Intelligent rail maintenance decision support system using KPIs

Jamshidi, Ali

DOI
10.4233/uuid:14048e52-00ad-49e8-9964-aa14e33673fd

Publication date
2019

Document Version
Final published version

Citation (APA)
Jamshidi, A. (2019). *Intelligent rail maintenance decision support system using KPIs*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:14048e52-00ad-49e8-9964-aa14e33673fd

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Intelligent rail maintenance decision support system using KPIs

Proefschrift

Ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen,

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op 3 September 2019 om 15:00 uur

door

ALI JAMSHIDI

Master of Science in Disaster Management (Civil Engineering). Tehran University, Iran

Dit proefschrift is goedgekeurd door de:

Promotors: Prof.dr.ir. Z. Li, Prof.dr.ir. R.P.B.J. Dollevoet

Copromotor: Dr.ir. A.A. Núñez Vicencio

Samenstelling van de promotiecommissie:

Rector Magnificus	Chairperson
Prof.dr.ir. Z. Li, promotor	Technische Universiteit Delft
Prof.dr.ir. R.P.B.J. Dollevoet, promotor	Technische Universiteit Delft
Dr.ir. A. Núñez Vicencio, copromotor	Technische Universiteit Delft

Onafhankelijke leden:

Prof.dr.ir. B. De Schutter	Technische Universiteit Delft
Prof.dr. I. Skrjanc	University of Ljubljana
Prof.dr. D. Galar	Lulea University of Technology
Dr.ir. A. Zoeteman	ProRail
Prof.dr.ir. M. Veljkovic	Technische Universiteit Delft

TU Delft

Delft University of Technology

This research is part of the NWO/ProRail project (Multi-party risk management and key performance indicator design at the whole system level, PYRAMIDS), project code 438-12-300.

Keywords: Key Performance Indicators, Rail Infrastructure, Rail Surface Defects, Axle Box Acceleration, Maintenance Decision Support System

Printed by: Ipskamp Printing, Enschede

ISBN: 978-94-6384-059-0

An electronic version of this dissertation is available at

http://repository.tudelft.nl

Dedicated to my parents

Acknowledgment

My PhD research was not possible in the absence of the support of my friends/colleagues. First of all, I would like to thank my daily supervisor, Alfredo Núñez, for his valuable supervision. I am keenly grateful for his creative and scientific insights improving the research quality and suggesting me interesting ideas. I would like to thank my main supervisor as well, prof. Zili Li, for his support and technical feedback. Also, Thanks to prof. Rolf Dollevoet for all his support and feedback. Finally, I am specially thankful of prof. Bart De Schutter and Robert Degenhart, for scientific guidance of the PYRAMIDS project and detailed feedback.

Many thanks to Shahrzad, Siamak and Meysam for our very fruitful collaborations. Their inputs certainly improved my quality of my research and publications. Special thanks go to Meysam for the many lunch and tea discussions that developed in a great friendship with him.

Further, I would like to thank everyone involved in the PYRAMIDS project, particularly Zhou Su. The project was a great chance to learn about cooperation and allowed me to realize of the importance of team work. I am glad that I have been contributing track visits and nightly track measurements as it taught me a lot regarding practical railway operations and also broadened my railway knowledge, although those visits and measurements were barely related to my own research. I am happy as well that I have been involved in organizing the colloquiums of Railway section as chairman and also social events of the section.

I also thank Railway engineering section's staffs at TUDelft: Jacqueline, Jan Moraal, Jurjen, Dirk and Valeri and my many current and former PhD colleagues and Touraj Saberivand for the cover of the thesis.

Table of content

Summary

Key Performance Indicators (KPIs) enable the infrastructure manager to keep the performance quality of the infrastructure at an acceptable level. A KPI must include specific features of the infrastructure such as functionality and criticality. The KPIs can be classified into three performance levels: (1) technical level KPIs; (2) tactical level KPIs and (3) global level KPIs. For instance, some KPIs are related to individual rail components (technical level) and some correspond to a bigger picture of the rail including multiple components (tactical level). The global level also gives an overview indication of the full-length rail based on what the infrastructure manager requires. Hence, to use every KPI correctly, the infrastructure manager should be aware of the proper KPIs level.

In this dissertation, an intelligent rail maintenance decision support system using KPIs is proposed. The thesis is composed of three parts: design of KPIs, rail degradation model and condition-based maintenance decision system.

In the first part, a design procedure of the KPIs for railway tracks is proposed. Among all factors affecting the railway performance, rail surface defects play an important role in influencing the rail performance. To measure rail health condition, three different data sources are used in this study: (1) axle box acceleration (ABA) (2) video images and (3) ultrasonic. Based on the measurement data, a set of KPIs is defined at the technical level for rail surface defects. Next, the KPIs are aggregated into track segments to facilitate maintenance decisions at the tactical level. At the end, a global level KPI can be obtained according to the tactical and technical level representing the whole system performance.

In the second part, degradation modeling is used for the prediction of the rail health condition. A predictive approach is required to anticipate what could occur in the rail over a given time period. Relying on field observations, measurements and mechanical understanding of how a rail defect grows, a degradation model is proposed to predict how fast a defect can evolve. Moreover, the growth is stochastic. Some defects can grow faster than others. To capture actual dynamic growth, the KPIs are provided with a scenario-based approach. Based on the degradation model, a set of predictive and robust KPIs is then defined. Chapter 2 and Chapter 3 of this dissertation show the KPI design and the degradation model. In Chapter 2, the focus is on designing a set of fuzzy KPIs for rail infrastructures. In Chapter 3, a risk KPI is presented. The risk KPI enables the infrastructure manager to perform an analysis on failure estimation of rail. The results can be insightful for maintenance plans.

In the third part, a maintenance decision system is proposed to use the KPIs for improving the system performance and to reduce maintenance operation costs. To do so, two approaches are given in order to establish the decision support system, namely (1) expert systems and (2) Pareto-based approaches. Chapter 4 and 5 describe the proposed approaches. Chapter 4 presents a methodology based on expert systems for rail maintenance. A combination of video images and ABA signals is used to give a list of defects associated with their kilometre positions in a track in the Dutch railway network. The idea for using both images and ABA signals was to increase the accuracy of the defect observation. Beside the rail observation, track characteristics are added as a set of “influential factors” to the maintenance decision system. By having both the rail observation and the influential factors, a list of decision rules is generated.

Therefore, an expert system is designed to provide systematic maintenance decision solutions. Moreover, a non-smooth optimization problem is formulated in the chapter to cluster the defects into different rail kilometer positions. The clustering model takes time and operation cost into account. In Chapter 5, a KPI-based degradation model is extracted directly from ABA signals. The methodology uses a Hilbert spectrum approach to detect damaged rail welds in a regional railway network. Next, an evolutionary multi-objective optimization model is proposed considering two objective functions, i.e. performance-related objective functions and cost-related objective functions. The objective functions are defined by using the predictive and robust KPIs. The aim is to have a set of Pareto solutions which minimize operation costs and maximize the rail performance.

At the end of the current dissertation, conclusions for each part are provided to not only present the capability of the proposed methodologies but also highlight assumptions, limits and potential improvements for each chapter. Future research is also included to show major potentialities that can be fulfilled for future researchers who are interested in the current topic. Moreover, main recommendations according to the acquired conclusions for ProRail are discussed.

Introduction

1.1 Introduction

During recent years, a shift from road to rail has been promoted in Europe. This shift will have advantages: an increase in transportation of passengers and goods by rail, mitigation of traffic congestion, reduction of air pollution by increasing the use of electrified railways based on resources, especially wind and solar energy, among other advantages. However, the increased use of railway implies a pressure on keeping the trains running over longer period, making monitoring and maintenance activities on the tracks more difficult to be scheduled. Thus, new integrated railway asset management is required to keep the railway system performance at the desired level. The current dissertation is focused on an intelligent rail maintenance decision support system using KPIs. This includes all the important indicators such as safety and life cycle costs based on the perspective of both the railway infrastructure manager and users [1]. In the current chapter, a brief description of rail asset management, railway track infrastructure and its components and also an overview of the dissertation are included.

1.2 Railway asset management

Asset management is a broad concept. To make the concept specific, it is important to define asset management properly [1]: "*The management of the railway assets involves a range of activities such as building, inspection, maintenance, enhancement and renewal aimed at optimising performance, risks and costs of the infrastructure*".

To keep control of the activities in the asset management, a decision-making support system is needed. Moreover, the decision support system should keep the infrastructure manager informed of the asset condition. A typical railway network consists of several assets. In The Netherlands, in which this PhD research is conducted, the railway network includes

6830 kilometres of tracks, 5100 tunnels and viaducts, 4500 kilometres of overhead wiring, 7508 switches, a signalling and safety control system and 388 stations [2]. The main focus in this dissertation is on railway track infrastructure as the railway track contributes significantly in keeping the railway network at the acceptable performance level. Furthermore, almost half of the annual railway maintenance budget in the Netherlands is allocated to the track infrastructure [2].

1.3 Overview of track structure

This section presents concisely basic concepts of a track structure. There are two types of tracks in The Netherlands: ballasted track and slab track. Figure 1.1 depicts the main structure of a ballast track including a zoom-in picture of a rail piece with a rail defect (squat) on it.

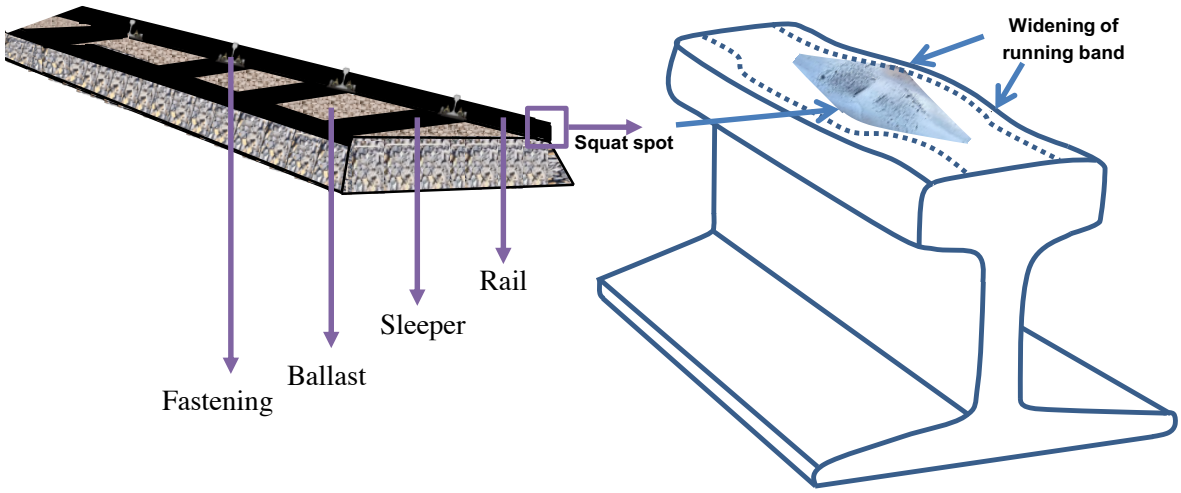


Figure 1.1 The structure of a track ballast with a zoom-in picture of a rail piece

A typical railway track is composed of the following components: rail, ballast, sleeper and fastening. These components are designed to transfer the loads caused by train to the track sub structures. To do so, a rail fastening system should fix rails to sleepers using clamps and rail pad. Then, sleepers transfer loads from rolling stocks to the track ballast. The ballast is composed of stones to keep control of stability of the whole track structure. Train and track are related to each other via the contact force between rail and wheel. This contact force is of major concern of railway engineering as it can cause huge challenges for infrastructure maintenance and costs. As the contact area between wheel and rail is small, it requires to bear a load, causing large contact stress. These stresses can potentially lead to material fatigue, or so-called Rolling Contact Fatigue (RCF). In the current dissertation, the RCF defects are used as an input for the proposed rail maintenance decision support system in order to include explicitly the rail health condition into the decision making.

1.4 Rail defects

RCF defects are important type of rail defects. Different factors such as axle load and track age can also influence the formation and evolution of the RCF defects which potentially cause rail break [3]. The analysis of rail failure risk due to the RCF defects is of the concerns of the infrastructure manager [4]. Moreover, there are critical track pieces, many are subject to rail defects, most importantly welds, insulated joints and switches. Thus, beside the common types of the rail defects such as squats, head checks and corrugations, damaged welds, switches and insulated joint can be defective and increase rail failure risk over rail life time [5]. In Figure 1.2 a, a squat is shown in different stages of growth from the initiation to the moderate size associated with its visual length and area. As can be seen in the figure, the defect evolves and gets more critical for the rail over time. In Figure 1.2 b, a rail weld prone to develop RCF defects is presented.

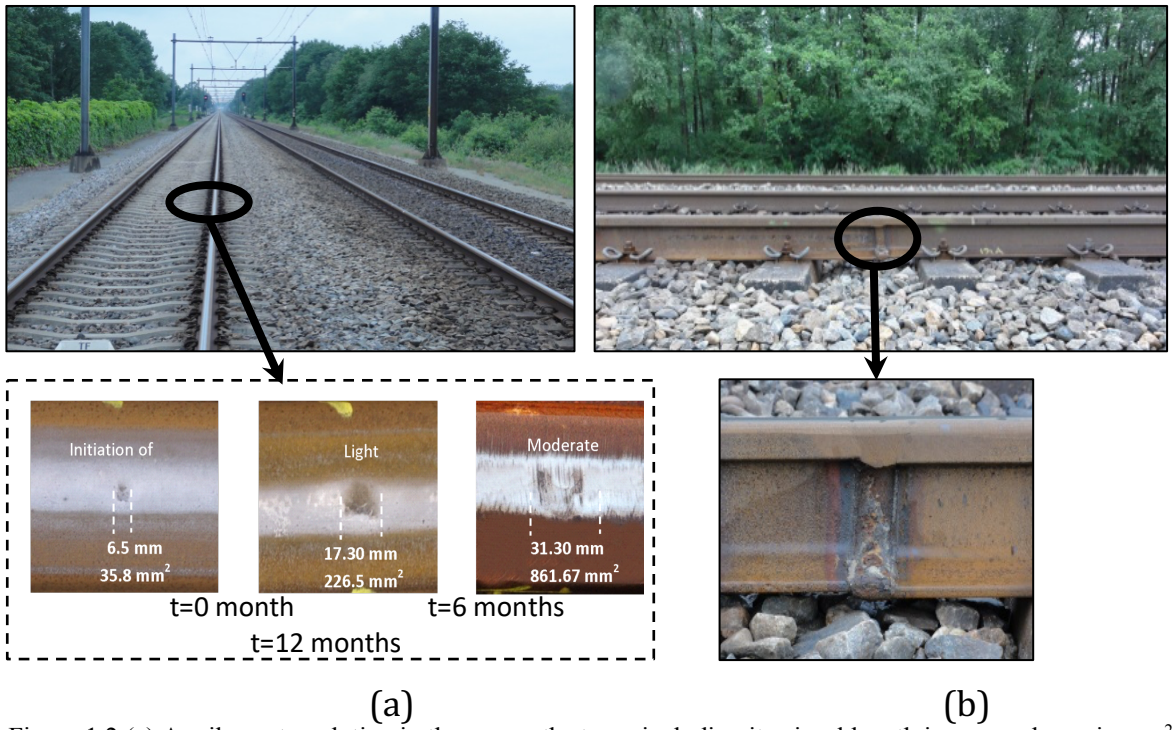


Figure 1.2 (a) A rail squat evolution in three growth stages including its visual length in mm and area in mm^2; (b) A damaged weld prone to RCF defect.

1.5 Performance analysis

The railway infrastructure performance is typically measured based on three factors including level of safety, availability and quality [6]. The performance measurement should also include all the critical health condition indicators of the infrastructure, such as life cycle costs of all its assets, social and environmental impacts, considering the perspective of both railway infrastructure manager and users. The infrastructure health condition is an influential information for the infrastructure manager to make proper maintenance decisions [7]. By having a monitoring system, it is possible to keep manage of the actual health condition of the infrastructure over number of trains passing by track. The obtained track traffic trend (over a given time horizon) is used to model the infrastructure degradation. Traffic can vary from one track to another at the same time horizon.

