, MDI in color maps, relative structural moisture in graphs, subgrade soil classification assessment, the track geometry deterioration rate, and track deflection.

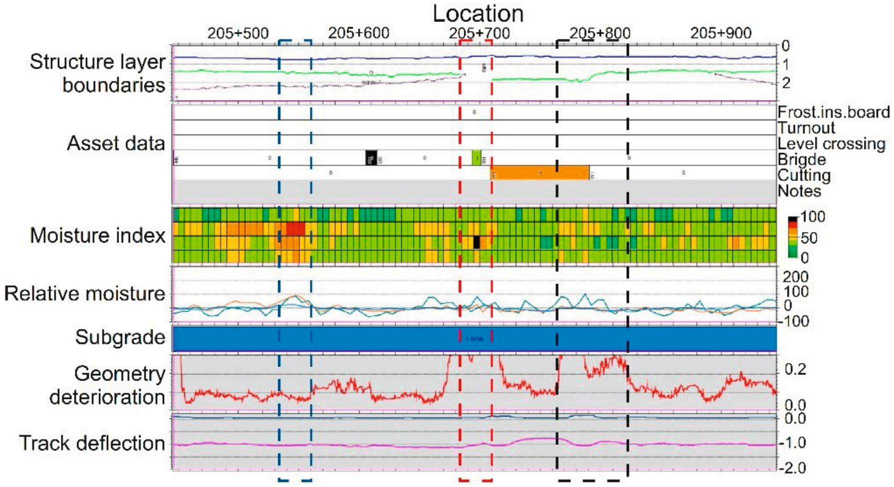


**Fig. 1.** Visualization of data in Rail Doctor®.

In Fig. 1, three areas are bordered by dashed lines. The left most bordered area exhibits high subballast and subgrade moisture on an over 2 m thick embankment. On that area, the track geometry deterioration rate is lower than average, and no track deflection variations are detected. The right most bordered area contains a soil cutting, where slightly moist subballast is observed, the track geometry deterioration rate is high, and track deflection is locally high, which is indicated in the mean and variance of track deflection. These two bordered areas are of interest with regard to subballast sampling.

Discontinuity in the rail or the track structure, such as bridges, turnouts, and culverts, could easily be detected when the data was visualized. Some of these areas were interpreted to be problematic due to their geometry history and deflection, which differed vastly from other sections of the track. An example of a problematic bridge transition can be found in the dashed area in the middle of Fig. 1. Two peaks in the track geometry deterioration rate are observed in the transitions to the bridge, and track deflection fluctuates. Track discontinuity areas, while interesting, were not the subject of this research.

### 3.5 Sampling and Laboratory Tests

Fifty subballast and ten ballast sampling points were selected along with 15 and 2 backup points for subballast and ballast, respectively. Various substructural conditions were required, therefore the subballast sampling points were selected as follows: 10 well performing points with a low moisture content, 10 well performing points with a high moisture content, and 30 problematic areas possibly due to substructure conditions. The ballast samples were selected from structures that were interpreted to be problematic according to the initial data.

The subballast samples were taken from two depths: 300–600 mm and 600–900 mm below the track bench (adjacent to the ballast shoulder). Some additional samples were taken from varying depths if a clear soil layer boundary was detected while taking the samples. Altogether, 118 subballast samples were taken.

All subballast samples were subjected to sieving in accordance with SFS-EN 933-1:2012 [16]. The grain size distribution and natural water content of all samples were investigated. Also, the coefficients of uniformity $C_u$ and curvature $C_c$ were calculated.

The grain size distributions were surprisingly similar throughout the track section. There were 33 samples in which the coefficient of uniformity $C_u$ was less than 5, meaning that the range of particle sizes was narrow. Thirty-one samples exceeded the limit values given in Finnish guidelines for subballast materials on the fine graded side of the grain size distribution scale. Nevertheless, only six samples exceeded the fines ($\leq$ 0.063 mm particle size) content limit of 4%. In some samples, a clear presence of the ballast material was detectable, but this was an expected result because the same observation was already determined while taking the samples.

The capillary rise test was subjected to 60 subballast samples which were taken mainly from the lower depth. Capillary rise was tested by placing samples in plastic tubes in shallow water for one week and measuring the highest visible waterline in the sample. This value represented the sample capillary rise. The average capillary rise of all samples was 31.5 cm. 5.5 cm and 66 cm were the minimum and maximum values, respectively. All samples exhibited very low fully saturated zones.

The ballast samples were retrieved and sieved according to a Finnish national guideline for ballast material sampling. The ballast sample was retrieved 30–40 cm below the bottom of the rail. A sample weighing 6–8 kg was taken, using a shovel, from between two sleeper ends from an area the size of 20 cm by 20 cm and 10 cm deep. Ballast samples were sieved, and material found on 1 mm, 8 mm, and 25 mm sieves was recorded. A ballast fouling index, described by the sum of the percent finer by weight on each of the aforementioned sieves, was used in determining sample quality. According to Finnish guidelines, if the ballast fouling index exceeds the limit value of 90, the ballast layer must be renewed or cleaned [17]. All ballast samples exhibited low ballast fouling indices, which were between 7.2 and 21.9. This means that all tested ballast materials are, according to the guideline, in very good condition.

# 4 Results and Discussion

## 4.1 Soil Sample Laboratory Test Result Analysis

Correlations between the soil sample laboratory test results and the initial data were investigated. Obtained relevant results are presented in this section. Fig. 2 and Fig. 3 show box plots of different property classes' track geometry deterioration rate of all samples and samples taken only from structures that do not have frost insulation boards. Fig. 4 presents a box plot of different property classes' d50 grain size concerning all samples. The boxes in the boxplot represents the second quartile of the data. The vertical lines (whiskers) represent the lowest and highest data points within the 1.5 inter quartile range of the lowest or highest quartile, respectively. Outlier points can be found outside the whiskers as empty dots. The crosses in the box plot are means and the horizontal lines are medians.

The properties in the graphs include capillary rise, frost insulation boards, MDI, d50 grain size, sample depth, material consistent with guideline grain size distribution for subballast (pass or fail), and coefficient of uniformity. Samples from structures with frost insulation boards were ignored in MDI boxes, because the GPR results beneath frost insulation boards are disrupted by the boards and do not represent true values.

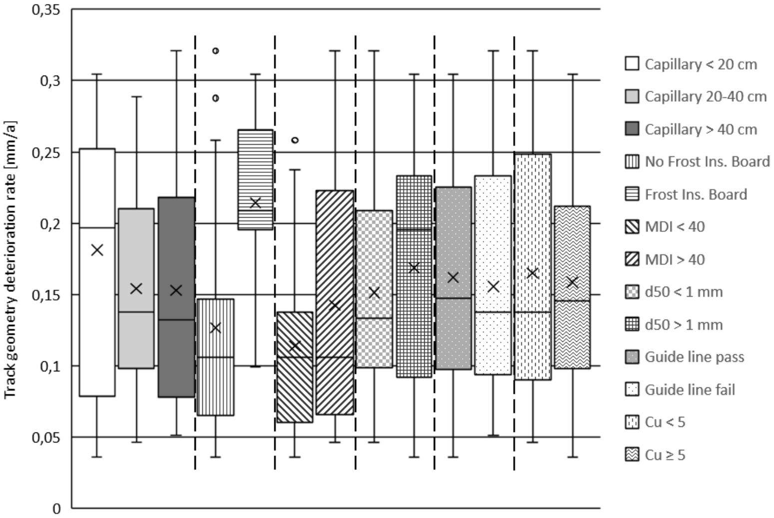


**Fig. 2.** Box plot of the track geometry deterioration rate of all samples.

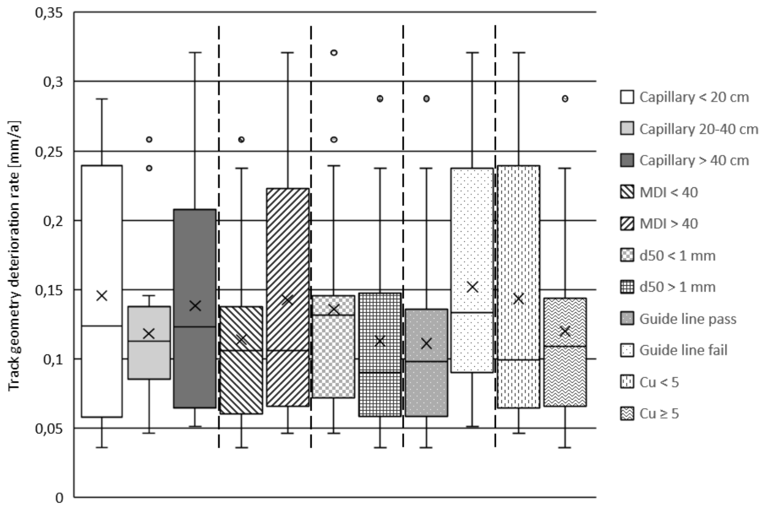


**Fig. 3.** Box plot of the track geometry deterioration rate of samples taken from structures that have no frost insulation boards.