1.5.1 Degradation analysis

The degradation of the infrastructure can be estimated relying on health condition over time, in particular for critical infrastructure like rails [8]. The degradation analysis assists to keep the infrastructure manager properly aware of the criticality of the rail infrastructure and evolution of the criticality over time. Geographical factors, infrastructure design and maintenance planning can hugely influence the degradation process of the rail infrastructure [9]. Rail degradation in a track is usually different from one rail kilometre position to another. The track positions that have rail defects are more degraded compared to the track pieces with less defects. Further RCF defects, damaged welds and insulated joints accelerate the rail degradation process [10].

1.5.2 Actual rail health condition monitoring

To estimate the actual rail degradation, different rail health condition monitoring methods are available e.g. eddy current [11], ultrasonic [12], Axle box acceleration (ABA) and video camera [13]. In recent years, the range of sensing technologies has expanded rapidly including guided-wave based monitoring [14], networking technologies and mobile ad hoc networking [15], and different types of the cheap wireless sensor networks (WSNs) [16], [17]. Relying on the chosen monitoring system, the actual rail health condition data are typically collected at certain periods. The measurements collect information that can help to reduce maintenance cost, unnecessary operations and also to focus on critical rail pieces where the actual health condition has reached the crucial level defined by the infrastructure manager.

1.6 Maintenance decision system

In Figure 1.3, a KPI-based rail maintenance decision making system is depicted. Three steps are presented including actual rail health condition monitoring, performance assessment and maintenance decision making. In the first step, a big data analysis approach is used in this dissertation to tackle the challenges related to the data size [17]. On the basis of the big data analysis of the rail measurement data, the degradation model is developed. Using the degradation analysis, a rail defect prediction model is proposed in order to define a set of robust and predictive KPIs. The obtained KPIs are used to support the infrastructure manager for proposing maintenance decisions. By having the rail health condition in a given maintenance time horizon, in the step 3, a maintenance decision making methodology is proposed to obtain maintenance decisions using the robust and predictive KPIs.

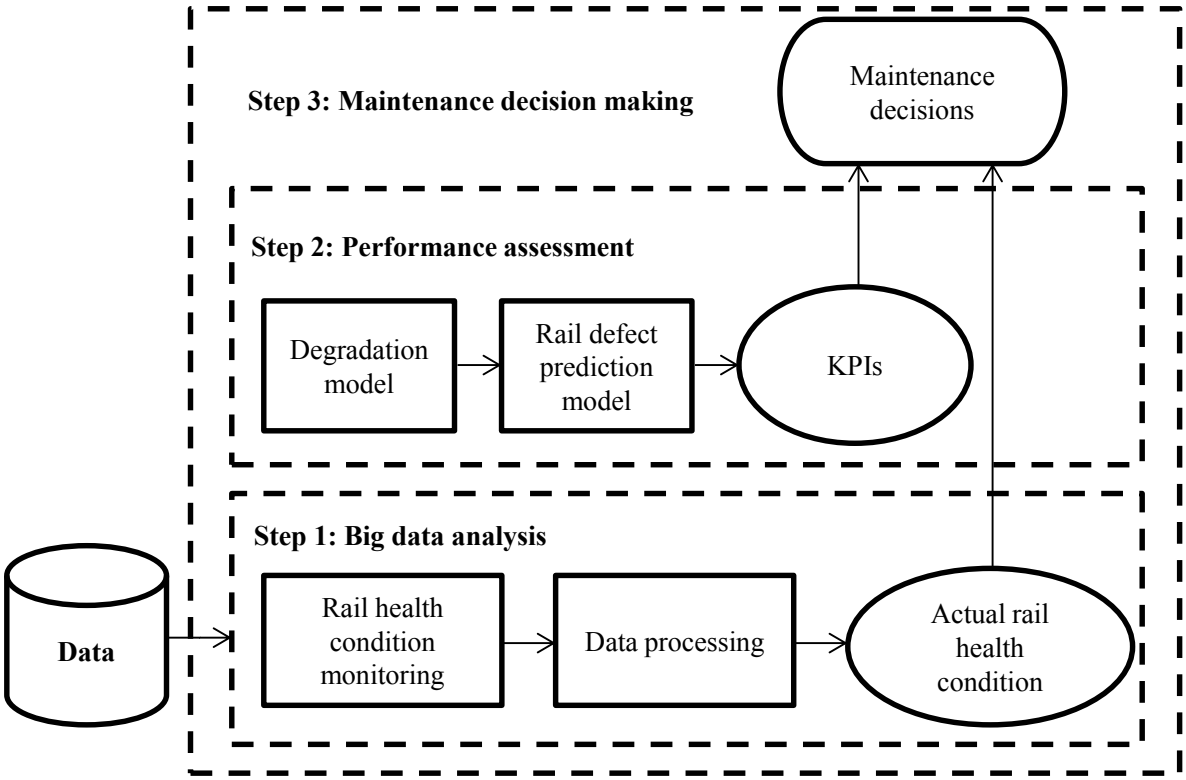


Figure 1.3 A global view of a KPI-based rail maintenance approach using big data analysis

1.7 Problem statement

To state the problem in the current thesis, a set of questions is defined to be answered over different chapters.
The research question is the following:
How can KPIs help the infrastructure manager to improve railway infrastructure performance?
This research question is divided into four sub-questions:

1. How to define predictive and robust rail KPIs for a railway network?
2. How to develop a set of rail maintenance decisions using the proposed KPIs?
3. How to estimate rail failure risk using a risk KPI?
4. How can infrastructure manager include simultaneously performance and cost to optimize rail maintenance decision making?

1.8 Outline of this dissertation

The outline of this dissertation is shown in Figure 1.4. The holistic framework of the proposed intelligent rail maintenance decision support system is divided into three parts including KPIs, degradation model and rail maintenance decision support system. Because the modelling of rail maintenance is covered extensively in the literature, it is only briefly discussed in the introduction of this dissertation. The effect of the defined KPIs on the rail health condition is discussed in Chapter 2. Chapter 3 presents a methodology to design risk KPIs including the probability of rail failure based on defect crack depth and traffic tonnage. The aim is to prioritize the rail kilometre positions in which the rail is prone to break. The methodology helps infrastructure manager to make proper decisions for rail replacement using the results of the

rail failure risk. Both Chapters 2 and 3 benefit from using degradation models over a prediction horizon and including relevant stochasticity. Chapter 4 and Chapter 5 propose rail maintenance methodologies based on the obtained KPIs. Chapter 4 defines and analyses the effect of track influential factors on rail maintenance decisions. To do so, the chapter presents an investigation on how to read the rail observation according to those influential factors using an expert system approach. The results ease finding most critical rail pieces for grinding operation. By having the predictive and robust KPIs and the proposed rail maintenance methodology, in Chapter 5, a multi-objective optimization model is included to analyse trade-offs between the rail performance and rail maintenance operation costs. The results present a Pareto front which can support the maintenance decisions for a railway network in regional network. At the end of the dissertation, Chapter 6 concisely concludes and discusses the dissertation achievements and gives some recommendations in order to make the proposed framework applicable for ProRail.

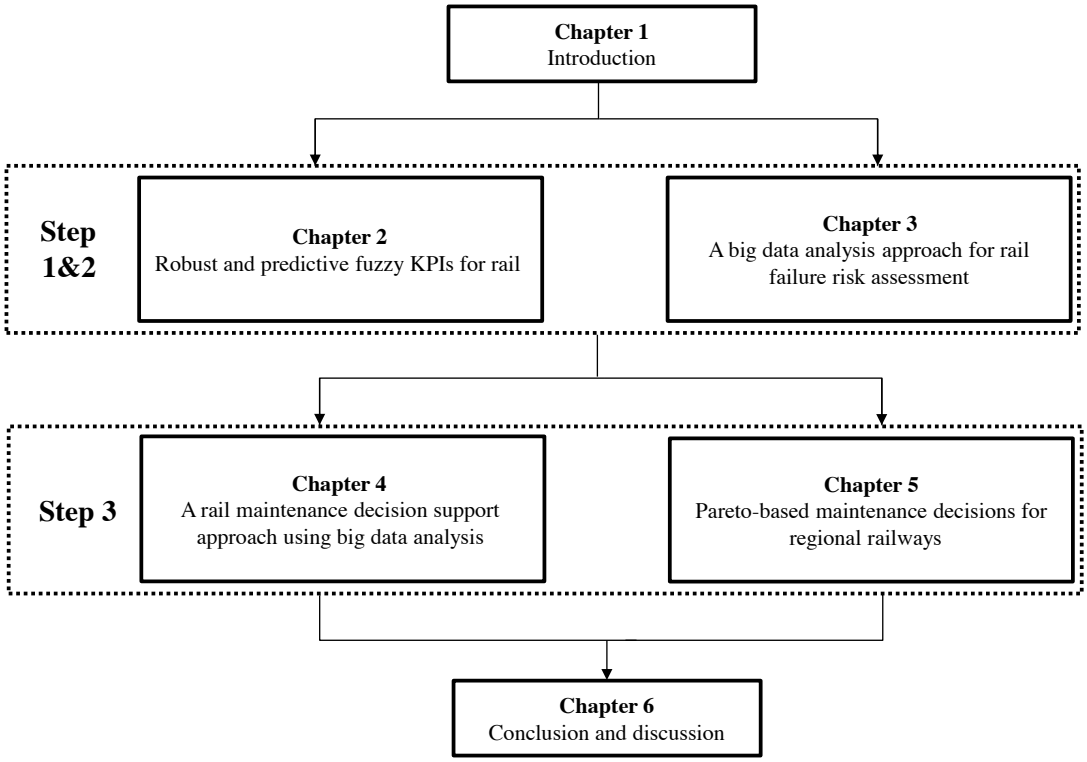


Figure 1.4 Outline of the dissertation.

1.9 References

[1] Rama, D., & Andrews, J. D. (2016). Railway infrastructure asset management: the whole-system life cost analysis. *IET Intelligent Transport Systems*, *10*(1), 58-64.

[2] Zoeteman, A., Dollevoet, R., & Li, Z. (2014). Dutch research results on wheel/rail interface management: 2001–2013 and beyond. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *228*(6), 642-651.

[3] Min, Y., Xiao, B., Dang, J., Yue, B., & Cheng, T. (2018). Real time detection system for rail surface defects based on machine vision. *EURASIP Journal on Image and Video Processing*, *2018*(1), 3.

[4] Liu, X., Saat, M., & Barkan, C. (2012). Analysis of causes of major train derailment and their effect on accident rates. *Transportation Research Record: Journal of the Transportation Research Board*, (2289), 154-163.

[5] Molodova, M., Li, Z., Núñez, A., & Dollevoet, R. (2014). Automatic detection of squats in railway infrastructure. *IEEE Transactions on Intelligent Transportation Systems*, *15*(5), 1980-1990.

[6] Parida, A., Kumar, U., Galar, D., & Stenström, C. (2015). Performance measurement and management for maintenance: a literature review. *Journal of Quality in Maintenance Engineering*, *21*(1), 2-33.

[7] Luan, X., Miao, J., Meng, L., Corman, F., & Lodewijks, G. (2017). Integrated optimization on train scheduling and preventive maintenance time slots planning. *Transportation Research Part C: Emerging Technologies*, *80*, 329-359.
[8] He, Q., Li, H., Bhattacharjya, D., Parikh, D. P., & Hampapur, A. (2015). Track geometry defect rectification based on track deterioration modelling and derailment risk assessment. *Journal of the Operational Research Society*, *66*(3), 392-404.
[9] Bai, L., Liu, R., Wang, F., Sun, Q., & Wang, F. (2017). Estimating railway rail service life: A rail-grid-based approach. *Transportation Research Part A: Policy and Practice*, *105*, 54-65.
[10] Chattopadhyay, G., & Kumar, S. (2009). Parameter Estimation for Rail Degradation Model. *International Journal of Performability Engineering*, *5*(2).
[11] Song, Z., Yamada, T., Shitara, H., & Takemura, Y. (2011). Detection of damage and crack in railhead by using eddy current testing. *Journal of Electromagnetic Analysis and Applications*, *3*(12), 546.
[12] Fan, Y., Dixon, S., Edwards, R. S., & Jian, X. (2007). Ultrasonic surface wave propagation and interaction with surface defects on rail track head. *Ndt & E International*,*40*(6), 471-477.
[13] Jamshidi, A., Hajizadeh, S., Su, Z., Naeimi, M., Núñez, A., Dollevoet, R., De Schutter, B., & Li, Z. (2018). A decision support approach for condition-based maintenance of rails based on big data analysis. *Transportation Research Part C: Emerging Technologies*, 95, 185-206.
[14] Mariani, S., Nguyen, T., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., Staszewski, W. J., ... & Carr, G. (2013). Noncontact ultrasonic guided wave inspection of rails. *Structural Health Monitoring*, *12*(5-6), 539-548.
[15] Hodge, V. J., O'Keefe, S., Weeks, M., & Moulds, A. (2015). Wireless sensor networks for condition monitoring in the railway industry: A survey. *IEEE Transactions on Intelligent Transportation Systems*, *16*(3), 1088-1106.
[16] Flammini, F., Gaglione, A., Ottello, F., Pappalardo, A., Pragliola, C., & Tedesco, A. (2010, October). Towards wireless sensor networks for railway infrastructure monitoring. *Electrical Systems for Aircraft, Railway and Ship Propulsion (ESARS),* 1-6.
[17] Santur, Y., Karaköse, M., & Akın, E. (2016). Learning Based Experimental Approach For Condition Monitoring Using Laser Cameras In Railway Tracks. *International Journal of Applied Mathematics, Electronics and Computers*, *4*(Special Issue-1), 1-5.
[18] Thaduri, A., Galar, D., & Kumar, U. (2015). Railway assets: A potential domain for big data analytics. *Procedia Computer Science*, *53*, 457-467.

Robust and predictive fuzzy KPIs

This chapter corresponds to the reference: *A. Jamshidi, A. Núñez, R. Dollevoet, and Z. Li, "Robust and predictive fuzzy key performance indicators for condition-based treatment of squats in railway infrastructures". Journal of Infrastructure Systems, Volume 23, Issue 3, September 2017, 04017006. DOI: 10.1061/(ASCE)IS.1943-555X.0000357.*

2.1 Introduction

During the recent years, a modal shift from road to rail has been promoted in Europe. The idea is to increase the share of transport demand for mobility of people and freights. Reduce road traffic congestion, make efficient use of the energy resources and tackle the major challenges of climate change. Major contributions are needed in the optimal management of railway assets, evolving towards a more automated predictive operation where functional assets are monitored. This includes all the important indicators such as economical, safety and societal impacts, considering the perspective of both railway infrastructure manager and users [1].

A typical set of railway assets is shown in Figure 2.1, and it includes the track, station, superstructure, sub-structure, communication, catenary, control room, signalling system, rolling stock, barrier, security and surrounding. In order to monitor and properly maintain the railway assets, it is necessary to measure the evolution of important health condition indicators over time, also called key performance indicators (KPIs), for each of the critical assets. For

example, in the Figure 2.1, $J_{label}^{Asset}(t)$ relates to the KPI for the health condition of an asset called "Asset", uniquely labelled as "label" at time *t*. In The Netherlands, the assets in the railway network include more than 3,000 km of track, 388 stations, being one of the densest networks in Europe. In this network, the design of an optimal maintenance plan for all its assets is a challenging problem. To optimally design the maintenance plans, the infrastructure manager requires to provide crucial information of each asset [2], and maintenance decision making considering risk averse situations [3]. Thus, the optimal maintenance plan is a necessity because of the high demand from users and government for a better quality of service, and the need of keeping costs as low as possible.

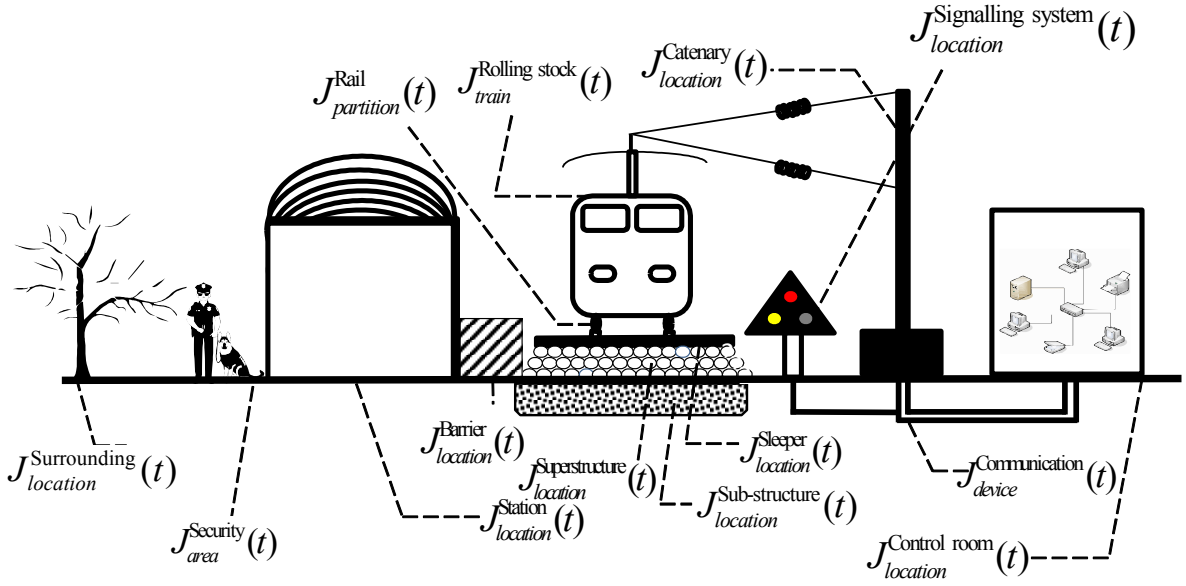


Figure 2.1 Main components of railway infrastructures.

Maintenance Performance Indicators evaluate the system performance and can be used to guarantee that these assets operate at an acceptable level of functionality and safety. In [4], a general systems framework is proposed using a hierarchical structure of multi-criteria maintenance performance measurements. In [5], the same framework is applied to the case of benchmarking railway infrastructure maintenance operations. Three different hierarchical levels are proposed: strategic level for top management decisions, tactical level for middle management and functional level for supervisors/operators. The general framework requires effective measurements of the health condition of the assets considering that the different assets degrade with different rates due to the effect of different exogenous sources. Particularly, the focus of this chapter is to design robust and predictive fuzzy performance indicators for health condition monitoring of railway tracks, considering a particular major type of Rolling Contact Fatigue (RCF) called squat (see [6]).

In The Netherlands over forty percent of the railway maintenance budget is allocated yearly to track maintenance [7]; [8]. The presence of RCFs accelerates track degradation which negatively influences its health condition. RCFs also increase the noise level that affects people living in the surroundings and in a worst case making a huge impact on safety as severe RCF's can result in derailment. For track maintenance to be effective, the planning should consider

not only costs but also the dynamics of RCFs. Complex interactions between environment, vehicles, wheels and track interface, structure and also different behaviours under maintenance operation such as grinding and rail replacement can be considered. In [9] rail degradation is modelled by a time to failure function using MGT (million gross tons) measurements and around 12 failure events, decision making is proposed in a Monte Carlo simulation setting. The maintenance operations are modelled as different cost functions, including rail grinding costs, track tamping costs, rail lubrication costs, among other maintenance operations. [10] assess the value of preventive maintenance in comparison with corrective maintenance. The idea is to analyse cost-benefit of using preventive maintenance including four different maintenance costs: maintenance inspections, repair of potential failures, repair of functional failures and service/production loss. In the case study for a Swedish railway line, the ten costliest railway sections are found to have three times the tonnage compared to the sections with the lowest costs, and also the costliest sections experience 4.5 times more track failures. The conclusion is that the railway sections with the lowest total maintenance cost have implemented more preventive maintenance actions.

In the literature, different studies have been carried out to present how a degradation model for tracks can be embedded on asset management to facilitate maintenance plans. Track geometry measurements relying on statistical analysis are used to capture the track degradation effect [11], [12], [13], [14], [15], [16], [17], [18]. In those papers, different time-dependent degradation models are proposed; they can all be used to improve maintenance interventions. Estimation of the track safety and considering the probability of rail break has also been investigated [19], [20], [21]. Detailed mechanical models can give many insights about the evolution of rail defects; however, the use of those models for maintenance planning operations requires sophisticated knowledge about the track and its operational conditions that are not always available or easy to obtain in practice. Fuzzy logic has increasingly been used in different fields; in particular, in the ones where uncertainties can influence the decision process. It is used to measure performance in different infrastructures by predicting failure of components [22], [23], optimizing asset condition [24], [25] and decision making [26]. In this chapter, the authors propose the use of an interval fuzzy model to capture the most important dynamics of squats in railway infrastructure, from the maintenance operation point of view. The authors aim to keep the prediction as simple as possible, but suitable enough to ease decision making in practice. The use of key performance indicators (KPIs) that are able to explicitly include the dynamics of the deterioration of the assets, together with an appropriate set of scenarios for the principal sources of stochasticity that might affect their performance are recommended. A fuzzy Takagi Sugeno (TS) interval model [27], [28], [29] is calibrated using real-life data collected over years of field tests and measurements. That helps obtaining numerical models capable to predict squat growth over a time horizon under different possible scenarios and under different maintenance decisions.

Based on the interval fuzzy models for squats, a condition-based methodology for rails is proposed in this chapter using different KPIs that are defined in a track-partition level which allows the grouping of defects located in a given track partition. In this methodology, the number and density of squats are considered over a prediction horizon under three different scenarios, vis. slow, average and fast growth. Then, to facilitate visualization of the track health condition and to ease the maintenance decision process, the chapter proposes a fuzzy global KPI based on fuzzy rules for each partition that merges the different KPIs over prediction horizon and scenarios. The methodology is evaluated with data from a Dutch railway track, relying on the use of technology-based Axle Box Acceleration (ABA) measurements, capable to detect the early stage squats on the rail [30], [31]. An introduction of the ABA measuring system is

described in Section 2.2, including background of the ABA measurement system and its application in rail condition monitoring based on ABA.

Figure 2.2 shows the flowchart of the proposed methodology divided in three steps. In Step 1, relying on ABA measurements, the health condition of the track and severity are estimated. A list of defects is assumed to be provided by the detection algorithm. In Step 2, using interval fuzzy TS model, the growth of each detected defect *i* is evaluated over time and different possible evolution scenarios are considered. Three models are evaluated, with grinding, replacement and without maintenance. The idea is to see the consequences of the maintenance operations on the detected squats for different scenarios over a prediction horizon. At the end, in Step 3, a global fuzzy KPI is used to describe the condition at a track partition level, for a given travel direction, left and right rails. The global fuzzy KPI at a partition, combines the effects of a vector of KPIs over a prediction horizon, considering three most representative defect evolution scenarios.

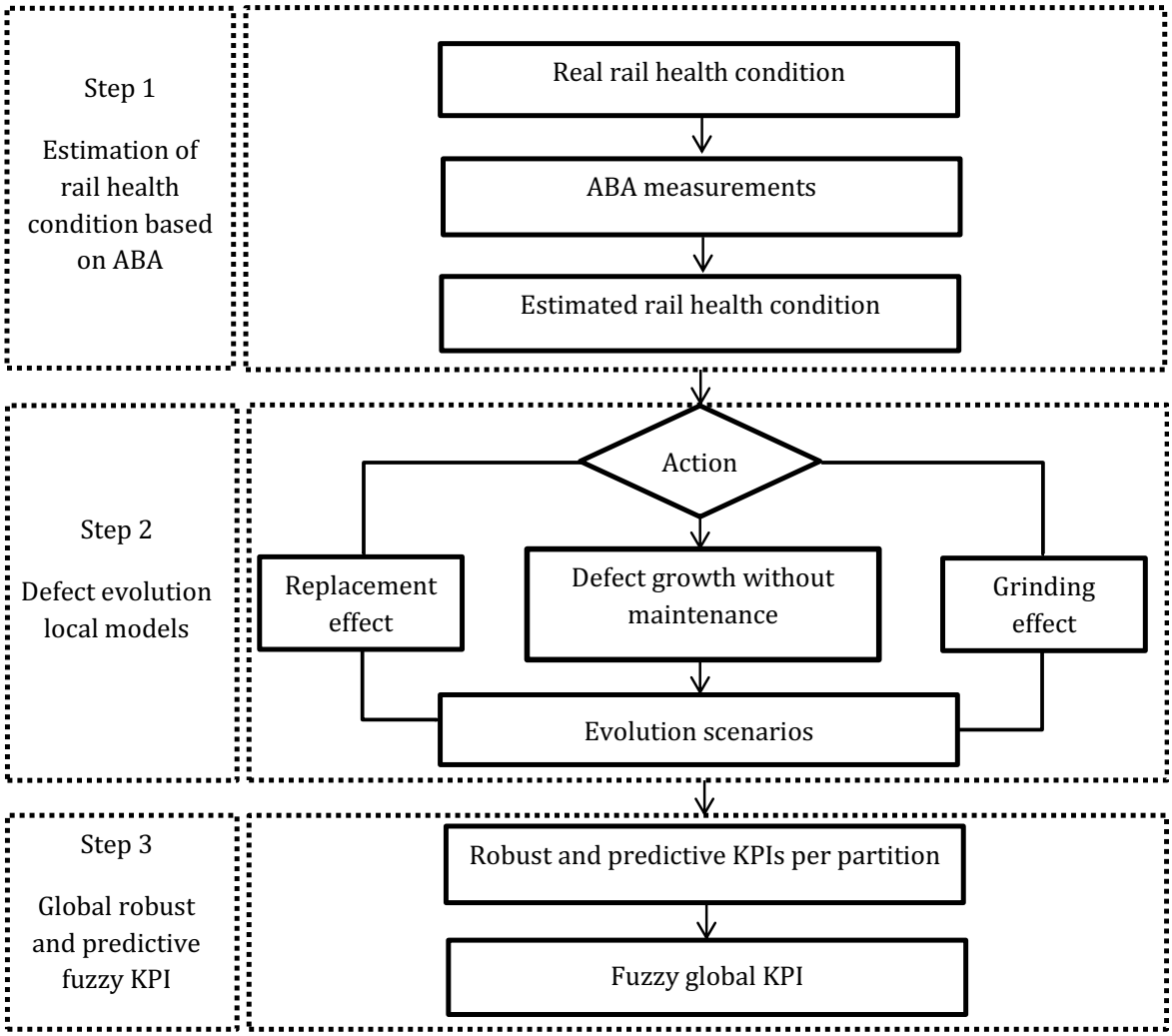


Figure 2.2 flowchart of the proposed methodology

The chapter is divided as follows. In Section 2.2, the main elements of the ABA based detection methods are presented. Fuzzy interval models for squats are presented in Section 2.3

for three cases: without maintenance, after grinding and after replacement. In Section 2.4, different KPIs are defined at a track partition level in order to aggregate the local dynamic behaviour of squats. Because of the number of scenarios and prediction horizon, the fuzzy global KPI is proposed to facilitate decision making. Later in Section 2.5, the numerical results and discussion are presented. Finally, conclusions and further research are discussed in Section 2.6.

2.2 ABA-based health condition monitoring in railways

2.2.1 Background of the ABA measurement system

There are different methods to diagnose the condition of rail defects, including ultrasonic measurements, eddy current testing, image recognition and guided-wave based monitoring among other technologies. Each of them has different advantages and disadvantages. In this chapter, a technology is needed capable to detect defects in an early stage, thus the authors consider the use of ABA measurements [32]; [30]. [31] investigated the feasibility of detecting early-stage squats using an ABA prototype. It is reported that squats could be detected by analysing the frequency content of the ABA signals in the wavelet power spectrum. In practice, the useful frequency band for early detection of squats ranges from 1000-2000 Hz and 200-400 Hz [30].

In the literature, it has been reported that ABA systems can be employed to detect surface rail defects like corrugation, squats and welds in poor condition. The ABA system offers the advantages of (1) having a lower cost than other types of detection methods, (2) it is easy to maintain and (3) can be implemented in-service on operational trains. Other significant advantages that ABA offers over similar measurement systems are (4) the ability to detect small defects with the absence of complicated instrumentation and (5) the ability to indicate the level of the dynamic contact force [33].

2.2.2 Rail condition monitoring-based on ABA

In this study, the authors are users of the ABA detection methodology presented in [31] and [30]; thus, it is assumed that a list of squats and their location are available. Let us define the counter of squat defects as $i=1,2,\ldots,N_{defects}$, where x_i represents the position of the squat i. The authors define $H(x,k)$ and $L(x,k)$ as the real rail condition and real squat length respectively, defined at position x and time step k. The authors only focus on positions x_i where squats are detected. To simplify the notation, it is assumed that $H_i(k)=H(x_i,k)$ and $L_i(k)=L(x_i,k)$ represent the severity and the length of squat i at time step k. To systematically classify squats in terms of severity, the authors follow the terminology used in [34], [35] and [36]. The definitions of these three references are compatible to one another. Although the transition between one class to the other is not always abrupt, the authors have defined fixed values for those transitions according to our experience. Depending on the squat length $L_i(k)$, measured in mm, the severity of the squat can be used to represent the health condition of the rail at location x_i as follows:

$$H_i(k)=\begin{cases} S & \text{if} \quad 0 \le L_i(k) < 8 \\ A & \text{if} \quad 8 \le L_i(k) < 30 \\ B & \text{if} \quad 30 \le L_i(k) < 50 \\ C & \text{if} \quad 50 \le L_i(k) < 60 \\ RC & \text{if} \quad L_i(k) \ge 60 \end{cases} \tag{2.1}$$

where *S* refers to a seed squat, *A* is a light squat (*A* squat), *B* is a moderate squat (*B* squat), *C* is a severe squat (*C* squat) and *RC* is a squat with risk of derailment. The boundaries were defined based on general guidelines to classify squats. Figure 2.3 depicts an example of defect growth collected from field measurements in the track Meppel-Leeuwarden. In the figure, the x-axis represents kilometre position of the track where the squats are located and the y-axis indicates the time in three different months, month 0 (moment of the measurement), month 6 and month 12. In the diagram, *A* squats are drawn as circles and *B* squats are squares. Different squats grow with different rates. In the average case, the track measurements show that it takes approximately 9 months for an *A* squat of 20 mm to evolve into an *B* squat of 30 mm.

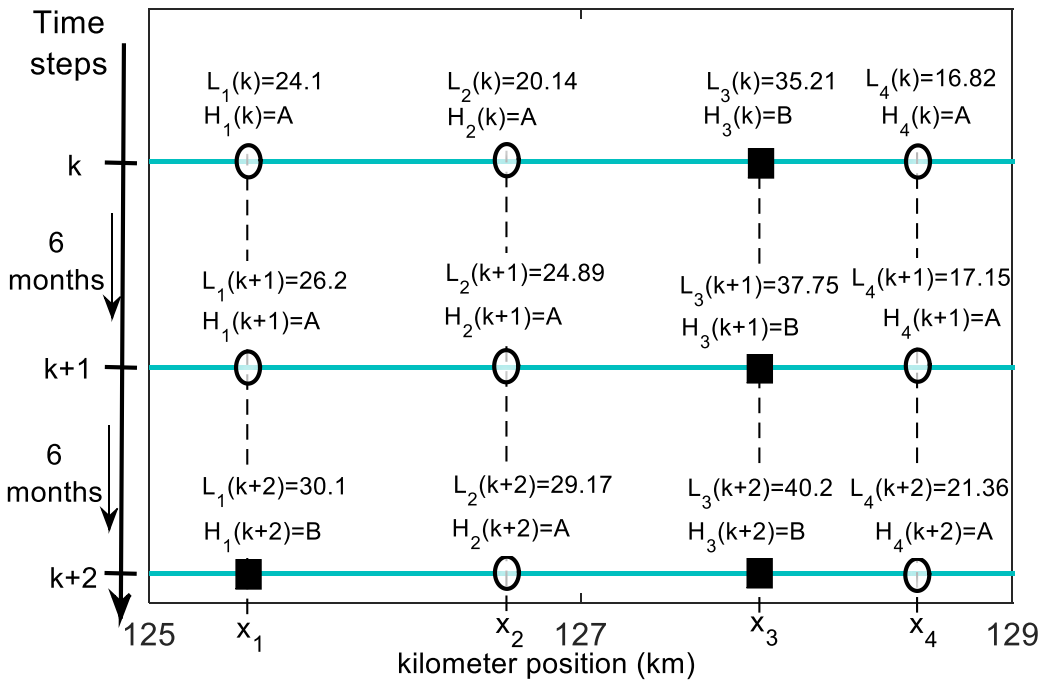


Figure 2.3 An example of defects evolution over time. The x axis is the kilometre position in the track, x_i the position of squat *i*, y axis is time every six months. In circles are *A* squats, squares are B squats.