Fig. 2 denotes that material properties had no practical effect on the track geometry deterioration rate if all samples are examined. Notable differences in the track geometry deterioration rate are observed only with regard to the MDI and frost insulation boards. High MDI and frost insulation boards are associated with a high track geometry deterioration rate. The effects of frost insulation boards, however, appear to be more dominant to the track geometry deterioration rate than subballast material properties.

If structures that have frost insulation boards are ignored, the influence of subballast material quality in the track geometry deterioration rate can be observed. In Fig. 3, the most distinct result in the sample material quality is that samples consistent with grain size distribution guidelines clearly exhibit lower track geometry deterioration rates compared to samples inconsistent with guidelines. The result, though, is not as notable as the effects of frost insulation boards in Fig. 2.

Further, in Fig. 3, a higher MDI, a smaller d50 grain size, and higher coefficient of uniformity values appear to have some correlation to higher track geometry deterioration rates. Unexpectedly, a low capillary rise appears to have a slightly higher track geometry deterioration rate compared with a high capillary rise.

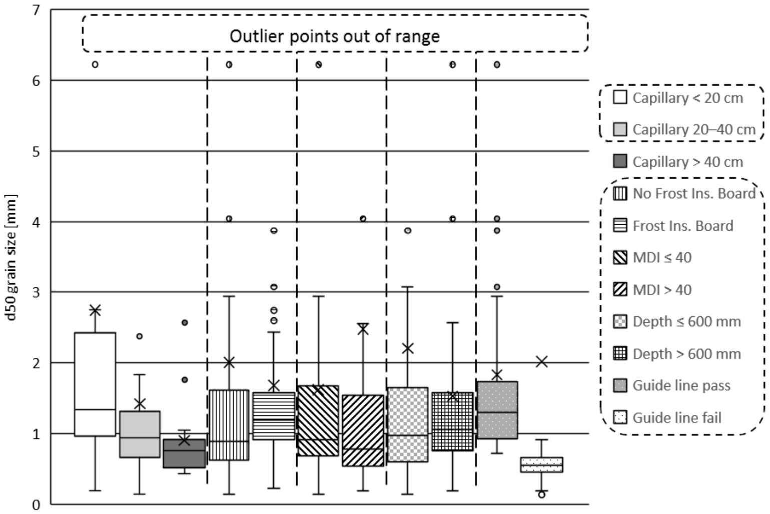


**Fig. 4.** Box plot of d50 grain size of all samples. Outlier points are outside of the graph range in all categories except 'Capillary > 40 cm'.

Fig. 4 represents the distributions of d50 grain size with regard to other parameters. Outlier points are present even out of the plotting range (14–35 mm) due to ballast material found in the subballast material, which increases the mean values in all bars except the > 40 cm capillary rise. d50 grain size and capillary rise correlated intuitively; A lower capillary rise was detected on samples that had larger size particles and vice versa. Structures with or without frost insulation boards did not exhibit much variation in their d50 grain size distributions. Higher MDI values had a higher mean d50 than lower MDI values, even though the medians and upper and lower quartiles were similar. Samples taken from a lower depth had slightly smaller d50 grain sizes, but again, the medians and upper and lower quartiles were quite uniform. Material in accordance with design guidelines was clearly coarser than material that did not meet the guideline limits.

In many cases, though, GPR interpretations can indicate a significantly moist subballast layer, yet no major variation in the subballast material moisture content or quality was detected in the samples taken from these locations, even when structures with frost insulation boards are ignored. This may indicate that GPR interpretations are influenced by the prevailing conditions of the track, which are dependent on many more factors than just the subballast material. Another factor influencing the correlation between GPR interpretations and subballast sample test results is the representativeness of samples taken from the slope of the embankment in relation to the material directly beneath the center line of the track.

### 4.2 Findings from Sample Locations

The GUHA data mining indicated that frost insulation boards were associated with problematic structures when installed in thick structures. However, the materials found in structures having frost insulation boards did not differ much from the non-problematic structures without frost insulation boards.

For instance, samples taken from around the 218-kilometer pole had no practical difference in their grain size distribution, natural water content, or capillary rise. However, the measured track geometry had deteriorated more than twice as fast on the section with frost insulation boards than that on the whole track section on average. As track structures or sub soil conditions do not vary much on this section, the only variable in the track structure is the frost insulation board.

Sampling revealed an extruded polystyrene foam frost insulation board installed directly underneath the ballast. The frost insulation board itself was in good condition, but the structure above and below the board was moist. Further, the board was covered with a fine graded material, which was most likely fouled ballast (Fig. 5).

**Fig. 5.** A frost insulation board underneath ballast at sample point 216+068. Photo credit: Toni Saarikoski.

The results indicate that more attention should be paid to the installation of frost insulation boards. The drawbacks of installing a frost insulation board directly under the ballast layer instead of installing it in the subballast should be investigated further because detrimental effects regarding structures' long-term performance may occur. Frost insulation boards may have a fouling effect on the ballast material due to increased stiffness variations in addition to drainage conditions that may be compromised.

Further, the ballast material of the sample point 216+068 was tested in accordance with guidelines in effect today in Finland. According to the results, the ballast material is of very good quality even though fouling is clearly detectable as presented in Fig. 5. This controversy results from taking the sample too high up in the ballast layer where a little amount of the fouled material can be found. These results give a good reason to review the guidelines in effect in Finland concerning ballast sampling and testing.

# 5 Conclusions

GUHA data mining and data analyses were used to assess the railway substructure conditions. Subballast and ballast sampling points were chosen according to the data mining and data analyses results. The material properties of the subballast samples and the initial data were used to examine substructural conditions and their effects on the track geometry deterioration rate.

The following conclusions were made in this research:

- The GUHA method is a novel approach to analyzing railway data. The practical benefits of using GUHA in analyzing railway data are best obtained in the early stages of designing maintenance or rehabilitation of an old railway structure, when maintenance data is abundant.
- The GUHA method could point out specific types of substructures to be problematic. The results from the GUHA data mining were in line with literature concerning explanations for track structure degradation.
- Using the knowledge gained from the GUHA method, the sampling can be focused on problematic structures, which would increase efficiency in design.
- The analyzed sampling data combined with the initial data indicated that frost insulation boards displayed a dominant correlation to the track geometry deterioration rate; structures with frost insulation boards were found to be more problematic than structures without frost insulation boards, even when the subballast material or substructure formation did not differ.
- However, if structures with frost insulation boards were ignored, material quality did exhibit correlations to the track geometry deterioration rate. In that case, a distinct correlation was detected between the subballast grain size distribution and track geometry deterioration rate. If the subballast sample grain size distribution was found to be consistent with current guidelines, a lower track geometry deterioration rate was observed.
- Individual material properties also resulted in an effect; coarse and well graded materials displayed a minor correlation to low track geometry deterioration rates. Further, a high MDI interpreted from GPR measurements displayed a considerable correlation to a high track geometry deterioration rate.
- GPR interpretations of soil moisture content did not always correlate with the subballast sample test results; some GPR interpretations indicated individual high substructure moisture contents in places where the substructure material was found to be dry and had a low fines content. This phenomenon was observed even when no GPR signal disturbing structures, such as frost insulation boards, were detected.
- Ballast sampling conducted in accordance with current Finnish guidelines did not give a true representation of the ballast material found in the structure, due to the fouled ballast material located lower than where the sample is to be taken.
- Future research should focus on depicting substructural conditions more explicitly in the initial data, especially the idea that parametrization of drainage conditions should be incorporated into future track maintenance data.