In this study, the ABA measurements are used to develop a model for defect evolution. For each squat, the related energy of the ABA is available using wavelet spectrum analysis and advanced signal processing methods [30]. Relying on the ABA measurement, the energy values of the ABA signals can be calculated at every position *x* at time step *k* as *E(x,k)*. From the energy signal, we are interested only in those locations with squats, namely $E_i(k)=E(x_i,k)$. For using the energy of the ABA signal to predict the squat length evolution, a correlation between the squat length and energy of the ABA signal was performed. Photographs from track visits of several years are used to measure the lengths of the squats and to fit the piecewise linear correlation model.

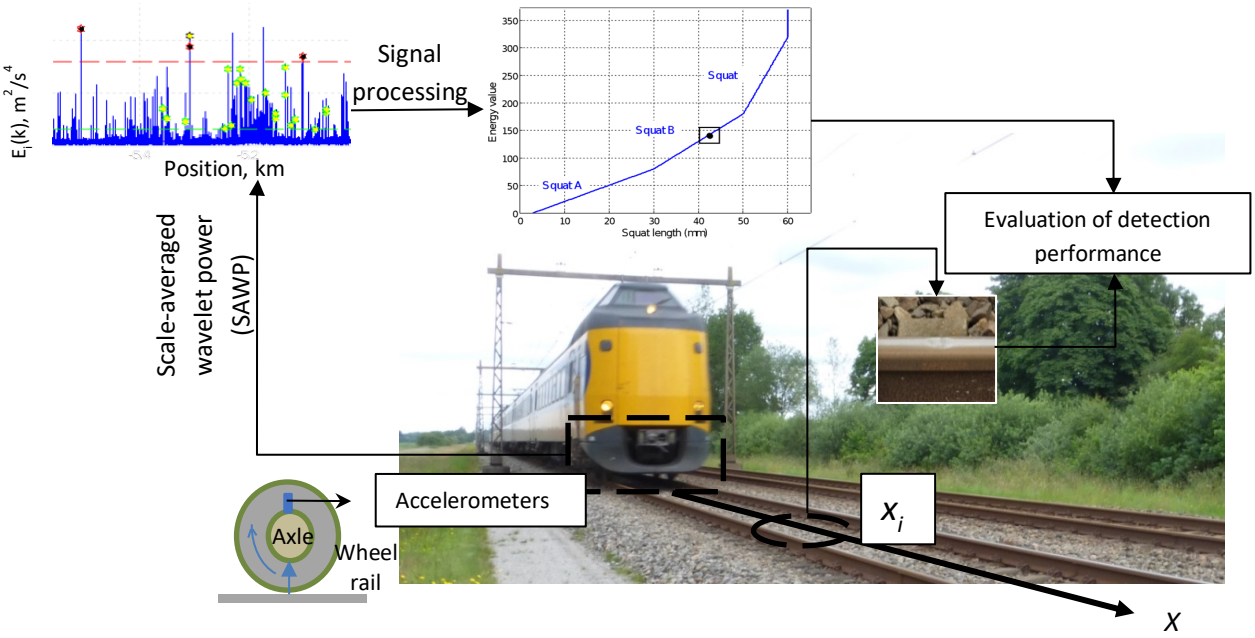


Figure 2.4 Global scheme of the main components of the Step 1: Estimation of track health condition based on ABA

The estimated length $\hat{L}_i(k)$ of squat *i* at time step *k* as function of the energy value $E_i(k)$ is given by:

$$\hat{L}_i(k) = \begin{cases} g_1 E_i(k) + q_1 & \text{if} \quad E_i(k) < 80 \\ g_2 E_i(k) + q_2 & \text{if} \quad 80 \le E_i(k) < 170 \\ g_3 E_i(k) + q_3 & \text{if} \quad 170 \le E_i(k) < 300 \\ g_4 E_i(k) + q_4 & \text{if} \quad E_i(k) \ge 300 \end{cases} \tag{2.2}$$

where the slope of local linear functions is g_m, m=1,…,4, and the bias q_m, m=1,…4, are adjusted to the specific track. For relation (2.2), we have been users of previous work of our group, [37], [31]. In general, we can say that the correlation coefficient and residual standard get affected by the speed of the measurement train. In this chapter, we assumed that the measurement is done at commercial speed as was done for the test measurement so far, and we have disregard segments that were measured out of a reasonable range of speed. A global view of the Step 1 of the methodology, estimation of track health condition based on ABA, is presented in Figure 2.4. As shown in the figure, in order to estimate the length $L_i(k)$, the energy value $E_i(k)$ is calculated using the ABA measurement. Hence, relying on the estimated squat lengths, the rail health condition $H_i(k)$ can be approximated. In the figure, a squat is detected with an energy value $E_i(k)$=145 m^2/s^4, the estimated squat length $L_i(k) = 43mm$ and the estimated health condition $\hat{H}_i(k) = B$.

2.3 Fuzzy interval models for squats

2.3.1 Maintenance oriented models for squats

Typically, maintenance slots in the Dutch railway network are decided based on long and short-term planning for preventive and corrective maintenance respectively. In the long term, the contractor should inform the asset manager at least one year before cyclic grinding for using the equipment needed. In the short term, normally, the maintenance is performed when the squats are in the last stage of growth (C squat). Thus, a predictive approach by employing well designed KPIs should aim to improve both short and long-term planning, (1) keeping a good balance between costs and health condition of the track, (2) simplifying the design of maintenance plan over the whole time horizon and (3) increasing indirectly the track safety.

The experimental results show that each squat can grow with a different rate. The estimation of squat lengths can be affected by the subjectivity of the human error. For instance, one source of uncertainty comes from the fact that visually only the rusty area of the defects is used to measuring the length, while the defect might be longer. Fuzzy systems can work under subjective environments. In the proposed methodology, the design of the global fuzzy KPI deals with the subjectivity. The definition of a low or a big number of defects will depend on the subjectivity of the infrastructure manager, and on how this information is incorporated for maintenance decision making. In order to generalize this characteristic, fuzzy confidence intervals can be used to capture the stochasticity of different scenarios for the squat growth. The upper bound of the interval represents a worst case scenario, while the lower bound represents a slow growth rate scenario. In the fuzzy interval approach, the average behaviour is given by a Takagi-Sugeno (TS) fuzzy model. This is used to approximate nonlinearities by smoothly interpolating affine local models. Each local model is involved in the global model based on the activation of a membership function. According to literature, the identification of fuzzy interval models is divided on three steps: clustering method to generate fuzzy rules, identification of the TS local linear parameters (average model), and identification of the fuzzy variance for each rule [29]; [38]. In this chapter, we use the fuzzy interval approach proposed in [28] and [38], which includes Gustafson Kessel clustering, local identification of the linear parameters and optimization of a parameter α to adjust the width of the interval, minimizing both area of the band and number of data points outside the band.

The general problem of interval defect evolution is as follows. Let's consider different defect growth scenarios $h = h_1,\ h_2,\ \ldots,\ h_H$, time steps $t = k, k+1, k+2, \ldots, k+N_p$, and $u(k)$ the maintenance action at time step *k*. The prediction model for the growth of a squat can be written as:

$$\hat{L}_i^h(k+1) = f_j^h\left(L_i(k), u(k)\right), \quad x_i \in \left[x_j, x_{j+1}\right) \tag{2.3}$$

where $\hat{L}_i^h(k+1)$ is an estimation of the length of the squat *i* located in the track partition *j* at the time step *k+1* considering the scenario *h*. The model considers the effect of maintenance $u(k)$ and the initial condition of the squat $L_i(k)$. Depending on the location of the squat *i* which is

x_i, the authors use a local model corresponding to the track partition *j* where the squat is located, $x_i \in \left[x_j, x_{j+1} \right)$. It is assumed that the dynamics for different squats are similar if they are in the same track partition under the same scenario. In this chapter, three maintenance actions are considered, $u(k) \in \{u_1, u_2, u_3\}$, where u_1 is no maintenance, u_2 is grinding and u_3 is replacement. Also, three scenarios are evaluated, $h = h_1,\ h_2,\ h_3$, where h_1 represents slow growth, h_2 average growth and h_3 is fast growth.

2.3.2 Dynamics of squats without maintenance

In the absence of maintenance, i.e. $u(k) = u_1$, the prediction model for the average growth scenario, h_2, is formulated based on TS fuzzy model:

$$\hat{L}_i^{h_2}(k+1) = f_j^{h_2}(L_i(k), u_1) = f_j^{\mathrm{TS}}(L_i(k)) = \sum_{r=1}^{N_R} \beta_{jr}(L_i(k)) L_{jr}(k) \tag{2.4}$$

$$L_{jr}(k) = a_{jr} L_i(k) + b_{jr} \tag{2.5}$$

$$\beta_{jr}(L_i(k)) = \frac{A_{jr}(L_i(k))}{\sum_{r=1}^{N_R} A_{jr}(L_i(k))} \tag{2.6}$$

where a_{jr}, b_{jr} are the parameters of the fuzzy local model on rule *r*, $r = 1, 2, ..., N_R$ and $\beta_{jr}(L_i(k))$ is the normalized activation degree of the rule *r*. In this chapter the authors will use Gaussians to model the membership degrees, $A_{jr}(L_i(k)) = \exp\left(-0.5 c_{jr,1} \left(L_i(k) - c_{jr,2} \right)^2 \right)$, defined by parameters $c_{jr,1}$ and $c_{jr,2}$ given by the Gustafson Kessel clustering algorithm. Once the TS model is obtained, the slow growth scenario and the fast growth scenario are used as lower and upper bound of the average growth scenario, $\hat{L}_i^{h_2}(k+1)$, respectively. The equations can be defined as:

$$\hat{L}_i^{h_3}(k+1) = \overline{f_j}^{\mathrm{TS}}(L_i(k)) = \sum_{r=1}^{N_R} \beta_{jr}(L_i(k)) \left(L_{jr}(k) + \alpha^{h_3} \Delta_{jr}(L_i(k)) \right) \tag{2.7}$$

$$\hat{L}_i^{h_1}(k+1) = \underline{f_j}^{\mathrm{TS}}(L_i(k)) = \sum_{r=1}^{N_R} \beta_{jr}(L_i(k)) \left(L_{jr}(k) - \alpha^{h_1} \Delta_{jr}(L_i(k)) \right) \tag{2.8}$$

$$\Delta_{jr}(L_i(k)) = \sigma_{jr} (1 + \psi_{jr}^T (\varphi_{jr} \varphi_{jr}^T)^{-1} \psi_{jr})^{0.5} \tag{2.9}$$

where $\hat{L}_i^{h_3}(k+1)$ is the estimated growth length of squat *i* in time step *k+1* in fast scenario, and $\hat{L}_i^{h_1}(k+1)$ is estimated growth length in slow scenario, α^{h_3} and α^{h_1} are tuning parameters in the fast growth scenario and the slow growth scenario respectively. Moreover, $\varphi_{jr} \varphi_{jr}^T$, $\psi_{jr} = \left[L_i(k), 1 \right]^T$ and σ_{jr} are covariance matrix, regression matrix and variance of the local

model. Figure 2.5 depicts the proposed fuzzy confidence interval model including 177 data points used to capture the squat evolution in different stages of growth. A subset of the data used for analysis is included in Table 1. *A* squat from 8 to 30 mm in length have no or shallow cracks. The *B* squats ranging from 30 to 50 mm grow quickly. The *B* squats evolve to *C* squats when the network of cracks beneath the squat gets further spread. All three stages are shown by reference photos of *A* squat, *B* squat and *C* squat in Figure 2.5.

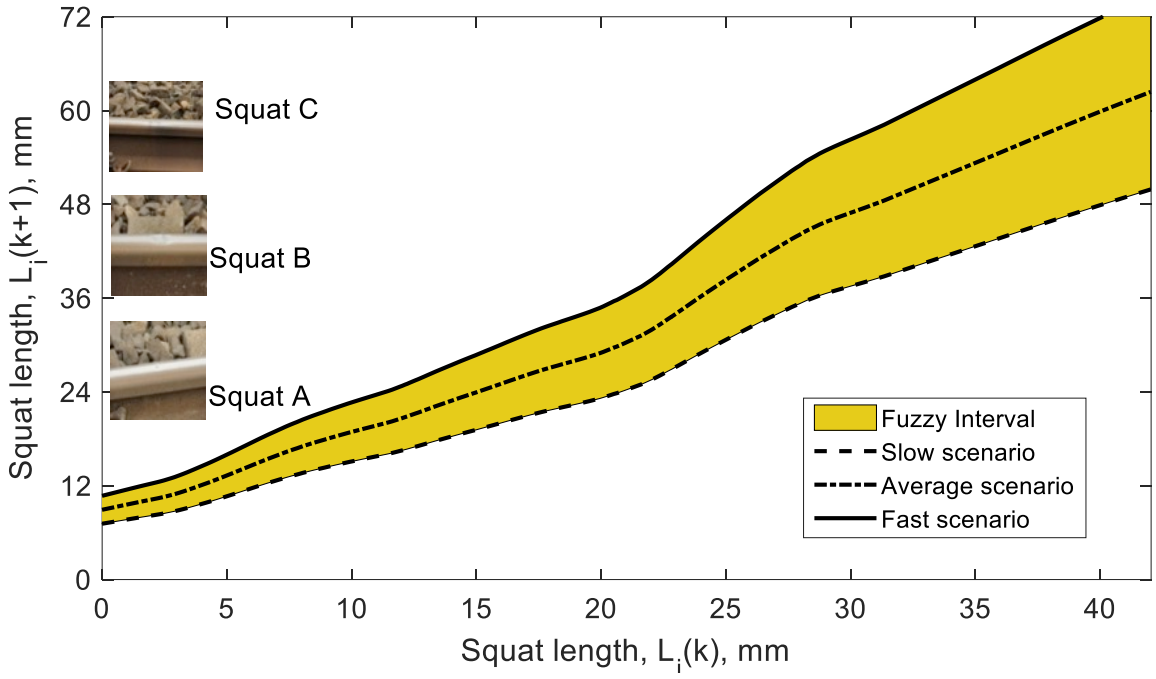


Figure 2.5 A schematic plot of interval fuzzy model for squat growth in the case study track.

Table 2.1 A subset of data used for squat analysis including defect position, km, and visual length, mm, at time k and k+1

Squat	Position, km	$L_i(k)$, mm	$L_i(k+1)$, mm	Squat	Position, km	$L_i(k)$, mm	$L_i(k+1)$, mm
1	104.8438	30.7260	34.7465	11	105.4613	22.8311	24.6695
2	105.1051	37.7420	40.5086	12	105.4953	19.5933	22.0216
3	105.1404	33.2264	37.0496	13	105.5827	14.5360	16.7962
4	105.2116	34.2207	37.7779	14	105.5852	19.5432	21.9787
5	105.3215	46.7870	49.1017	15	105.6353	11.0032	13.9019
6	105.3901	33.0151	36.8862	16	105.6591	25.1642	27.19551
7	105.4195	19.1797	21.6607	17	105.7462	15.4564	17.7552
8	105.4269	20.2236	22.5435	18	106.3105	28.7262	32.2116
9	105.4344	9.4918	12.4747	19	106.8735	55.1141	57.1707
10	105.4561	33.2798	37.0903	20	107.2845	17.8761	20.4044

2.3.3 Rail grinding effect

Squats can be effectively treated by grinding when they are in an early stage of growth. Cyclic rail grinding not only keeps control of maintaining the rail profiles but also to plan track maintenance efficiently [39]. Figure 2.6 depicts squat growth before and after grinding where black points show those squats that did not disappear after grinding. As seen in the figure, some

A squats are located in the effective zone of grinding such that these squats have a zero length after grinding. Those A squats that are imminent to become B squats are located in the ineffective zone for grinding as well as B squats and C squats. Moreover, three growth scenarios in the effective zone are specified to capture the squat evolution rate. Even though grinding severe squats postpones rail replacement, it could accelerate squat evolution as the cracks are not totally disappeared.

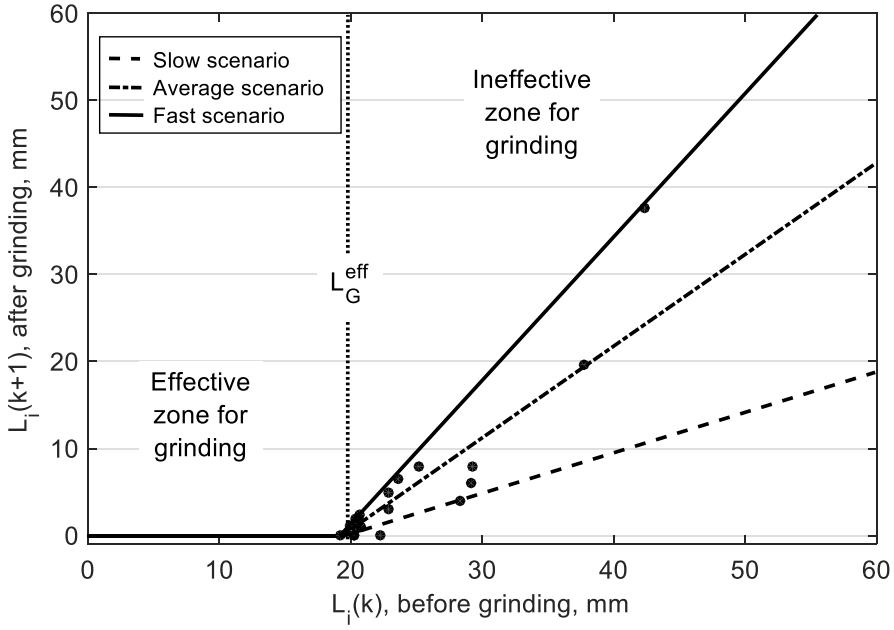


Figure 2.6 Squat growth before grinding and after grinding classified in two effective and ineffective zones for grinding operations. In this case, the depth of the grinding was around 1.0 mm.

The growth model for squat *i* by considering grinding effect can be expressed as:

$$\widehat{L_i^h}(k+1)=\begin{cases}0 & L_i(k)\le L_G^{eff} \quad \text{effective zone for grinding}\\ z_G^h(L_i(k)-L_G^{eff} & L_i(k)\le L_G^{eff} \quad \text{ineffective zone for grinding}\end{cases} \tag{2.10}$$

where L_G^{eff} is the critical squat length that estimate effectivity of grinding, L_G^{eff} is around 20 mm in Figure 2.6 for a grinding depth of 1.0 mm, z_G^h is the slope of the linear model in the ineffective zone for grinding for different scenarios *h,* slow, average and fast growth scenarios.

2.3.4 Rail replacement effect

When the squat severity becomes worse and cracks are grown considerably, grinding is not efficient anymore. Therefore, replacement is the only solution. As replacing a piece of rail takes time and it is costly, an optimal decision making for when and where the rail should be replaced is important. As in a track with curvature, two rail have different degradation behaviors due to the centrifugal force [39], usually only the most needed rail is replaced. Rail replacement is performed using welds to connect the new rail with the old one. After replacement, the rail surface defects will totally disappear by the installation of new rail whereas development of new squats will depend on various factors, like track conditions, MGT, and other different

factors. In the case of the welds, because they are composed by materials with different properties than the rails, they are prone to squat defect appearance [40]. Figure 2.7a and 2.7b show squat growth before and after rail replacement. Figure 2.7a shows the squat growth between welds where all the squats will disappear after replacement. The model assumes that no squats will appear during a long horizon by considering that new developed squats can be detected in the next measurement campaign. Figure 2.7b shows squat growth on the welds in a period after replacement. The exact time instant when the growth starts is related to the quality of the weld. This means that for those welds that have good quality, the starting point would be much later. If squat is positioned between two welds, then the squat length after the rail replacement should be equal to zero during a time horizon N_1. The growth model on the weld can be expressed according to the time N_2 when squat can appear. Before time $k+ N_2$ no squat is present in the weld, while at $k+ N_2+1$ the squat will start to appear and evolved based on the proposed growth scenarios.

$$\hat{L}_i^h(x_{w_1}, k+t) = 0 \qquad t = 1,2,...,N_1, h = h_1, h_2, h_3 \tag{2.11}$$
$$\hat{L}_i^h(x_{w_2}, k+t) = 0 \qquad t = 1,2,...,N_2, h = h_1, h_2, h_3$$

$$\hat{L}_i^h(x_{w_2}, k+N_2+1) = \begin{cases} \underline{f}_i^{TS}(\Delta L_i) & if \quad h = h_1 \\ f_i^{TS}(\Delta L_i) & if \quad h = h_2 \\ \overline{f}_i^{TS}(\Delta L_i) & if \quad h = h_3 \end{cases} \tag{2.12}$$

where x_{w_1} is some position between the welds, x_{w_2} is the location of the weld, and ΔL_i is small value that triggers the growth when the squat i starts evolving at the thermite weld at time instant $k+ N_2+1$. After the squat appears, the interval fuzzy model will capture its evolution over time.

2.4 KPIs for rail health condition

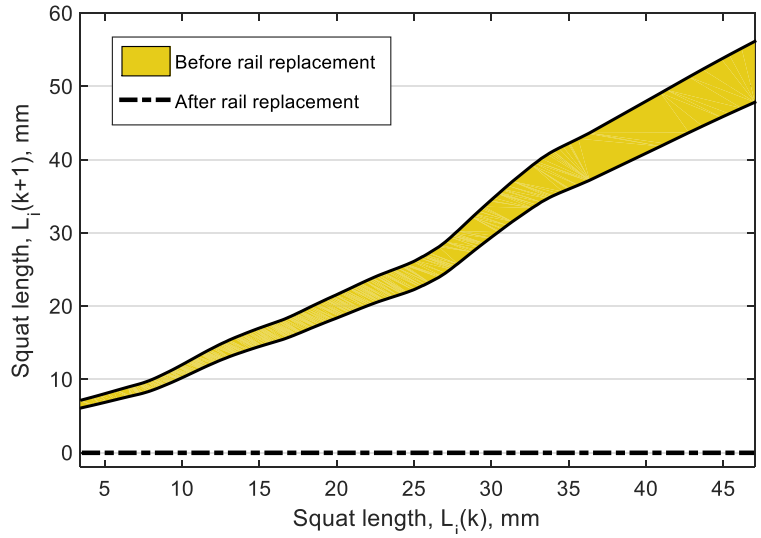
2.4.1 KPI description

The monitoring of the evolution of a single squat might not be practical from the maintenance perspective. Aggregated information over bigger track partitions can facilitate infrastructure manager decisions over the maintenance plans. In the case of squats, the authors propose key performance indicators (KPI's) considering the number of *A*, *B* and *C* squats and the number of squats with potential risk of rail break called *RC* squats, at different time *t* and different growth scenario *h*. Moreover, as significant number of *B* and *C* squats near to each other indicate a high potential risk to track safety, a KPI is proposed relying on a measure of density of squats *B* and *C*.

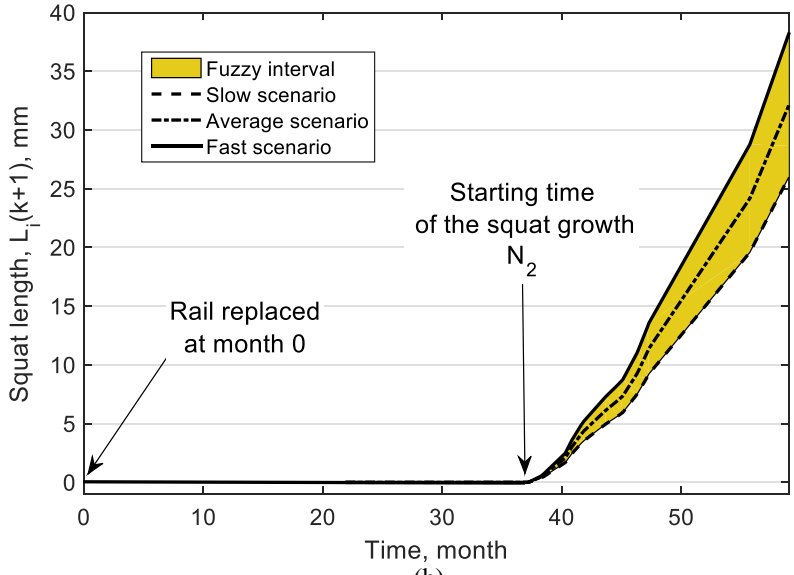
The function $\delta_{h,j}^d(x,k)$ is provided by the ABA detection algorithm, for the current instant of measurement *k*. The function equals to 1 if a squat type $d \in \{\mathrm{A,B,C,RC}\}$ is located at position *x*, instant *k*, partition *j* and growth scenario *h* and equals to zero otherwise. Used as initial condition, and relying on the interval fuzzy model, it is possible to predict $\delta_{h,j}^d(x,t)$ for any time horizon, $t=1,...,N_P$. The growth of new squats during the prediction horizon is not considered in this work, because it is assumed that new squats will be detected in the next measurement

campaign at instant *k+1*, where the models can be updated according to the new conditions. The KPIs of squat numbers at partition *j*, instant *t*, scenario *h*, can be expressed as:

$$\begin{aligned} y_{h,j}^{A}(t) &= \sum_{x\in\left[x_j,x_{j+1}\right)} \delta_{h,j}^{A}(x,t) \\ y_{h,j}^{B}(t) &= \sum_{x\in\left[x_j,x_{j+1}\right)} \delta_{h,j}^{B}(x,t) \\ y_{h,j}^{C}(t) &= \sum_{x\in\left[x_j,x_{j+1}\right)} \delta_{h,j}^{C}(x,t) \\ y_{h,j}^{RC}(t) &= \sum_{x\in\left[x_j,x_{j+1}\right)} \delta_{h,j}^{RC}(x,t) \end{aligned} \tag{2.13}$$



(a)



(b)

Figure 2.7 (a) After rail replacement with a piece of new rail free of damage, the length of squats $L_i(k+1)$ will become zero no matter their initial length $L_i(k)$; (b) on welds after rail replacement a squat is prone to appear.

Also, to estimate the density of *B* and *C* squats $d_{h,j}^{BC}(x,t)$, a window is defined around the coordinate x (in this chapter, the window is 50 m in track length). The function $d_h^{BC}(x,t)$ equals the number of squats B or C in the moving window $[x-0.025, x+0.025]$. The KPI density for partition *j*, instant *t*, and scenario *h* can be defined as the area of the density function as follows:

$$y_{h,j}^{dBC}(t) = \frac{\sum_{x \in [x_j, x_{j+1})} d_h^{BC}(x,t)}{x_{j+1} - x_j} \tag{2.14}$$

A vector containing all the KPIs called $y_{h,j}(t)$ for partition *j*, instant *t*, and scenario *h* is defined as follows:

$$y_{h,j}(t) = \left[y_{h,j}^{A}(t), y_{h,j}^{B}(t), y_{h,j}^{C}(t), y_{h,j}^{RC}(t), y_{h,j}^{dBC}(t) \right]^T \tag{2.15}$$

where $y_{h,j}^{A}(t)$, $y_{h,j}^{B}(t)$, $y_{h,j}^{C}(t)$, $y_{h,j}^{RC}(t)$ and $y_{h,j}^{dBC}(t)$ are the number of *A* squats, *B* squats, *C* squats, *RC* squat and the density of *B* squats *C* squats, respectively. Due to the large number of KPI's obtained in terms of all the growth scenarios and predictions over time, the authors propose two simple steps to include the effect of the trajectories of the KPIs into one global KPI:

Step 1: First, transform the vector $y_{h,j}(t)$ for each partition *j*, scenario *h* and instant *t*, into a single KPI using a fuzzy expert system $y_{h,j}^{M}(t) = f_{\text{Mamdani}}\left(y_{h,j}^{A}(t), y_{h,j}^{B}(t), y_{h,j}^{C}(t), y_{h,j}^{RC}(t), y_{h,j}^{dBC}(t) \right)$.

Step 2: Then, aggregate the single KPI over the set of scenarios and over the prediction horizon, for each partition *j*. This results into a single global KPI for the current instant *k*, $J_j^{Rail}(k)$:

$$J_j^{Rail}(k) = f_{\text{aggregate}}\left(y_{h_1,j}^{M}(k),...,y_{h_1,j}^{M}(k+N_P),....,y_{h_H,j}^{M}(k),...,y_{h_H,j}^{M}(k+N_P) \right) \tag{2.16}$$

2.4.2 Mamdani fuzzy KPI

For Step 1, a Mamdani fuzzy expert system is used to calculate a single KPI [41]. Even though the Mamdani fuzzy system approach was proposed more than 40 years ago, it is still popular because of its simplicity and interpretability [42]; [43]; [44]. In this case, 32 fuzzy if-then rules are generated. The aim is to assign a membership degree to each KPI to represent the contribution of each KPI in the rail health condition:

$$\begin{aligned} &If\ y_{h,j}^{A}(t)\ is\ A_1^r\ and\ \ y_{h,j}^{B}(t)\ is\ A_2^r\ and\ y_{h,j}^{C}(t)\ is\ A_3^r\ and\ y_{h,j}^{RC}(t)\ is\ A_4^r\ and\ y_{h,j}^{dBC}(t)\ is\ A_5^r \\ &\qquad\qquad then\ y_{h,j}^{M}(t)\ is\ G^r \end{aligned} \tag{2.17}$$

where A_1^r, A_2^r, A_3^r, A_4^r, A_5^r and G^r are the membership functions for rule *r* and $y_{h,j}^{M}(t)$ is the output Mamdani KPI. The KPIs are first normalized, then Gaussian membership functions are used to fuzzify the KPIs. Also, to defuzzify, centre of gravity method is applied in order to obtain crisp value at the end. Furthermore, relying on the fuzzy rules, interdependency of KPIs and Mamdani KPI are captured as shown in Figure 2.8. In this figure, it is presented how Mamdani

KPI models the influence in the health of the track of two KPIs, varying from fully healthy (equals to zero) to completely unhealthy (equals to one), while all the other KPIs are assumed to be fully healthy (equals to zero). Four plots are presented. In Figure 2.8(a), a higher value for the *BC* density is much relevant than the contribution of the number of *B* squat. In Figure 2.8(b), a high number of *C* squats makes the most significant impact on the rail health condition. The rail condition will get highly unhealthy with high values of either density of the *BC* squats or number of *C* squat. In Figure 2.8(c), a high number of *RC* squats will influence much strongly on the health condition than the number of *A* squats. In the last plot, Figure 2.8(d), a high number of *A* squats or *B* squats will not have strong influence in the short term (the condition moves between the values 0.28 to 0.37). However, the number of *B* squats effects more negatively the rail health condition than the number of *A* squats. Figure 2.8, shows the intuitive fact that rail condition gets worse with the increasing number of squats from *A*, *B*, *C* to *RC*.

In general, the number of *A* squats will not have significant impact on the current rail health condition. However, in the long term, if not ground, A squats will evolve into severe defects. In order to capture this and other dynamic effects, the prediction model is used, and the global KPI is calculated over time and under different scenarios.

2.4.3 Fuzzy global KPI

Relying on defined Mamdani KPIs $y_{h,j}^{M}(t)$, a fuzzy global indicator is calculated to give a KPI over growth scenarios in partition *j*:

$$J_j^{Rail}(k) = \frac{\sum_{h\in\{h_1,h_2,h_3\}} \sum_{t=k}^{k+N_p} w_h.w_t.y_{h,j}^{M}(t)}{\sum_{h\in\{h_1,h_2,h_3\}} \sum_{t=k}^{k+N_p} w_h.w_t} \tag{2.18}$$

where $J_j^{Rail}(k)$ is fuzzy global indicator, w_h is the weight for scenario given by the infrastructure manager to indicate the improtance of each scenario. Moreover, to take the exponential effect of time into account w_t is defined. In this way, the authors aggregate different KPIs into a single one, that captures together stochasticity and evolution over time.