# References

1. Li, D.: 25 years of heavy axle load railway subgrade research at the Facility for Accelerated Service Testing (FAST). Transportation Geotechnics 17, 51–60 (2018). doi: 10.1016/j.trgeo.2018.09.003
2. Li, D., Hyslip, J., Sussmann, T., Chrismer, S.: Railway geotechnics. CRC Press, Taylor & Francis Group, Boca Raton (2016). doi:10.1201/b18982
3. Andrade, A., Teixeira, P.: A Bayesian model to assess rail track geometry degradation through its life-cycle. Research in Transportation Economics 36(2012) 1–8 (2018). doi:10.1016/j.retrec.2012.03.011
4. Andrade, A., Teixeira, P.: Statistical modelling of railway track geometry degradation using Hierarchical Bayesian models. Reliability Engineering and System Safety 142(2015) 169–183 (2015). doi:10.1016/j.ress.2015.05.009
5. Vale, C., Lurdes, S.: Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line. Reliability Engineering & System Safety Vol. 116, s. 91–98 (2013). doi:10.1016/j.ress.2013.02.010
6. Andrews, J., Prescott, D., Rozières, F.: A stochastic model for railway track asset management. Reliability Engineering and System Safety 130, 76–84 (2014). doi:10.1016/j.ress.2014.04.021
7. Esveld, C.: Modern railway track 2nd edn. MRT-Prod, Zaltbommel (2001).
8. Shenton, M.: Ballast deformation and track deterioration. Track technology, 253–265 (1985).
9. Li, D., Selig, E.: Evaluation of railway subgrade problems. Transportation research record, 1489, 17–23 (1995).
10. Hájek, P., Havránek, T.: Mechanizing Hypothesis Formation: Mathematical Foundations for a General Theory. 1st ed. Springer, Berlin/Heidelberg (1978).
11. Rauch, J.: Observational Calculi and Association Rules. Studies in Computational Intelligence 469. Springer, Heidelberg (2013). doi:10.1007/978-3-642-11737-4
12. Berka, P.: Practical aspects of data mining using LISp-Miner. Computing and Informatics Vol. 35, 528–554 (2016).
13. Finnish Transport Infrastructure Agency (in Finnish: Liikennevirasto): Ratatekniset ohjeet (RATO) osa 3 Radan rakenne, annex 5 (2018).
14. Arnold, G., Fon Sing, P., Saarenketo, T., Saarenpää, T.: Pavement moisture measurement to indicate risk to pavement life. New Zealand Transport Agency research report 611, p. 22 (2017).
15. Luomala, H., Rantala, T., Kolisoja, P., Mäkelä, E.: Assessment of track quality using continuous track stiffness measurements. In: 3rd International Symposium Railway Geotechnical Engineering: 23-24 November 2017, pp. 218–289. IFSTTAR, Marne La Vallee, France (2017).
16. SFS-EN 933-1. Tests for geometrical properties of aggregates. Part 1: Determination of particle size distribution. Sieving method. (2012).
17. Finnish Transport Infrastructure Agency (in Finnish: Ratahallintokeskus): Ratatekniset ohjeet ja määräykset (RAMO) osa 15 Radan kunnossapito (2002).

# PUBLICATION
# IV

## Framework for implementing track deterioration analytics into railway asset management

Sauni, M., Luomala, H., Kolisoja, P. and Vaismaa, K.

Built Environment Project and Asset Management, Vol. 12 No. 6, pp. 871–886.
10.1108/BEPAM-04-2022-0058

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2044-124X.htm

# Framework for implementing track deterioration analytics into railway asset management

Mikko Sauni, Heikki Luomala, Pauli Kolisoja and Kalle Vaismaa
*Research Centre Terra, Tampere University, Tampere, Finland*

Received 2 May 2022
Revised 22 August 2022
Accepted 28 September 2022

## Abstract

**Purpose** – Recent research outputs can be difficult to implement into ongoing safety critical processes. Hence, research is well beyond current practices in railway asset management. This paper demonstrates the process of creating tangible change within a railway asset management organization by introducing a framework for advancing track geometry deterioration analyses (TGDA) in practice.

**Design/methodology/approach** – The research was conducted in three parts: (1) maturity models were reviewed and adapted as the basis for the framework, (2) the initial maturity level was investigated by conducting semi-structured expert interviews, and (3) a framework for development was created in cooperation with stakeholders during three workshops. The methodology and findings were tested and applied in the Finnish state rail network asset management.

**Findings** – The main output of this study is the framework for advancing TGDA in railway asset management. The novel framework provides structure for controlled incremental development, which is essential when altering a safety critical process.

**Practical implications** – The research process was successfully applied in Finland. Following the steps presented in this article, any organization can apply the framework to plan their development schemes for railway asset management.

**Originality/value** – Full-scale implementation of novel models and methods is often overlooked, which prevents practical asset management from obtaining tangible benefits from research. This research provides an innovative approach in narrowing the overlooked research gap and brings research results within the reach of practitioners.

**Keywords** Asset management, Railway, Track geometry, Deterioration, Framework, Maturity model

**Paper type** Research paper

## 1. Introduction

The majority of rail infrastructure funding in the EU gets spent on maintenance and renewals (M&R) (Commission, 2021). This is because railway tracks endure strenuous loading and harsh weather conditions in daily operations, resulting in structural deterioration. Furthermore, as railways are safety critical infrastructure with regard to, for instance, high-speed passenger traffic and hazardous cargo, their safety needs to be closely monitored. The primary means of monitoring the condition and safety of railway tracks include conducting track geometry measurements using a specific track recording car. The track recoding car measures the relative position of the rails, thus providing detailed information on the condition of the tracks and the safety of operations. Recently, track recording car measurements have become a source of

---

The authors would like to thank the Finnish Transport Infrastructure Agency for the valued co-operation.

*Funding:* This research was funded by the Finnish Transport Infrastructure Agency ROMHA program and the Tampere University doctoral school.

Built Environment Project and Asset Management
Vol. 12 No. 6, 2022
pp. 871-886
Emerald Publishing Limited
2044-124X
DOI 10.1108/BEPAM-04-2022-0058

increasing research interest. The literature reviews by Higgins and Liu (2018) and Soleimanmeigouni *et al.* (2018b) present the growing amount of research published on the topic. One driving force for the interest in track recording car measurements is the advances in data analytics. Novel data analytics method and software are incorporated across all industries, railways being no exception. The data from track recording cars are usually time series data from several years with high accuracy and decent measurement alignment, making the data attractive for further analysis. Track recording car data analyses have been used, for instance, to analyse the effectiveness of maintenance (Soleimanmeigouni *et al.*, 2018a), predict unplanned maintenance needs (Andrade and Teixeira, 2014) and investigate root causes of problematic track deterioration behaviour (Sauni *et al.*, 2020). All this information is vital to successful asset management, for example, in selecting the timing and means of M&R.

However, it is exactly here, in the implementation of research results into practical asset management, where the development lags. In some organisations, practical asset management revolves more around the personnel's expertise and experience rather than on a systemic process. Systemic refers to a documented data-based process in this context. The problem is that if practical asset management does not utilize track geometry deterioration modelling, the maintenance actions may be timed poorly or have little impact, which will lead to repetitive and inefficient M&R. Therefore, it is important to investigate the maturity of current practices on track geometry deterioration analyses (TGDA) and form a framework tailored for advancing them. With a controlled process and a documented framework for advancing TGDA, M&R can be allocated more efficiently in the future. This type of controlled process development is made possible by applying maturity models (Albliwi *et al.*, 2014; Helgesson *et al.*, 2012). However, there are no currently available maturity models for TGDA process improvement, and the available generic maturity models are typically too general for this specific task, as they are created for organization-wide development.

The aim of this study was, therefore, to create a framework for implementing TGDA development in railway asset management. The study was divided into three goals:

1. Adapt a maturity model for advancing TGDA
2. Investigate railway asset managers' maturity level in TGDA
3. Provide a tangible framework with which railway asset managers can advance their maturity in TGDA

Consequently, the study was conducted in three parts (Figure 1). First, a generic maturity model for track geometry management was developed according to literature on asset management maturity models. Following this, semi-structured interviews of track asset management professionals were conducted to define the current maturity level. Finally, workshops were held with track maintenance and asset management professionals to create concrete steps for incrementally advancing the maturity of track geometry management. The study was done in the context of Finnish railway asset management, and Finland was used as a case example of implementation in parts 2 and 3 of this study. The main contribution of this paper is the framework for advancing the maturity level of TGDA in railway asset management. This study also provides a means for determining the maturity level of TGDA in a railway asset management organization. The rest of this paper is organized as follows. First, the background of Finnish rail network ownership and management is elaborated to bring context to the case examples. Also, the background on TGDA is elaborated. Second, the three-part process is presented. Finally, findings and conclusion are provided.

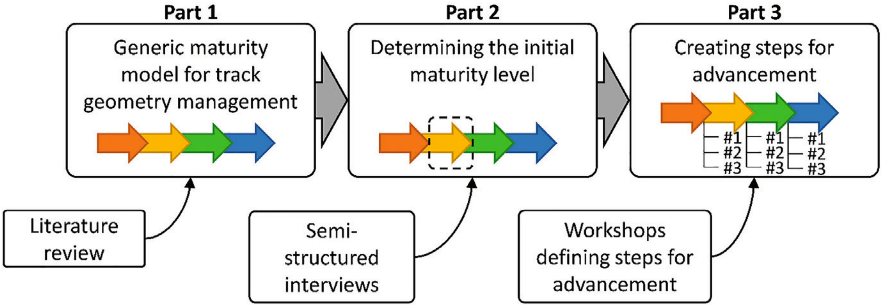


**Figure 1.**
Research methodology

## 2. Background

### 2.1 Finnish rail network ownership and management

The reader must consider that this study was done in the context of Finnish rail network ownership and management. Even though the framework created in this study is meant to be generic, the underlying organizational arrangements inevitably affect the way the framework is used. Therefore, this section is dedicated to elaborating the basic structure of Finnish rail network ownership and management.