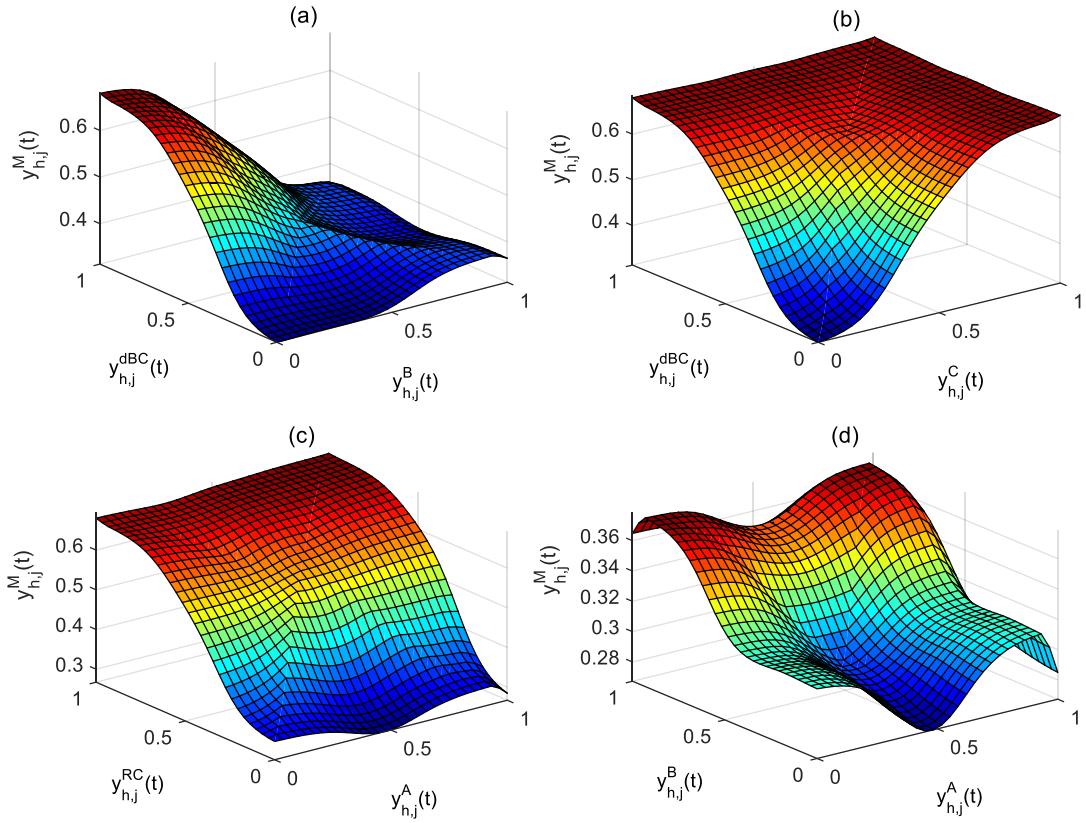


Figure 2.8 Interdependency of KPIs over Mamdani KPI, $y^{M}_{h,j}(t)$.

2.5 Numerical results

2.5.1 Fuzzy confidence interval

This section summarizes the simulation results to predict the squats length. A data set of squat lengths collected from different track visits are used to evaluate the performance of the squat growth model. Identification data and validation data for the interval fuzzy TS model are selected randomly, using 60% of the data for identification and 40% for validation (see Figure 2.9).

To optimize the number of clusters, models from two to ten clusters are tested. For each number of clusters, the root mean square (RMS) of the prediction error is used to determine the best model. During the training, the tuning parameters of the confidence interval fuzzy model are considered the same for the lower and upper fuzzy bounds. The idea is to obtain the optimum parameter α using a band as confidence interval. The results should be based on a minimum number of data points outside the band whereas the band is as narrow as possible. Figure 2.10(a) depicts the Pareto front of the normalized area of the band versus the normalized number of data points outside the band ranging α from 0 to 40. Figure 2.10(b) shows how α behaves in terms of area of the band. As shown in Figure 2.10(b), the area will reach a maximum value if α equals 32. In reality, the variance of the worst case scenario is much larger than the best case scenario; thus the assumption of a fixed α must be relaxed. Using full trajectories of different squats, ad-hoc α^{h_1} and α^{h_3} were obtained to better fit the dynamics. The use of the interval fuzzy

model for prediction is presented in Figure 2.11, with α=1.5 and modified parameters $\alpha^{h_1} = 0.32 \cdot \alpha$, and $\alpha^{h_3} = 1.7 \cdot \alpha$. The squat length grows from a small defect of 8 mm to a severe squat of 60 mm. An important characteristic is when the prediction model reaches the highest bound 60 mm. This happens for squats of 48 mm for the one-step ahead prediction (within 6 months), and it will happen for squats of 18 mm in the case of four-step ahead prediction (within 24 months). For testing purposes, the authors have evaluated this model with another data set of the trajectories presented in [45]. All of them are contained within the interval model.

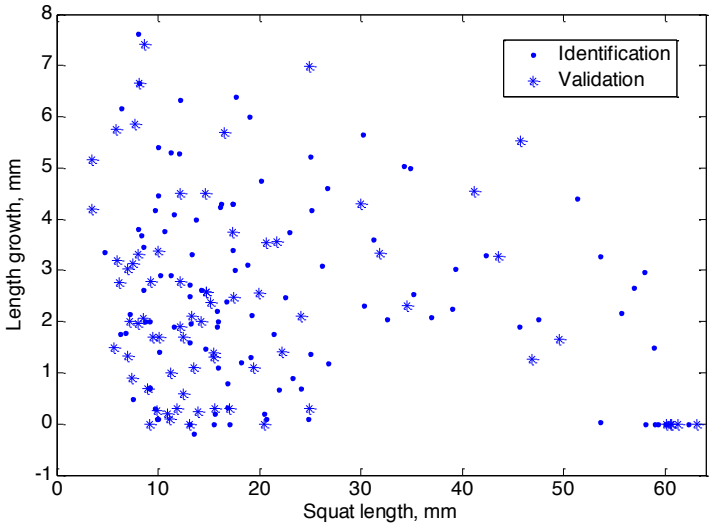


Figure 2.9 Validation and identification data for the squat length.

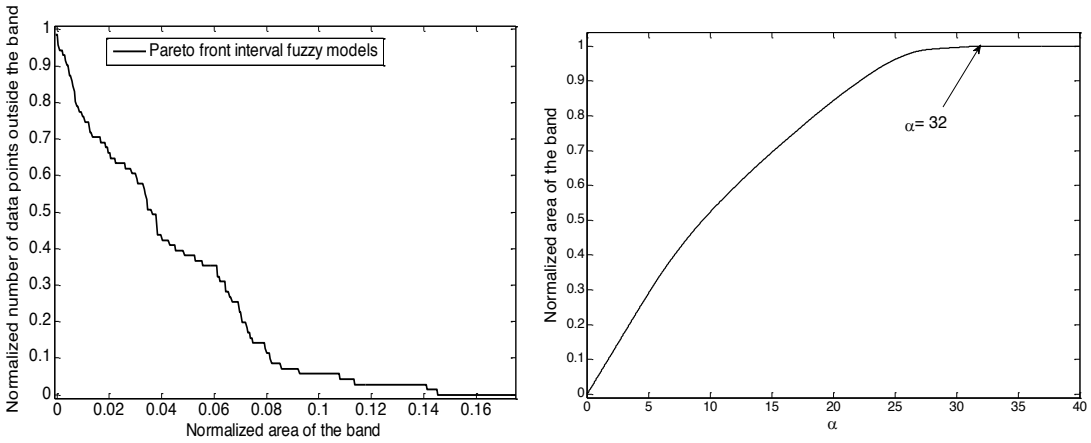


Figure 2.10 (a) pareto front of number of data point outside v.s area of the band. (b) area of the band over α.

2.5.2 Fuzzy global KPI for track health condition

The full track of the Meppel-Leeuwarden is used to illustrate the proposed methodology. The Figure 2.12 shows a simple map of the track and the four partitions j_1, j_2, j_3 and j_4. The partitions can be adapted according to the maintenance plans or other design considerations. The partitions in this chapter are all around 10 kilometres long, except the last one which is 15 kilometres long. Meppel is at kilometre 105, Leeuwarden is at 150, the partitions are defined between the kilometres: $x_{j_1} = 105$, $x_{j_2} = 115$, $x_{j_3} = 125, x_{j_4} = 135$ and $x_{j_5} = 150$.

Figure 2.13 shows the different KPIs regarding squats number over four step-ahead prediction when no maintenance is performed. All the cases are calculated for the scenarios slow growth (in blue), average growth (in yellow) and fast growth (in red). In Figure 2.13(a), the number of *A* squats tends to get reduced over time, as they are becoming *B* squats. In Figure 2.13(b), the number of *B* squat increases because of the *A* squats becoming *B* squats, but after t=12, the number of *B* squat decreases as most of them are becoming *C* squat. When no corrective maintenance is performed, it can be seen from Figure 2.13(c) that after t=12, a huge number of C squats are in the track (worst case scenario), which is a very expensive situation as the only solution will be to replace the rails. In Figure 2.13(d), it is possible to see the moment when operational risk locations start to appear, indicating that maintenance should be done before the worst case scenario indicates their appearance.

Figure 2.14(a) shows how potential risk squats will start to appear over time. Figure 2.14b shows the KPI related to density of *B* and *C* squats. As seen in Figure 2.14(a), the first squats with high potential risk of derailment, *RC* squats, appear for the worst case scenario at t=12, in four kilometer positions $\{130.9,132.0,132.5,133.0\}$. Three of those four locations were already detected at t=0 in Figure 2.14(b), while all of them are already present in the B-C squat density signal at t=6 for all the scenarios. It means that within the first 12 month, the infrastructure manager is expected to take actions, to prevent risk of derailment.

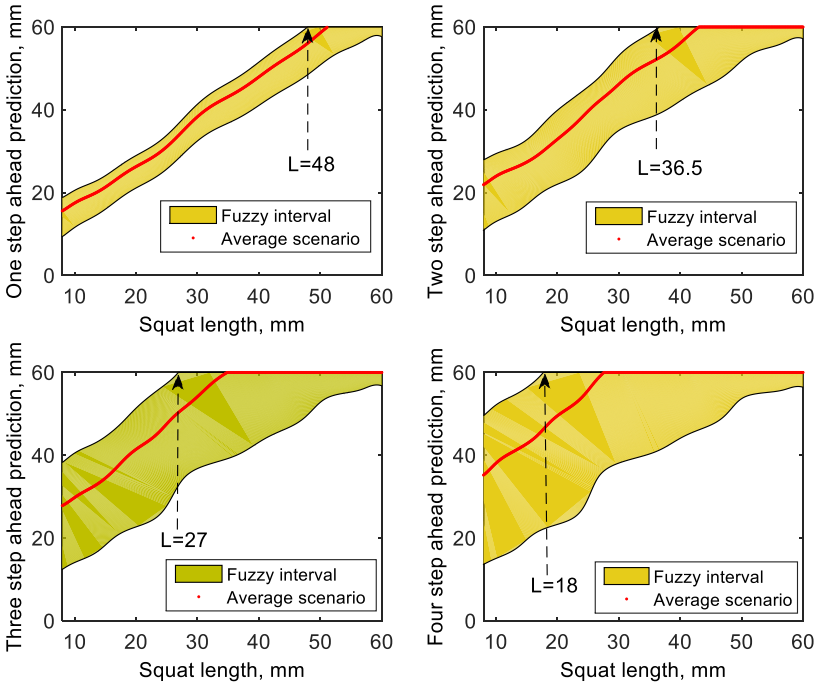


Figure 2.11 Interval fuzzy model predictions, one, two, three and four steps-ahead.

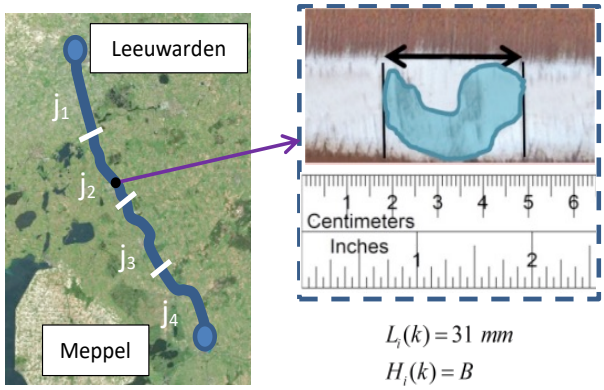


Figure 2.12 Schematic track map between two stations, Meppel and Groningen, divided into four partitions, j_1, j_2, j_3, and j_4

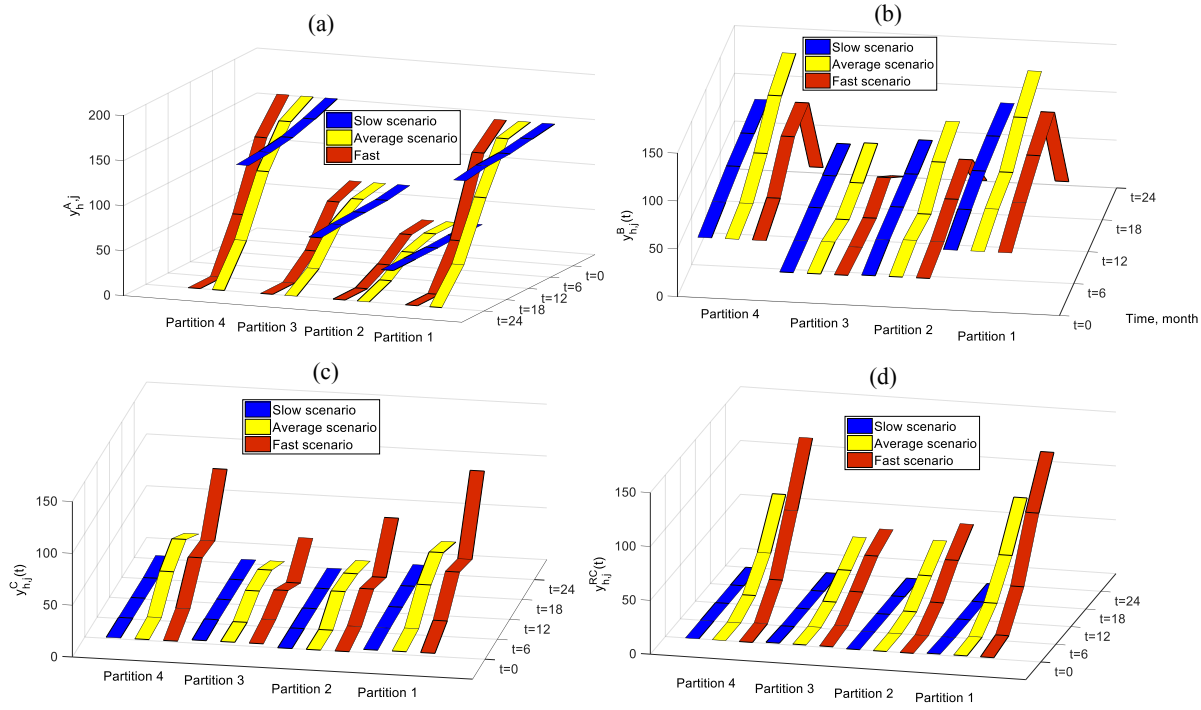


Figure 2.13 Squat number KPIs for the slow, average and fast growth scenarios in the absence of maintenance operation, (a) number of *A* squat, (b) number of *B* squat, (c) number of *C* squat and (d) number of *RC* squats.

Figure 2.15 collects all the scenarios and the signals over the whole prediction horizon, to indicate a single global fuzzy KPI for each track partition. Three cases are considered, no maintenance, grinding at t=0, and local rail replacement at t=0 for each severe squat. Maintenance considerably can improve the rail health condition, but to be fully efficient a combination of both grinding and replacement is necessary.

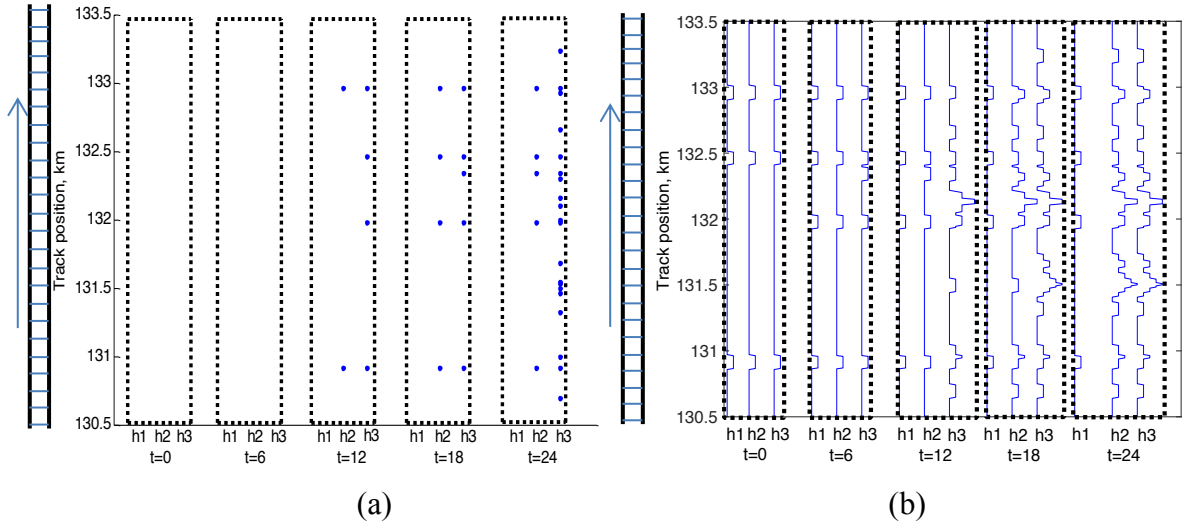


Figure 2.14 For track position between 130.5 and 133.5km, predictions over 24 months and three scenarios for: (a) Potential risk locations, (b) *B*-*C* squats density.

After the maintenance operations, the condition is in the average condition range, where the potential risk of derailment is considerably lower during the prediction horizon. The following result allows the infrastructure manager to decide how to manage the rail in the future at each track partition. In the absence of maintenance operation, a cost of zero euro with the clear consequence of the bad rail health condition is expected. In the case of the grinding effect and the replacement effect, the results can be applied as an effective factor for cost analysis of the track maintenance plan.

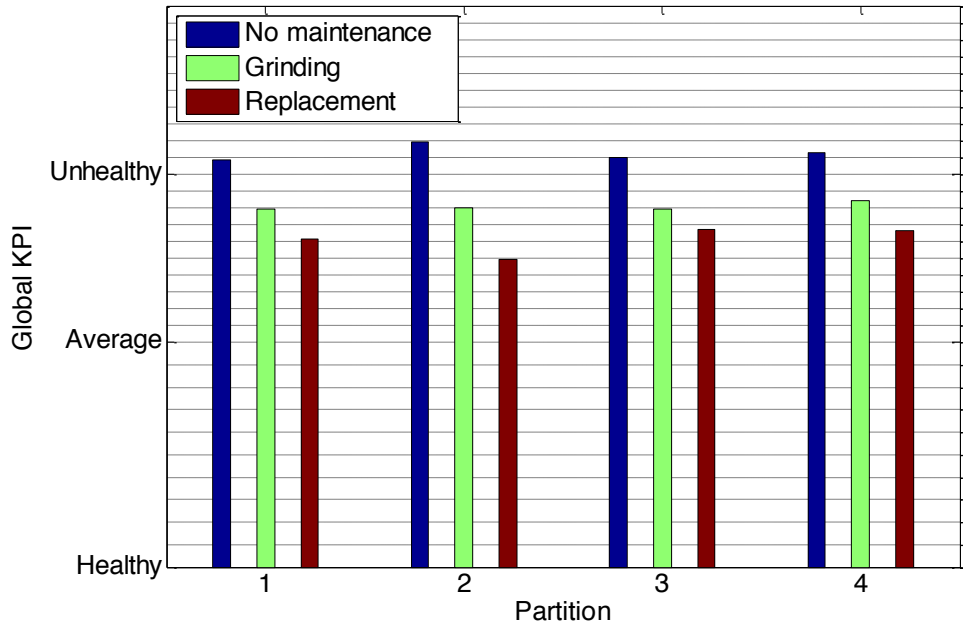


Figure 2.15 Fuzzy global KPIs

2.6 Conclusion and future research

In this chapter, a condition-based monitoring methodology is developed for a type of surface defect in the rail called “squats”. This methodology is employed to construct an interval-based TS fuzzy prediction modelling in order to monitor the track condition over maintenance time horizon per track partition.

The idea of using fuzzy interval is to capture all the possible growth scenarios. Based on the interval fuzzy models for squats, a condition-based methodology for railway tracks is proposed using different KPIs defined in a track-partition level, allowing the grouping of defects located in a given track partition. In the methodology, number and density of squats are considered over a prediction horizon under three different scenarios, slow, average and fast growth. Then, to facilitate visualization of the rail health condition and to ease the maintenance decision process, the authors propose a fuzzy global KPI based on fuzzy rules for each partition, that combine the different KPIs over prediction horizon and scenarios. Hence, the proposed methodology adds value by defining fuzzy global KPIs which are predictable over time to facilitate maintenance decision making of the rail. As an example, the KPIs obtained are presented for the track Meppel-Leeuwarden.

As a further research, the study will be oriented into an optimization-based methodology to reduce life cycle costs effectively and to fit the methodology much closely to the real-life maintenance operations. The use of new predictive and robust KPIs defined for different parties will be considered, including infrastructure manager, rolling stock manager, contractors and users.

2.7 References

[1] Zoeteman, A. (2001). Life cycle cost analysis for managing rail infrastructure. *European Journal of Transport and Infrastructure Research*, *1*(4), 391-413.

[2] Stenström, C., Parida, A., Lundberg, J., & Kumar, U. (2015). Development of an integrity index for benchmarking and monitoring rail infrastructure: application of composite indicators. *International Journal of Rail Transportation*, *3*(2), 61-80.

[3] Tyrrell Rockafellar, R., & Royset, J. O. (2015). Engineering decisions under risk averseness. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *1*(2), 04015003.

[4] Parida, A., & Chattopadhyay, G. (2007). Development of a multi-criteria hierarchical framework for maintenance performance measurement (MPM). *Journal of Quality in Maintenance Engineering*, *13*(3), 241-258.

[5] Åhrén, T., & Parida, A. (2009). Maintenance performance indicators (MPIs) for benchmarking the railway infrastructure: a case study. *Benchmarking: An International Journal*, *16*(2), 247-258.

[6] Li, Z., Molodova, M., Núñez, A., & Dollevoet, R. (2015). Improvements in axle box acceleration measurements for the detection of light squats in railway infrastructure. *IEEE Transactions on Industrial Electronics*, *62*(7), 4385-4397.

[7] Zoeteman, A., & van Meer, G. (2006). A Yardstick for Condition-Based and Differentiated Planning of Track and Turnout Renewal: A Major Step towards Full Decision Support. *Proceedings 7th World Congress on Railway Research,* 4-8.

[8] Zoeteman, A., Dollevoet, R., & Li, Z. (2014). Dutch research results on wheel/rail interface management: 2001–2013 and beyond. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *228*(6), 642-651.

[9] Patra, A. P., Söderholm, P., & Kumar, U. (2009). Uncertainty estimation in railway track life-cycle cost: a case study from Swedish National Rail Administration. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *223*(3), 285-293.

[10] Stenström, C., Norrbin, P., Parida, A., & Kumar, U. (2016). Preventive and corrective maintenance–cost comparison and cost–benefit analysis. *Structure and Infrastructure Engineering*, *12*(5), 603-617.
[11] Sadeghi, J., & Askarinejad, H. (2010). Development of improved railway track degradation models. *Structure and Infrastructure Engineering*, *6*(6), 675-688.
[12] Andrade, A. R., & Teixeira, P. F. (2011). Uncertainty in rail-track geometry degradation: Lisbon-Oporto line case study. *Journal of Transportation Engineering*, *137*(3), 193-200.
[13] Andrade, A. R., & Teixeira, P. F. (2012). A Bayesian model to assess rail track geometry degradation through its life-cycle. *Research in Transportation Economics*, *36*(1), 1-8.
[14] Andrews, J. (2013). A modelling approach to railway track asset management. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *227*(1), 56-73.
[15] Vale, C., & Lurdes, S. M. (2013). Stochastic model for the geometrical rail track degradation process in the Portuguese Railway Northern Line. *Reliability Engineering & System Safety*, *116*, 91-98.
[16] Nathanail, E. (2014). Framework for monitoring and assessing performance quality of railway network infrastructure: Hellenic Railways case study. *Journal of Infrastructure Systems*,*20*(4), 04014019.
[17] Guler, H. (2014). Prediction of railway track geometry deterioration using artificial neural networks: a case study for Turkish state railways. *Structure and Infrastructure Engineering*, *10*(5), 614-626.
[18] Weston, P., Roberts, C., Yeo, G., & Stewart, E. (2015). Perspectives on railway track geometry condition monitoring from in-service railway vehicles. *Vehicle System Dynamics*, *53*(7), 1063-1091.
[19] Schafer, D. H., & Barkan, C. P. (2008). A prediction model for broken rails and an analysis of their economic impact. *American Railway Engineering and Maintenance of Way Association (AREMA) Annual Conference*, Salt Lake City. September 2018.
[20] Burstow, M. C., Watson, A. S., & Beagles, M. (2002). Application of the whole life rail model to control rolling contact fatigue. In *Proceedings of the International Conference Railway Engineering 2002,* 3-4 July, London, UK.
[21] Sandström, J., & Ekberg, A. (2009). Predicting crack growth and risks of rail breaks due to wheel flat impacts in heavy haul operations. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *223*(2), 153-161.
[22] Senouci, A., El-Abbasy, M. S., & Zayed, T. (2014). Fuzzy-based model for predicting failure of oil pipelines. *Journal of Infrastructure Systems*, *20*(4), 04014018.
[23] Sadiq, R., Rajani, B., & Kleiner, Y. (2004). Fuzzy-based method to evaluate soil corrosively for prediction of water main deterioration. *Journal of Infrastructure Systems*, *10*(4), 149-156.
[24] Xu, J., Tu, Y., & Lei, X. (2013). Applying multi-objective bi-level optimization under fuzzy random environment to traffic assignment problem: Case study of a large-scale construction project. *Journal of Infrastructure Systems*, *20*(3), 05014003.
[25] Wang, K., & Liu, F. (1997). Fuzzy set-based and performance-oriented pavement network optimization system. *Journal of infrastructure systems*, *3*(4), 154-159.
[26] Khatri, K. B., Vairavamoorthy, K., & Akinyemi, E. (2011). Framework for computing a performance index for urban infrastructure systems using a fuzzy set approach. *Journal of Infrastructure Systems*, *17*(4), 163-175.
[27] Škrjanc, I., Blažič, S., & Agamennoni, O. (2005). Identification of dynamical systems with a robust interval fuzzy model. *Automatica*, *41*(2), 327-332.
[28] Núñez, A., & De Schutter, B. (2012). Distributed identification of fuzzy confidence intervals for traffic measurements. *51st IEEE Annual Conference on Decision and Control (CDC)*, 6995-7000, 10-13 December, California, USA.
[29] Sáez, D., Ávila, F., Olivares, D., Cañizares, C., & Marín, L. (2015). Fuzzy prediction interval models for forecasting renewable resources and loads in microgrids. *IEEE Transactions on Smart Grid*, *6*(2), 548-556.
[30] Molodova, M., Li, Z., Núñez, A., & Dollevoet, R. (2014). Automatic detection of squats in railway infrastructure. *IEEE Transactions on Intelligent Transportation Systems*, *15*(5), 1980-1990.
[31] Li, Z., Molodova, M., Núñez, A., & Dollevoet, R. (2015). Improvements in axle box acceleration measurements for the detection of light squats in railway infrastructure. *IEEE Transactions on Industrial Electronics*, *62*(7), 4385-4397.
[32] Li, Z., Zhao, X., Esveld, C., Dollevoet, R., & Molodova, M. (2008). An investigation into the causes of squats—Correlation analysis and numerical modeling. *Wear*, *265*(9-10), 1349-1355.
[33] Molodova, M., Li, Z., Núñez, A., & Dollevoet, R. (2015). Parametric study of axle box acceleration at squats. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *229*(8), 841-851.
[34] Smulders, J. (2003). Management and research tackle rolling contact fatigue. *Railway Gazette International*, *158(7), 439–442*.
[35] UIC Code, (2002). Rail Defects. *International Union of Railways, 4th edition*, Paris, France.

[36] Rail Damages, (2001). *The Blue Book of RailTrack*. U.K.
[37] Li, Z., Dollevoet, R., Molodova, M., & Zhao, X. (2011). Squat growth—Some observations and the validation of numerical predictions. *Wear*, *271*(1-2), 148-157.
[38] Škrjanc, I. (2011). Fuzzy confidence interval for pH titration curve. *Applied Mathematical Modelling*, *35*(8), 4083-4090.
[39] Magel, E. E., & Kalousek, J. (2002). The application of contact mechanics to rail profile design and rail grinding. *Wear*, *253*(1-2), 308-316.
[40] Lewis, R., & Olofsson, U. (2009). Basic tribology of the wheel–rail contact. In *Wheel–rail interface handbook*, 34-57.
[41] Mamdani, E. H., & Baaklini, N. (1975). Prescriptive method for deriving control policy in a fuzzy-logic controller. *Electronics Letters*, *11*(25), 625-626.
[42] Camastra, F., Ciaramella, A., Giovannelli, V., Lener, M., Rastelli, V., Staiano, A., & Starace, A. (2015). A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference. *Expert Systems with Applications*, *42*(3), 1710-1716.
[43] Mohammad, R., Mostafa, A., Abbas, M., & Farouq, H. M. (2015). Prediction of representative deformation modulus of longwall panel roof rock strata using Mamdani fuzzy system. *International Journal of Mining Science and Technology*, *25*(1), 23-30.
[44] Kisi, O. (2013). Applicability of Mamdani and Sugeno fuzzy genetic approaches for modeling reference evapotranspiration. *Journal of Hydrology*, *504*, 160-170.
[45] Li, Z., Molodova, M., Zhao, X., & Dollevoet, R. (2010). Squat treatment by way of minimum action based on early detection to reduce life cycle costs. *Joint Rail Conference (JRC),* 305-311, 27-29 April, Urbana, USA.

A big data analysis approach for rail failure risk assessment

This chapter corresponds to the reference: *A. Jamshidi, S. Faghih-Roohi, S. Hajizadeh, A. Núñez, R. Babuška, R. Dollevoet, Z. Li and B. De Schutter, "A big data analysis approach for rail failure risk assessment". Risk Analysis, Volume 37, Issue 8, August 2017, Pages: 1495-1507. DOI: 10.1111/risa.12836*

3.1 Introduction

Among all transportation infrastructure, the railway network is one of the most successful transport systems for reducing transportation cost, traffic congestion, and air pollution emission levels. On the one hand, the increase in usage of the railway network requires a systematic monitoring plan to keep the trains running in a safe way as well as with the least possible disruptions [1]. On the other hand, a large amount of data is collected by frequent measurements from the monitoring systems of the infrastructure and the assets involved in the railway operations. This data should be controlled, stored, and processed, such that it can be employed to take all necessary actions to guarantee the rail asset quality level desired by the infrastructure manager [2]. The large amount of data should be processed into actionable knowledge within a certain time period [3].