The state-owned rail network in Finland is around 6,000 km long. The track network is mostly single track, and around half the length of the network is electrified. The maximum axle weight is primarily 22.5 tonnes; however, some lines have 20 or 25 tonne maximum axle weights. The maximum speed for passenger trains is 220 and 120 km/h for cargo trains, but most of the network is limited to slower speeds than the maximum. The division of responsibilities for managing the Finnish state-owned rail network is presented in Figure 2. Management of the state-owned rail network is run by the Finnish Transport Infrastructure Agency (FTIA) or Väylävirasto in Finnish. The FTIA is steered by the Ministry of transport and communications (LVM), which is a branch of the Finnish government. The permits to run the rail network are controlled by the Finnish Transport and Communications Agency, Traficom. The FTIA's role is to be the infrastructure owner and organize transportation on the network in accordance with LVM steering while satisfying Traficom's requirements.

The FTIA outsources its daily track management and M&R to private companies. The FTIA sets the guidelines on which the operations are based. The FTIA also tenders and supervises the contracts for track management, track maintenance and track geometry measurements. Track managers are private consultant companies who are responsible for managing and supervising daily M&R. Daily M&R is conducted by private rail construction and maintenance companies. Track geometry measurements are conducted by a private company. Periodical measurements are performed using one track recording car for the whole network. A new contract for the track recording car was tendered in 2016, and in 2021, the new track recording car started commercial operation. Therefore, analysing the track geometry measurement results is very topical in Finland, as new policies and practices are being formed.

### 2.2 Track geometry deterioration analysis

Track geometry describes the position and location of the rails. Track geometry can be measured using either absolute or relative measurements. Absolute measurements are generally performed using a total station or a GPS measurement device to provide coordinates for the rails in a specified coordinate system. Relative measurements, on the other hand, provide measurement data about deviations from an ideal geometry, thus describing

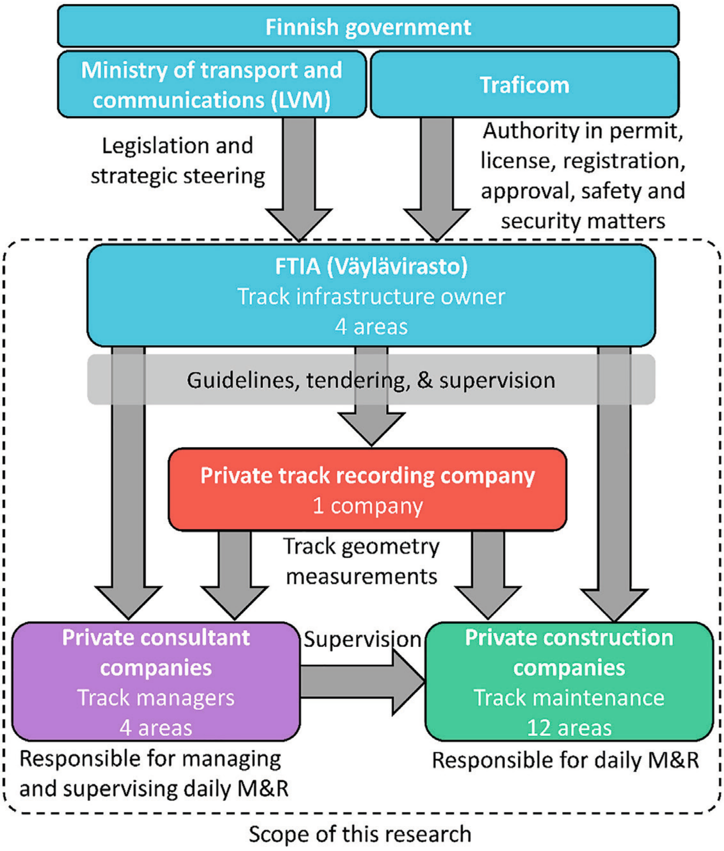


**Figure 2.**
Division of responsibilities in Finnish rail network ownership and management

the smoothness of the track geometry. Relative measurements are usually performed using a track recording car. Relative track geometry measurements are predominantly used for statistical TGDA because they are conducted recurrently, continuously, and they offer information about the quality of the structures rather than positioning information. Therefore, analyses based only on relative track geometry measurements are discussed from hereon. It is also worth mentioning that instead of using track geometry measurements, track geometry deterioration could also be approached with mechanistic models. However, their use is not regarded in this study, as they are better suited for individual structure resilience analyses rather than complete sections of track with varying structure types (Elkhoury *et al.*, 2018).

The basic use cases for (relative) track geometry measurements are the inspection of safety and quality of the tracks. If the measurements reveal deviations exceeding a safety threshold, traffic is restricted and immediate maintenance actions are taken (Figure 3a). If the measurements reveal only poor quality, but the safety limits are not exceeded, maintenance is planned to be conducted soon, but not immediately. The limit values for safety and maintenance limits are presented in the international standard series EN-13848, and they are usually specified further in national guidelines.

More advanced use of the track geometry measurements includes collecting data from several measurements and forming time series data. Typically, the standard deviation (SD) of

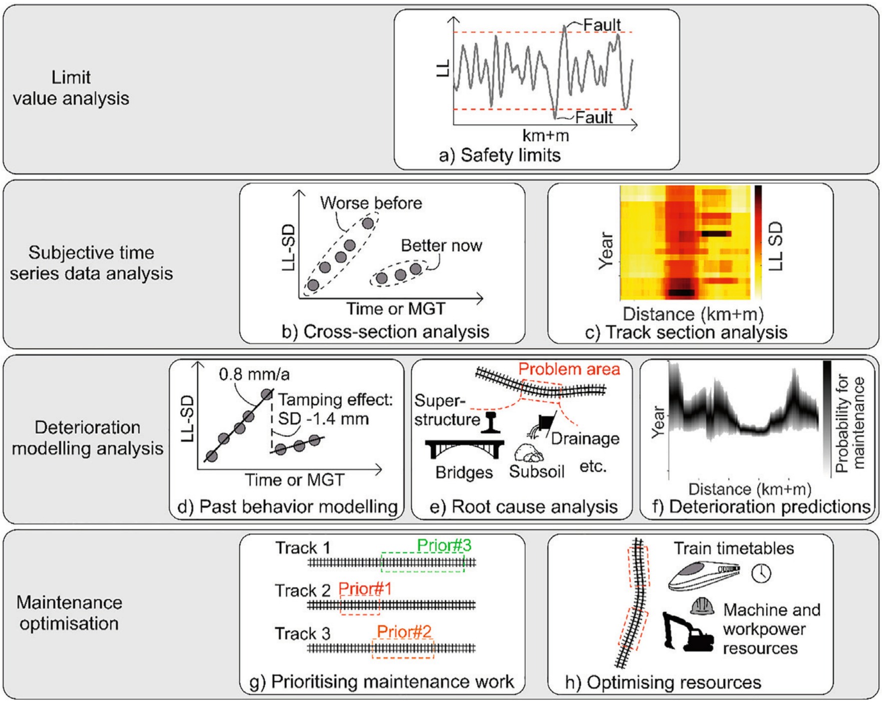


**Figure 3.**
Summary of different types of track geometry deterioration analyses

the vertical track geometry measurement signal, or in different terminology, the longitudinal level (LL) is used in the time series data analyses (Higgins and Liu, 2018). The SD provides a smooth parameter, which is easy to interpret and align amongst different measurements. The LL SD values can be plotted and examined manually, and interesting trends can be observed with suitable tools, for example, individual cross section measurement histories or heatmaps (Figure 3b and c). However, for more detailed analyses, track geometry behaviour is modelled using some mathematical idealization. The basic modelling methods include linear, exponential and logarithmic models (Neuhold *et al.*, 2020). In these approaches, the maintenance intervals need to be defined first, either from the maintenance history or by evaluating decreases in the LL SD (Sauni *et al.*, 2022). Additionally, some more complex models, such as stochastic and probabilistic models, have also been used to model the behaviour of track geometry (Elkhoury *et al.*, 2018; Higgins and Liu, 2018). Regardless of the model used, the result of track geometry deterioration modelling is generally a numerical description of the track geometry deterioration behaviour. The numerical values representing the behaviour can be used to compare the track geometry deterioration rate of different areas and time periods, or to evaluate past maintenance effectiveness (Figure 3d). This information can be combined with other asset data to investigate the root causes for track geometry issues (Figure 3e) (Sauni *et al.*, 2022). Track geometry deterioration modelling can be used to predict future behaviour based on the track geometry deterioration history (Figure 3f) (Sauni *et al.*, 2022). The predictions can be used to prioritize maintenance and optimise resources before safety limits are exceeded (Figure 3g and h). Prioritizing maintenance based on track geometry deterioration modelling is necessary, as there are usually more repair needs than there are available funding. After the maintenance needs have been prioritized, maintenance resources (track work machines and timetables) can be optimized. Maintenance prioritization and resource optimization can be combined in some organizations, if the maintenance is conducted by the track owner, see for example (Bressi *et al.*, 2021). Otherwise, if a company, responsible only for track maintenance but not track ownership, performs the prioritization and resource optimization, they can emphasize resource optimization to ease their work schedule, instead of focusing on which segments of the track require the most immediate attention. All the analyses mentioned above are summarized in Figure 3.

### *2.3 Maturity models*

Maturity models provide a good basis for controlled incremental development, as their background is in managing large software projects (Paulk *et al.*, 1993). The capability maturity model (CMM) can be considered one of the original maturity models, from which many variants have been developed, each with their own characteristics (Albliwi *et al.*, 2014; CMMI Product Team, 2010; Helgesson *et al.*, 2012; Paulk *et al.*, 1993; Poeppelbuss *et al.*, 2011). These maturity models have been used, for instance, in setting future goals for development with high success (Herbsleb and Goldenson, 1995).

Nonetheless, the models are not without criticism. One major critique is that the readymade models, such as the CMM (Paulk *et al.*, 1993), do not cover every aspect of an organization (Albliwi *et al.*, 2014). Furthermore, Poeppelbuss *et al.* (2011) present three general challenges associated with using maturity models: (1) vastness of theoretical research, (2) empirical assessment of maturity levels and (3) the lack of one linear sequence for development in practice.

In this study, the comprehensiveness of maturity models is not as important as their adaptability in defining the maturity levels within these models. This is due to the research focusing on a clearly defined process, TGDA, rather than a whole organization. As for the general challenges associated with maturity models, this study applies only the principles of past maturity models, which eliminates the need for a readymade model (challenge 1). The assessment of maturity levels was based on the interviews and workshops held with relevant stakeholders which provided a comprehensive assessment (challenge 2). Finally, the end results, a framework for TGDA development, will be based on a maturity model, but the process will not be strictly linear (challenge 3). Rather, the process will describe the order of the steps required for advancing TGDA.

The process of creating a maturity model has been reported in previous studies (de Bruin *et al.*, 2005; Maier *et al.*, 2012). This includes, at least, phases for planning, developing, evaluating and maintaining the model (Maier *et al.*, 2012). Planning and developing a maturity model require vast domain knowledge that must be obtained from industry experts, by conducting surveys, interviews and workshops (Maier *et al.*, 2012). Model evaluation can be done in different ways: evaluation by the model authors, evaluation involving the industry experts and evaluation through practical case-use (Helgesson *et al.*, 2012). The created model must also be maintained by re-evaluating the current maturity level and revising goals.

Similar to the current study, the supporting ideals of maturity models have been applied to different applications in previous research, for example, building information models (Eadie *et al.*, 2015) and railway cybersecurity (Kour *et al.*, 2019). Maturity models have even been integrated into railway operations in the International union of railways (UIC) application guide for asset management (UIC, 2016). However, the maturity model reported by the UIC covers the overall maturity of an entire asset management organization which is too general a starting point for specific process development, such as TGDA. Therefore, this study applied and modified the established maturity model as the basis for TGDA development, as is reported in section 3 of this paper, to fill this gap in research.

The principal justification for using a maturity model as the basis for TGDA development was the possibility of assessing current maturity and setting intermediate goals for tracking development. These strengths have been observed in previous research on the topic (Hirose *et al.*, 2020). Incremental evolution rather than sudden revolution is preferred also in this case because the development concerns an ongoing safety critical process, which cannot be disturbed. More specifically, track geometry measurements are used to determine whether it is safe to conduct rail traffic. If this process is seized, or the results interpretation is disturbed, this might result in unacceptable track irregularities going unnoticed, which can cause train derailments. Furthermore, incremental development helps to form logical progress for the development as the next maturity level should not be pursued until the conditions of the current maturity level are satisfied. This prevents, for example, implementing elaborate deterioration models before data production and pre-processing are in order. Finally, it should be noted that the primary objective of this research is not only to create a maturity model but also to develop the competence of an asset management organization. To achieve this result, maturity models are utilized as the vehicle for implementing development.

## 3. Research process

### *3.1 Part 1: adapting a generic maturity model for TGDA development*

In this study, the maturity model from UIC (2016) was applied to track geometry deterioration management. The initial version of the TGDA specific maturity model was created by the authors of this paper based on the different types of TGDA (*cf.* Figure 3). The further development and evaluation of the model were performed with industry stakeholders in the workshops reported in part 3 of this research. A consensus over the contents of the maturity model and the respective framework was reached during the workshops.

The generic maturity model was modified into a four-level model with the following levels (Figure 4). The first level is ensuring traffic safety, as chaotic track geometry management is not an option for a responsible asset manager. Ensuring safety means periodically measuring track geometry to reveal locations with deviations exceeding safety thresholds and requiring immediate maintenance. The next level is monitoring track quality, which includes, for instance, collecting measurement time series data in a database. These data can be analysed subjectively to reveal areas with recurrent problems and progressive deterioration. The third level, track geometry management, includes connecting other asset management systems and data to track geometry measurement databases and modelling track geometry deterioration. With these advancements, for example, the root causes of track geometry anomalies can be investigated. The last level, optimizing track geometry, contains optimizing and prioritizing maintenance according to available maintenance resources, track repair time and track class, for instance. The excellent level was not considered because excellence can be defined as having fully optimized maintenance.

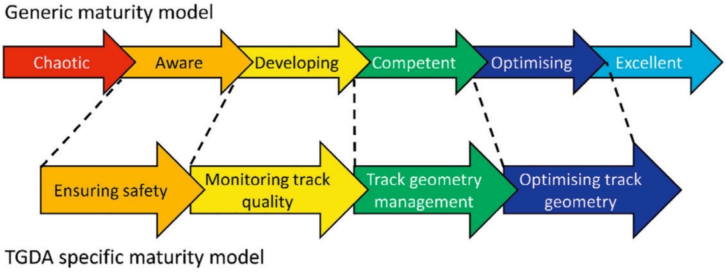


**Figure 4.**
A generic maturity model applied to TGDA

The TGDA specific maturity model was used as the basis for investigating the maturity level in Finland (part 2 of this study) and for determining a detailed framework for future development (part 3 of this study). Detailed contents for each level in the maturity model were researched in part 3 of this study.

### *3.2 Part 2: investigating the maturity level in TGDA*

The aim of this part of the study was to find a suitable way of assessing the current maturity level in TGDA. The current maturity level must be investigated first because it is pertinent to create the framework based on actual needs from the industry, as the framework is to be implemented in practice. Thus, the current processes and development needs for TGDA must be investigated by interviewing experts in the field. In the case of Finland, the interviewees included experts from all private companies that had either track M&R or track management contracts with the FTIA and FTIA's own personnel. Most interviews were group interviews comprise experts from the same organization. The interviewees were a representative sample of Finnish railway asset management as all track management areas and organizations were represented. The interviewees included:

- 5 track maintenance experts from 3 track construction companies
- 12 asset management experts from 4 track asset management companies
- 5 track inspection and maintenance experts from the infrastructure owner

The interviewees were highly experienced with 18 years of experience from the railway sector on average.

The interviews were conducted as semi-structured interviews. The rationale behind choosing semi-structured interviews as the mode of surveying included:

- Low number of interviewees, $n = 22$
- Exploratory nature of the interviews
- Possibility of group interviews

The low number of interviewees was due to the limited number of people working closely with track geometry data in Finland. Furthermore, the interviews were exploratory as there was little written about current practices in Finland. Also, many participants wished to be interviewed in groups along with colleagues from their organization to allow for colleagues to supplement their answers. Semi-structured interviews allowed for taking all these into consideration while still having some control on the topics that were discussed in the interviews.

All interviews followed the same format. The interviews were segmented into three themes with relevant subquestions. The subquestions were used to generate discussion and to guide conversation if needed, but the participants were free to answer as they pleased, and follow-up questions not belonging to the standard form were presented as conversations diverged. The themes and questions were identical for every interviewee regardless of their position or organization. The basic structure of the interviews is presented in Table 1. The interviews were conducted, reported and analysed in Finnish, but the form and conclusions were translated into English for this paper.

The interview structure had a larger number of simpler questions introduced first to get the interviewees talking and relaxed about answering. Later in the interview, the questions were more open ended and there were fewer of them to allow the interviewee to answer in greater length, and possibly even wander off topic. The purpose of theme 1 and its subquestions was to investigate the current use of track geometry measurement results and

| Theme | Sub questions |
|---|---|
| T1: Current use of track geometry measurement results | Q1: For what purpose do you use track geometry measurement results in your line of work, and what information do you require from them?<br>Q2: Who handles track geometry measurement results in your organization, and are there differences between the use-cases of different personnel within your organization?<br>Q3: Which guidelines do you follow in analysing track geometry measurement results, and what other guidance do you know of related to the topic?<br>Q4: Are there deficiencies in the guidelines related to your use?<br>Q5: What procedures, related to your work, are conducted/ordered in different circumstances according to track geometry measurement results?<br>Q6: Do you use track geometry measurement results for some other purposes besides analysing the condition of the track, for example, contractual purposes or work planning? |
| T2: Procedures for analysing track geometry measurement results | Q7: Do you refine the track geometry measurement results (e.g., with statistics, models or key figures) in addition to the results provided to you?<br>Q8: Do you know of some methods for refining the measurement results that would be suitable for your use but are not currently in use?<br>Q9: What other sources of information do you use when analysing track geometry measurement results (e.g., plans, maps, photos and reports)?<br>Q10: What other sources of information would you require to aid track geometry measurement result analysis, but they are not currently available? |
| T3: The potential of track geometry measurement result analysis | Q11: What could be achieved by analysing track geometry measurement results if current problems did not exist?<br>Q12: What do you wish to be changed in the processes of analysing track geometry measurement results?<br>Q13: What directions for future development do you know of, or would hope to see, regarding track geometry measurement analysis? |

**Table 1.**
The structure of the interviews

describe the general process of handling the data in the organization. This theme was especially important because the framework was to be built upon current practices, which were unknown beforehand. Theme 2 discussed the analysis of track geometry measurement results in a narrower focus. Special attention was paid to the further handling and refining of the results by the interviewees. The ways interviewee have had to work around and complement the current processes would tell a lot about what deficiencies current processes have. Questions 6, 7 and 9 had examples within them, which could be interpreted to be leading the interviewee on. However, this was a deliberate choice to have the examples presented with the questions, as an expert on the matter, that can consider the examples to be self-evident and not mention them otherwise. Theme 3 focused on getting the interviewee to reflect on what the limitations to the current analyses really are, what could be done to change them and what would be the effect.

### *3.3 Part 3: creating a framework based on the TGDA maturity model*

Once the maturity model and current maturity level were investigated, it was time to create a framework for advancing TGDA development. The framework was designed in a set of three workshops. The topics of workshops were (1) knowledge areas, (2) development paths and (3) implementation plan. Knowledge areas refer to the categories which form the structure of the framework. Development paths refer to the tangible contents of maturity levels. The workshops were held with 2-month intervals in the winter of 2021–2022. The contents of each workshop regarding the framework are shown in Figure 5.

The goal of the first workshop was to determine the knowledge areas that will structure the framework. The workshop was held online on Teams, and the group work was done on

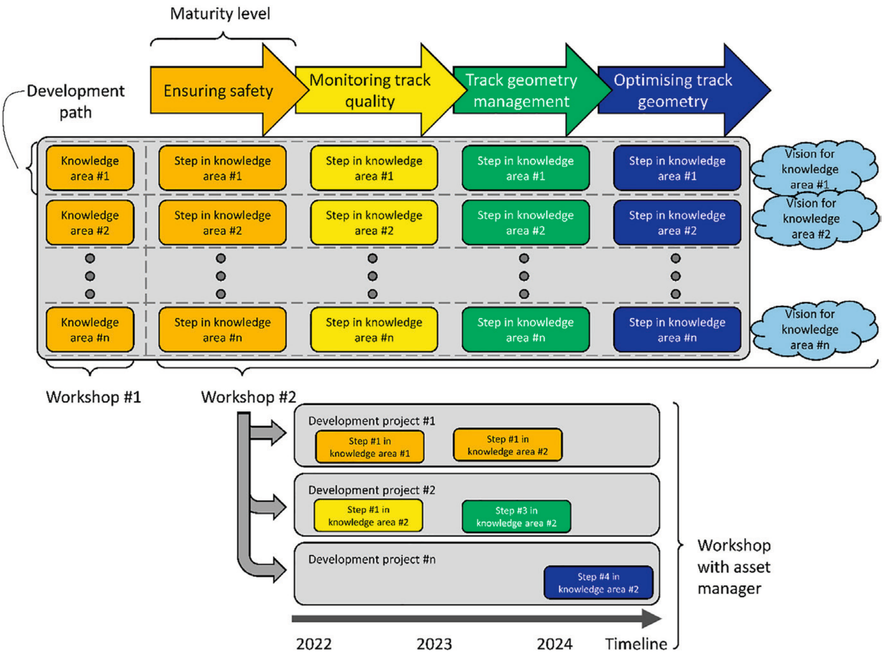


**Figure 5.**
The design of the workshops

the whiteboard application Flinga, which could be operated freely by any participant. There were 23 participants and four organizers who were divided into groups of 3–4 people, each group with their own whiteboard. The participants were divided into groups based on their affiliation so that infrastructure owners, asset managers and maintenance personnel were mixed and represented as diversely as possible in different groups. The participants were first asked to come up with possible knowledge areas by answering a supporting question: “What areas or processes are affected by or connected to track geometry measurements in your line of work?” From here on, the participants created mind maps of the most essential knowledge areas and operations related to them. These mind maps were the result of the first workshop. The mind maps were later analysed using ATLAS.ti to identify the most frequently mentioned topics. Overall, 323 observations in 65 codes and 8 code groups were created. From these codes and code groups, six knowledge areas were created to be further developed in the second workshop.

The goal of the second workshop was to create development paths to the six knowledge areas obtained from the first workshop’s results. The knowledge areas were presented to the participants along with preliminary visions for the future of said areas. The participants were divided into six groups, and each group was given one knowledge area. The first task was to challenge and supplement the given preliminary visions. After this, blank four-level maturity models, as described in part 1 of this study, were given, and participants were asked to fill in the models with concrete actions for each stage. Then, the groups were rotated twice so that they could comment and supplement the previous groups’ work.

Before the third workshop, an implementation plan was developed in cooperation with the asset manager. The implementation plan included placing the steps from the framework on a timeline within a relevant process, whether it be a development project, contract or guideline. In this way, the contents of the framework could be implemented concretely as the next development milestones for said processes. The third workshop concentrated on commenting and supplementing the framework and implementation plan. In addition to getting much valued feedback on the framework and implementation plan, the final workshop played a role in presenting the results and engaging different organisations in the forthcoming development.

## 4. Results and discussion

In part 1 of this study, a maturity model was adapted for incremental advancement of TGDA in railway asset management. The developed four-level maturity model functions as the basis, on which the framework was built in part 3 of this study. The new knowledge produced concerned the application of maturity models into a novel domain, TGDA in railway asset management. The purpose of this maturity model was only to create the structure for the developed framework, not to be tested and validated as a stand-alone maturity model, as is common with applied maturity models (Helgesson *et al.*, 2012). The validity of the maturity model and consensus over the contents of the model were verified during the industry stakeholder workshops.

The interviews, held in part 2 of this study, were successful in determining the current maturity of TGDA in Finland. The exploratory nature of the semi-structured interview provided a systematic way of collecting data while enabling leeway for the interviewees' answers. The interviews revealed different use-cases and user types, which helped in designing and supplementing the framework. This novel information was utilized in constructing the initial framework for part 3 of this study.

As a conclusion from the interviews, the maturity level of TGDA was primarily at the *monitoring track quality* level (Figure 6). In Figure 6, the observations in the green area indicate that the level is satisfied to that degree. The observations in the red area indicate that the level is not satisfied with regard to the comments. Figure 6 does not depict the maturity of

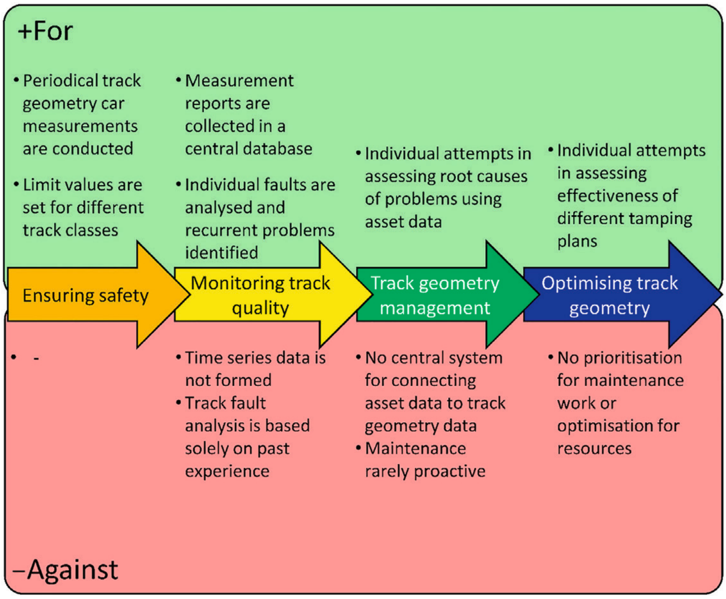


**Figure 6.**
The initial maturity level of TGDA in Finland

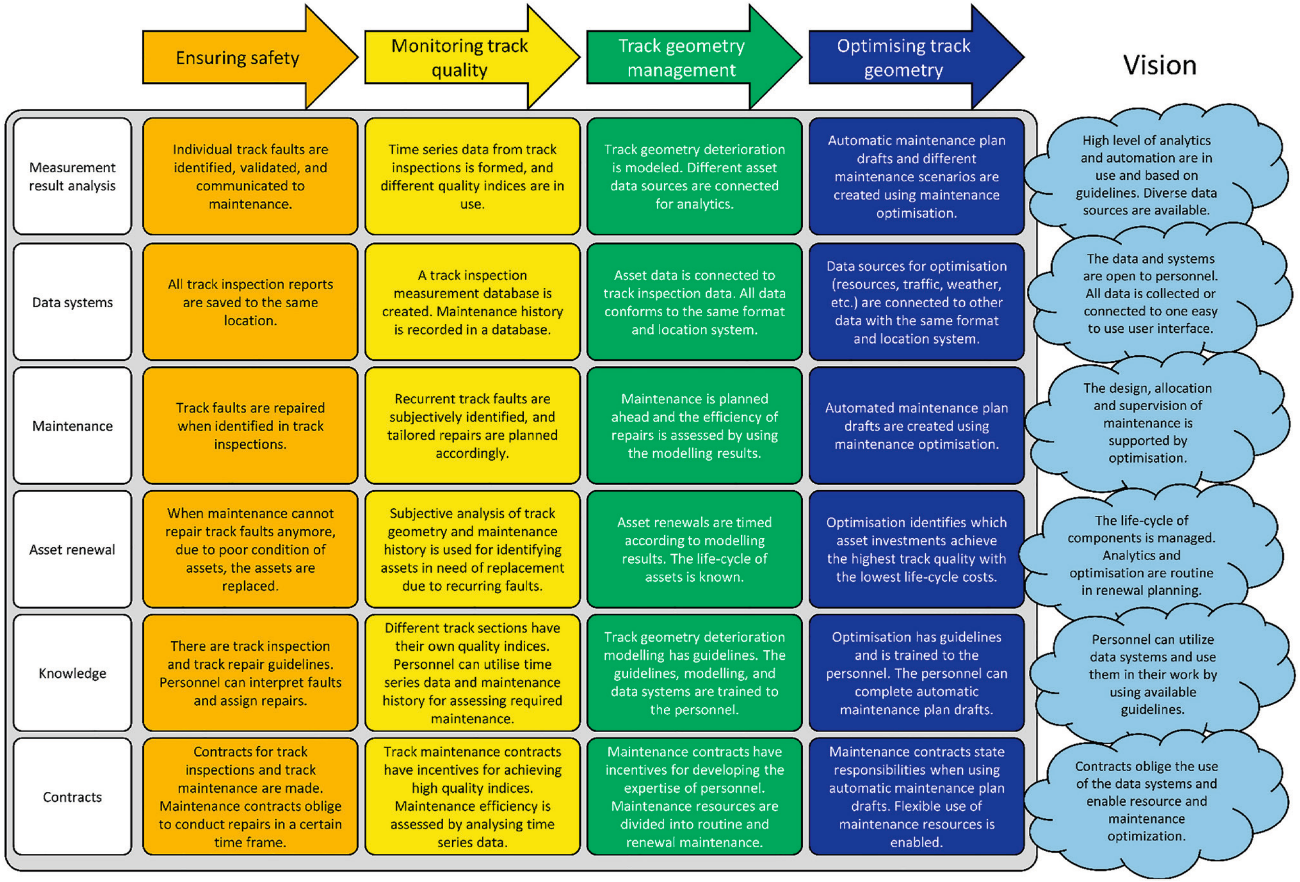
asset management throughout the asset management organization but only the maturity of the TGDA. There were some observations of more developed analyses made by individuals on their own. The individual more developed analyses included assessing what quality levels could be achieved with different tamping plans and connecting asset data to recurring problem areas. However, overall, current practices in the industry are more focused on the *monitoring track quality* maturity level. The focus in TGDA has been in identifying problem areas on track sections and to plan their remediation. For example, numerical track geometry measurement data are not shared, there are no methods in use for modelling track geometry deterioration, and asset data are not generally connected to track geometry measurement data.

The end-result of the workshops was the framework for advancing TGDA in Finland (Figure 7). The framework included six development paths: Measurement result analysis (1) referred to the process after measurement when visualizations and analytics are produced for the user. Data systems (2) covers the software needed to store data and create the analytics. Maintenance (3) refers to designing, conducting and supervising the maintenance actions included in the current maintenance contracts, whereas asset renewals (4) indicate repairs and investments not included in current maintenance contracts (e.g. large-scale track renewals). Knowledge (5) includes the ability to utilize the results as well as the required guidelines and training. Lastly, there are the contracts (6) needed to acquire the services required to achieve a certain maturity level. Each development path has a vision depicting the ultimate goal of said path and four maturity levels presenting incremental steps in progressing towards the vision. The maturity levels increase to the right. A maturity level contains all the requirements from previous levels, thus making the development cumulative. Before advancing to the next maturity level, all development paths should satisfy the current maturity level. In this way, the development is incremental and builds upon implemented practices, which is important in creating tangible progress during development. Also, as track geometry condition monitoring is a safety critical process, new features must be implemented one-by-one while making sure the current process is not disturbed.

The framework is a tool for advancing TGDA in railway asset management. The framework enables examining the development in the present and distant future at the same time, all the while maintaining focus on the correct order for the development. This is achieved by following the development paths in the framework. This was found very helpful when turning visions into actual development projects.

The workshops yielded valuable information to supplement the framework. If the framework was created using only available literature, many pragmatic aspects would have been overlooked. For example, the diversity of different TGDA user types and use cases would not have been uncovered without the workshops. Additionally, the workshops engaged the stakeholders in the development. As the framework was developed in the workshops with the stakeholders, the development was transparent, and the stakeholders had an influence on the framework. This is believed to reduce resistance to change and provide community support for the succeeding development projects. Similar observations regarding the benefits of workshops have been made in previous research (Ørngreen and Levinsen, 2017; Phaal *et al.*, 2007).

The limitations of this framework concern the influences of the Finnish railway operating environment on the study. The perspective in the framework was a buyer–supplier model, in which the infrastructure owner acts as the buyer who has the responsibility for implementing the development. Additionally, the participants were solely from Finland, limiting the different operating environment experiences obtained in the interviews and workshops. Nevertheless, the steps within the framework were designed to be universally applicable, but the global validation of the framework was left as a source of future research. It is also worth noting that the steps within the framework are not equal in effort. Therefore, the



| | Ensuring safety | Monitoring track quality | Track geometry management | Optimising track geometry | Vision |
|---|---|---|---|---|---|
| Measurement result analysis | Individual track faults are identified, validated, and communicated to maintenance. | Time series data from track inspections is formed, and different quality indices are in use. | Track geometry deterioration is modeled. Different asset data sources are connected for analytics. | Automatic maintenance plan drafts and different maintenance scenarios are created using maintenance optimisation. | High level of analytics and automation are in use and based on guidelines. Diverse data sources are available. |
| Data systems | All track inspection reports are saved to the same location. | A track inspection measurement database is created. Maintenance history is recorded in a database. | Asset data is connected to track inspection data. All data conforms to the same format and location system. | Data sources for optimisation (resources, traffic, weather, etc.) are connected to other data with the same format and location system. | The data and systems are open to personnel. All data is collected or connected to one easy to use user interface. |
| Maintenance | Track faults are repaired when identified in track inspections. | Recurrent track faults are subjectively identified, and tailored repairs are planned accordingly. | Maintenance is planned ahead and the efficiency of repairs is assessed by using the modelling results. | Automated maintenance plan drafts are created using maintenance optimisation. | The design, allocation and supervision of maintenance is supported by optimisation. |
| Asset renewal | When maintenance cannot repair track faults anymore, due to poor condition of assets, the assets are replaced. | Subjective analysis of track geometry and maintenance history is used for identifying assets in need of replacement due to recurring faults. | Asset renewals are timed according to modelling results. The life-cycle of assets is known. | Optimisation identifies which asset investments achieve the highest track quality with the lowest life-cycle costs. | The life-cycle of components is managed. Analytics and optimisation are routine in renewal planning. |
| Knowledge | There are track inspection and track repair guidelines. Personnel can interpret faults and assign repairs. | Different track sections have their own quality indices. Personnel can utilise time series data and maintenance history for assessing required maintenance. | Track geometry deterioration modelling has guidelines. The guidelines, modelling, and data systems are trained to the personnel. | Optimisation has guidelines and is trained to the personnel. The personnel can complete automatic maintenance plan drafts. | Personnel can utilize data systems and use them in their work by using available guidelines. |
| Contracts | Contracts for track inspections and track maintenance are made. Maintenance contracts oblige to conduct repairs in a certain time frame. | Track maintenance contracts have incentives for achieving high quality indices. Maintenance efficiency is assessed by analysing time series data. | Maintenance contracts have incentives for developing the expertise of personnel. Maintenance resources are divided into routine and renewal maintenance. | Maintenance contracts state responsibilities when using automatic maintenance plan drafts. Flexible use of maintenance resources is enabled. | Contracts oblige the use of the data systems and enable resource and maintenance optimization. |

**Figure 7.**
Framework for advancing TGDA

implementation of a step must be individually planned, as one step may require years of development, whereas others only slight amendments to guidelines. The framework should not be seen as a development project but as a tool for turning a vision into a series of development projects.

The practical implications of the TGDA development framework include improvements to the way data are utilized in safety and condition monitoring in railway asset management. Currently, much of the data are subjectively assessed, which creates opportunities for human error. With advanced TGDA, human errors in safety and condition monitoring can be avoided with, for example, automatic alerts and predictive analytics. Furthermore, maintenance can be planned more efficiently, thus reducing costs by eliminating redundant maintenance. These benefits are obtainable by any asset management organization that increases their capabilities in data analytics.

## 5. Conclusion

In recent years, research on TGDA has evolved greatly, and novel information on the condition of railway tracks can be produced to streamline the use of maintenance resources. However, the implementation of TGDA into railway asset management is lagging due to the complexity of altering ongoing safety critical processes. Therefore, the implementation of TGDA requires in-depth research to narrow the gap between research and practice to obtain tangible societal benefits from previous research.

In this study, a framework for implementing TGDA into railway asset management was developed. The framework was developed, tested and applied in the Finnish state rail network asset management. The framework was established in three parts: (1) a maturity model was adapted as the basis for the framework, (2) semi-structured interviews were conducted to evaluate the current maturity level and (3) workshops were held to construct the detailed content of the framework.

The main contribution of this study is the novel framework presented in Figure 7. When an asset manager identifies their placement within the framework and applies the framework into designing their development projects, they can create a vision that can be reached with incremental development. This is especially useful when the asset manager wants to create a long-term strategy for TGDA development, while keeping the implementation of development highly practical. Furthermore, the incremental and cumulative progress achieved with using the framework is much easier to communicate to stakeholders and implement than abrupt revolution.

The practical implication of this study is the possibility for an asset manager to advance their TGDA, thus improving the efficiency of condition monitoring, which reduces safety risks and maintenance costs. The framework was successfully tested and applied in Finnish state rail network asset management; The current maturity level in TGDA was identified and development paths were tailored. However, the limitation of this study was that the framework was validated only in the workshops in a Finnish operating environment, even though the framework was designed to be universally applicable. The validation of the framework into different operating environments is a source of future research. Further research on the topic could also include different asset management processes, for example, the life-cycle management of track components.

## References

Albliwi, S.A., Antony, J. and Arshed, N. (2014), “Critical literature review on maturity models for business process excellence”, *2014 IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 79-83.

Andrade, A.R. and Teixeira, P.F. (2014), "Unplanned-maintenance needs related to rail track geometry", *Proceedings of the Institution of Civil Engineers - Transport*, ICE Publishing, Vol. 167 No. 6, pp. 400-410.

Bressi, S., Santos, J. and Losa, M. (2021), "Optimization of maintenance strategies for railway track-bed considering probabilistic degradation models and different reliability levels", *Reliability Engineering and System Safety*, Vol. 207, 107359.

CMMI Product Team (2010), *CMMI for Development, Version 1.3*, Carnegie Mellon University, Software Engineering Institute.

Commission, E. (2021), "Seventh monitoring report on the development of the rail market under article 15(4) of directive 2012/34/EU of the European parliament and of the council, European commission", Brussels, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2021:0005:FIN

de Bruin, T., Rosemann, M., Freeze, R. and Kulkarni, U. (2005), "Understanding the main phases of developing a maturity assessment model", *ACIS 2005 Proceedings - 16th Australasian Conference on Information Systems*, No. December.

Eadie, R., Browne, M., Odeyinka, H., McKeown, C., McNiff, S., Mc Keown, C. and Mc Niff, S. (2015), "A survey of current status of and perceived changes required for BIM adoption in the UK", *Built Environment Project and Asset Management*, Emerald Group Publishing, Vol. 5 No. 1, pp. 4-21.

Elkhoury, N., Hitihamillage, L., Moridpour, S. and Robert, D. (2018), "Degradation prediction of rail tracks: a review of the existing literatur", *The Open Transportation Journal*, Vol. 12, pp. 88-104.

Helgesson, Y.Y.L., Höst, M. and Weyns, K. (2012), "A review of methods for evaluation of maturity models for process improvement", *Journal of Software: Evolution and Process*, John Wiley & Sons, Vol. 24 No. 4, pp. 436-454.

Herbsleb, J.D. and Goldenson, D.R. (1995), "Systematic survey of CMM experience and results", *Proceedings - International Conference on Software Engineering*, pp. 323-330.

Higgins, C. and Liu, X. (2018), "Modeling of track geometry degradation and decisions on safety and maintenance: a literature review and possible future research directions", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 5, pp. 1385-1397.

Hirose, Y., Phaal, R., Farrukh, C., Gerdsri, N. and Lee, S. (2020), "Characterizing maturity levels for organization-wide roadmapping implementation", *IEEE Engineering Management Review*, Vol. 48 No. 4, pp. 133-143.

Kour, R., Karim, R. and Thaduri, A. (2019), "Cybersecurity for railways – a maturity model", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 234 No. 10, pp. 1129-1148.

Maier, A.M., Moultrie, J. and Clarkson, P.J. (2012), "Assessing organizational capabilities: reviewing and guiding the development of maturity grids", *IEEE Transactions on Engineering Management*, IEEE, Vol. 59 No. 1, pp. 138-159.

Neuhold, J., Vidovic, I. and Marschnig, S. (2020), "Preparing track geometry data for automated maintenance planning", *Journal of Transportation Engineering, Part A: Systems*, American Society of Civil Engineers, Vol. 146 No. 5, 4020032.

Ørngreen, R. and Levinsen, K. (2017), "Workshops as a research methodology", *Electronic Journal of E-Learning*, Vol. 15 No. 1, pp. 70-81.

Paulk, M.C., Curtis, B., Chrissis, M.B. and Weber, C.V. (1993), "Capability maturity model, version 1.1", *IEEE Software*, Vol. 10 No. 4, pp. 18-27.

Phaal, R., Farrukh, C.J.P. and Probert, D.R. (2007), "Strategic roadmapping: a workshop-based approach for identifying and exploring strategic issues and opportunities", *EMJ - Engineering Management Journal*, Vol. 19 No. 1, pp. 3-12.

Poeppelbuss, J., Niehaves, B., Simons, A. and Becker, J. (2011), “Maturity models in information systems research: literature search and analysis”, *Communications of the Association for Information Systems*, Vol. 29, doi: 10.17705/1cais.02927.

Sauni, M., Luomala, H., Kolisoja, P. and Turunen, E. (2020), “Investigating root causes of railway track geometry deterioration – a data mining approach”, *Frontiers in Built Environment*, Vol. 6, p. 122.

Sauni, M., Luomala, H., Kolisoja, P. and Nummi, T. (2022), “Advancing railway asset management using track geometry deterioration modelling visualization”, *Journal of Transportation Engineering, Part A: Systems*, Vol. 148 No. 2, doi: 10.1061/JTEPBS.0000626.

Soleimanmeigouni, I., Ahmadi, A., Arasteh Khouy, I. and Letot, C. (2018a), “Evaluation of the effect of tamping on the track geometry condition: a case study”, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 2, pp. 408-420.

Soleimanmeigouni, I., Ahmadi, A. and Kumar, U. (2018b), “Track geometry degradation and maintenance modelling: a review”, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 232 No. 1, pp. 73-102.

UIC (2016), *UIC Railway Application Guide: Practical Implementation of Asset Management through ISO55001*, UIC Railway Technical Publications, available at: http://www.nvfnorden.org/library/Files/Utskott-och-tema/Drift-och-underhall/Møter-og-protokoller/UIC ISO 55000 Guidelines_final.pdf

**Corresponding author**
Mikko Sauni can be contacted at: mikko.sauni@tuni.fi