Risk is intuitively connected to decision making under uncertainty [4]. Recent developments in big data analytics for uncertainty management and risk assessment of

industrial systems have been studied on [5], and [6]. Risk assessment of large-scale systems is of current interest across many application domains such as healthcare [7], environmental safety [8]; [9], transportation [10]; [11]; [12]; [13], business [14], and product development [15]. In particular for railway applications, risk assessment is critical for the prediction of infrastructure health condition within a given time period. Continuous monitoring of railway systems can guarantee the availability of data that can be used to assess the risk of infrastructure failures. Also, the database constructed from continuous monitoring of data will become larger and larger over time. Thus, applying a big data analysis approach is necessary in order to adequately monitor the infrastructure condition [16]. The rail track is an important infrastructure in a railway network and should be taken into account in terms of budget allocation [17]. As a high percentage of failures occurring in the railway infrastructure is directly related to the rail, it is important to assess the failure risk of rails. The rail risk assessment involves detecting the rail defects that can potentially result in rail break and derailment in extreme cases [18]; [19]; [20]. Rail surface defects are caused by different factors such as fatigue due to a large number of trains passing over rail components at especially welds, joints, and switches [21]. Early detection of surface defects is important to mitigate disastrous consequences of rail breaks. There are different methods to diagnose the condition of rail defects, including ultrasonic measurements [22], eddy current testing [23], and guided-wave based monitoring [24]. In general, these methods are not able to detect defects in an early stage of growth, i.e. not until the defects are severe. In particular, detection of defects at the late stage of growth imposes extra operation and maintenance costs due to the fact that the only solution is to replace the rail.

To address the limitations of the current measurement methods, the use of video cameras installed on trains has become popular [25]; [26]. The use of video cameras avoids the error-prone, costly, and time-consuming process of manual rail monitoring. Moreover, the videos taken from side cameras enable the infrastructure manager to capture the real condition of other track components such as fasteners, switches, and sleepers. Using video cameras, one can simply monitor whether the visible defects are at the early or late stage of growth. This means that the infrastructure manager has the opportunity to observe how the defect evolves over time in order to take actions at the right moment and to focus on the most urgent places for maintenance operations. This can lead to a significant reduction in the operation costs induced by the defects and it can prevent potential risks of rail breaks, reducing the risk of derailment. Due to the large amount and the high resolution of the videos taken over the rail, an automatic detection algorithm is required to process the huge number of images from those videos. The main contribution of this chapter is to assess rail failure risk based on an integrated framework that merges the information of two defect-related variables: visual length and crack growth. There is no similar approach in the literature for risk assessment of rail failure that considers both variables. This is due to the fact that in this case a big data analysis problem has to be faced, as a result of which usually railway maintenance managers look at only one type of data and ignore the other influencing factors. We propose a risk function (equation1) as a composition of three functions: the probability function, the crack growth function, and the partially inversed severity function. To evaluate these functions, we apply several techniques, including a deep convolutional neural network for image processing and defect detection, an N-step ahead prediction model for defect severity and crack growth analysis, and a Bayesian inference model for failure probability estimation. To implement our proposed framework, a particular type of surface defect in railway networks called squat is considered in the case study. Furthermore, we give a proposed classification of the squats in terms of the visual length. Thus, squats are classified according to different severities. These classes can be used later for condition-based maintenance where we have different maintenance operations for different stages of the growth (rail grinding for light squats and replacement for severe squats).

However, our approach can be generalized and applied for similar cases when there is a need to analyze a huge amount of image data for assessment of failure probability and risk function. For example, in a recent work by [28] satellite images have been employed to assess flood hazard risk. Moreover, in the field of health science, abnormality detection using image processing has become very popular [29]. There are many cases in the literature where image data is used to deal with risk assessment problem [28]; [29]; [30]; [31]; [32] In all these cases, as long as the focus is to detect abnormalities and failures among a big database of images, the risk assessment approach proposed in the current chapter is applicable for merging attained information from images.

This chapter is organized as follows. In Section 3.2, the proposed failure risk assessment model is presented including the model framework. Section 3.3 addresses a real-life case study of the Dutch railway network. Section 3.4 presents the results and discussions. Finally, in Section 3.5, conclusions are presented.

3.2 Failure risk assessment model

3.2.1 The proposed framework

In this section, we propose a failure risk framework for analysing the rail surface defects. The proposed framework is depicted in Figure 3.1. Video images, ultrasonic detection [22], and eddy current testing [23] can all be used to detect the defects that can lead to rail break. In this chapter, we rely on both ultrasonic detection method and video images. On the one hand, with ultrasonic measurement, we derive a general characteristic of crack growth. One the other hand, with video image, we analyze the growth of the visual length of defects which are detected among huge number of rail images. Then, a sample of the visual length of the detected defects is chosen for the assessment of failure risk model. The approach can be employed for any type of rail defects.

In this framework, a large amount of image data is automatically processed by a deep convolutional neural network to detect squats in Step 1 (see details in Section 3.4). The visual lengths of defects are measured from the defect detected from the video images, and then used for defect severity analysis in Step 2 (see details in Section 3.2).

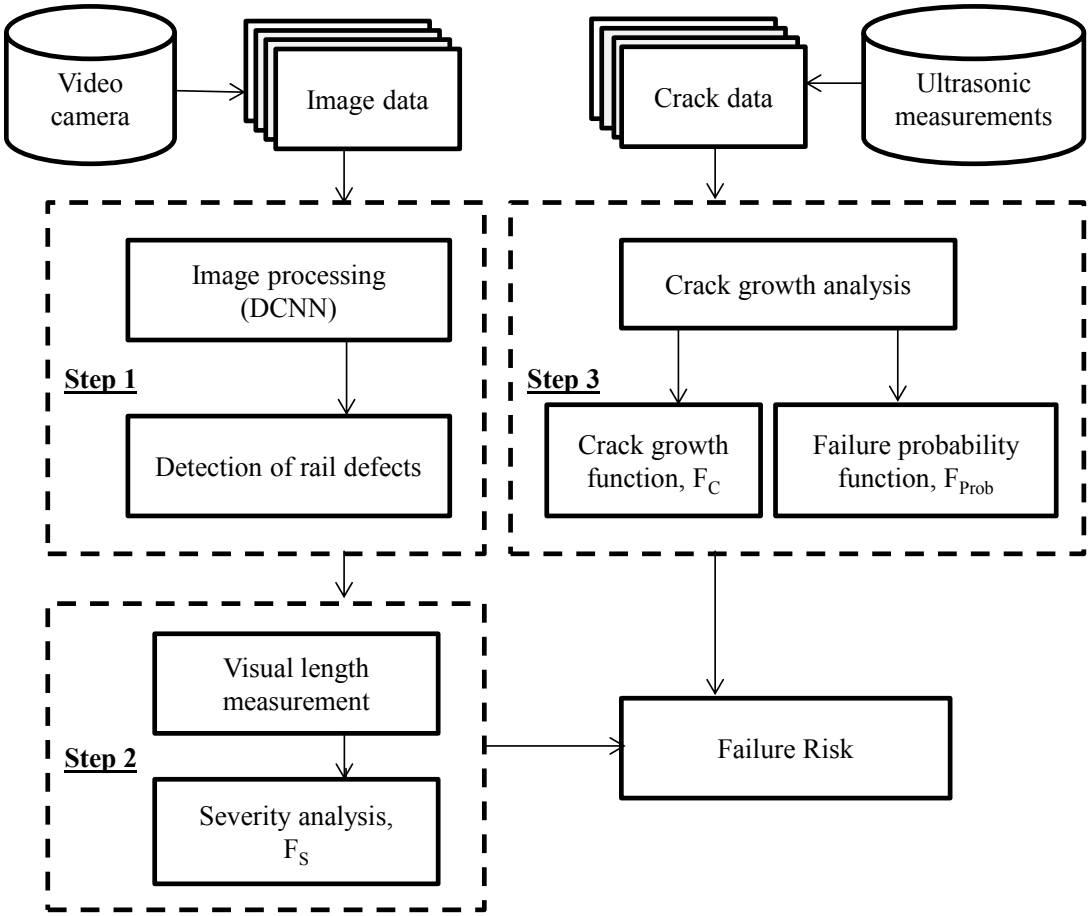


Figure 3.1 Flowchart of the proposed methodology

In Step 3, a crack growth analysis is performed to estimate the crack growth as a function of Million Gross Tons (MGT) by using the data from ultrasonic measurements (see details in Section 3.2.3). In addition, the probability of rail failure as a function of crack growth is estimated using the crack growth data. Finally, we propose to assess the risk of rail failure with the composition of the probability function, the crack growth function and the partial inverse of severity function:

$$\text{Risk} \propto F_{\text{Prob}}\left(F_C\left(F_{S,inv.}\left(V_1, V_2\right)\right)\right) \tag{3.1}$$

where V_1 and V_2 are two consecutive measurements of visual length for a defect, detected by analysis of image data, and $F_{S,inv.}$ relates V_1 and V_2 to MGT. The function F_C relates the estimate of MGT to crack growth, and function F_{Prob} estimates the probability of failure based on the estimate of crack growth. Thus, the risk is approximated relying on the failure probability achieved in (3.1). It means that the failure probability represents the risk of failure within a given MGT.

3.2.2 Severity analysis

This section aims to model the visual length of defects based on the Million Gross Tons (MGT). MGT is a measurement unit to show the total weight of freight and passenger trains that pass over a given track in a given time horizon. Thus, the MGT can directly influence the growth of defects in the sense that an increase in the MGT accelerates the defect evolution process and the tracks with a lower train occupation are expected to have a lower degradation rate than the busy tracks.

The defects are automatically detected using the image processing method described in Section 2.4. We measure the visual lengths of the detected defects to use in severity analysis. We consider visual length as an indicator of a defect severity. Analysis of rail image data shows that the visual length of defects can grow with different rates as the MGT increases. To capture the dynamics of the growth, we keep track of the growth for each individual squat to determine the increase of the visual length in each MGT step. A generic function is used to model the growth. The function can be applied relying on different methods where two consecutive data measurements are available. We present the benefits of using an N-step ahead prediction model for the prediction of squat's growth in our recent studies. For details, see [33]; [34]. Thus, considering the index m as an increment counter for a step value of MGT, $M^h(m)$, we use an *N*-step ahead prediction model to describe the growth of visual length at different growth scenarios $h = h_1,\ h_2,\ \ldots,\ h_H$:

$$\begin{cases} \hat{V}_i^h(m+1) = F_S^h\left(\hat{V}_i^h(m), M^h(m)\right) & for\ m = 0,1,...,N-1 \\ \hat{V}_i^h(0) = V_i^h(0) \end{cases} \tag{3.2}$$

where $\hat{V}_i^h(m)$ is the estimate of the visual length for each individual squat *i* at step m assuming scenario *h*, $F_S^h(\cdot)$ is the one-step ahead prediction function, and $V_i^h(0)$ is the visual length measurement at the current step. By partially inversing $F_S^h(\cdot)$, we get $F_{S,inv.}^h$ as a function of the visual length in two consecutive MGT steps. In case of scarce data for the total amount of MGT in each step, an approximation can be made for the prediction model (3.2):

$$\hat{V}_i^h(m+1) = F_{S,approx.}^h\left(\hat{V}_i^h(m)\right) \tag{3.3}$$

In this chapter, a fixed increment of the MGT (*m*) is assumed and the step value for MGT holds the same over a given prediction time horizon. Then, we apply function $F_{S.approx.}^h$ in an *N*-step ahead fashion to reconstruct F_S^h. This allows to formulate the relation between visual length and MGT at step m. Once F_S^h is formulated, we can partially inverse it to get $F_{S,inv.}^h$ as follows:

$$MGT^h(m) = F_{S,inv.}^h\left(\hat{V}_i^h(m+1), \hat{V}_i^h(m)\right) \tag{3.4}$$

3.2.3 Crack growth analysis

The crack growth of defects is an important factor in rail breaks. Independent of the defect severity, the growth of the crack length depends on the traffic load (MGT). The idea in this chapter is to analyze the data measured by ultrasonic detection technique and to present a function for estimation of the crack growth over the MGT [33]; [34]:

$$\Delta\hat{L}_i^h(m) = F_C^h\left(\hat{M}^h(m)\right) \qquad for\ m = 0,1,...,N-1 \tag{3.5}$$

where $\Delta\hat{L}_i^h(m)$ is the estimate of the crack growth length for defect *i* at MGT step m assuming scenario *h* and $F_C^h(.)$ is the crack growth function. We will use a similar approach as described in Section 3.2.2 to assess the crack growth function. Regarding the crack growth data, assume the crack growth length is ΔL, containing I measurements total $(\Delta L_1, \Delta L_2, ..., \Delta L_I)$.Then the failure event can be defined as:

$$\bigcup_{i=1}^{I}\{\Delta L_i > d_i\} \tag{3.6}$$

where d_i is the critical level for the *i*th measurement. This formula implies that a failure occurs if the crack growth length exceeds the critical level. A logistic function is appropriate for these data since the variable is binomial meaning that the system fails if the measurement value satisfies (3.6), otherwise there is no failure [35]. Therefore, a logistic function is considered for the likelihood of rail failure probability $f(\Delta L|(a,b))$ with parameters *a* (intercept) and *b* (slope). Recently, the Bayesian inference model has been employed extensively to assess model uncertainty and robustness for stochastic data behaviors [36]; [37]; [38]. Using a Bayesian inference model, variations of the model parameters can be considered as a step-wise degradation process. According to Bayes theorem, if prior knowledge about the parameters θ is represented by its probability density distribution $\pi_0(\theta)$, and if the statistical observations of crack growth length have likelihood $f(\Delta L|\theta)$, then rail failure probability can be expressed as posterior distribution π:

$$\pi(\theta|\Delta L) = \frac{f(\Delta L|\theta)\pi_0(\theta)}{f(\theta)} \propto f(\Delta L|\theta)\pi_0(\theta), \qquad \theta \in (a,b) \tag{3.7}$$

Typically, Monte Carlo methods are used in Bayesian data analysis to derive the posterior distribution [39]; [40]. The aim of using a Monte Carlo method is to generate random samples from the posterior distribution in order to use them when it is impossible to analytically compute the posterior distribution. Among all the Monte Carlo methods, slice sampling is easier to implement as only the posterior needs to be specified [41]; [42]. The slice sampling algorithm selects samples uniformly from the region under the density function. Therefore, in this chapter, a slice sampling algorithm is selected to capture the failure probability function.

3.2.4 Analysis of rail image data

We consider a railway health monitoring situation where a huge amount of video data is regularly collected. Subsequently, the video data needs to be analyzed in order to detect defects with a potential risk of rail break. The data is collected by a set of high frame rate cameras that are mounted on a measurement train. The video recordings cover the entire length of the measured distance on the rail track. The mounted cameras capture the rails from several angles to look at different components. The top view camera is aimed at the rail surface defects, with each frame covering a length of 15 cm of the track along the longitudinal direction. The recordings are preprocessed into video compilations where consecutive frames have a few millimeters of overlap and the effects of variations in the train speed are removed. Recordings made from (bi)monthly measurements of roughly 6500 kilometers of rail amount to producing thousands of Gigabytes. Every 4 Gigabytes of data covers 16 kilometers of rail track. As a result, for recording videos of the whole Dutch rail network, almost 10 terabytes of data are required per year.

To be able to automatically extract defect information from the data, we train and apply a deep convolutional neural network (DCNN) [43] to detect and classify the defects. Recently, application of DCNN has become very popular in the domain of big data due to the increases in the size of available training sets and algorithmic advances such as the use of piecewise linear units and dropout training [44]; [45]; [46]. By passing through a number of convolutional layers, the images are fed to the DCNN to train a set of shared neuron weights, referred to as filters. Convolution filters detect distinguishing features and form what is called a feature map. We use Rectified Linear Unit (ReLU) [47] activation function which is a piecewise linear function that outputs the input directly if it is positive, otherwise, it outputs zero. The function comes after the convolution steps. We also use max-pooling layers to efficiently down-sample the outcome of each layer. Moreover, to prevent overfitting to the training data, we use dropout layers before each convolutional layer. Overfitting occurs when a classifier is fitted too closely to the sample data set that is unable to accurately describe the entire population, resulting into a high error over the test data. The dropout layer is known to prevent this by randomly disabling some activations from the previous layer [48]. The convolutional and pooling layer are finally attached to a sequence of three fully-connected layers to get class predictions. The DCNN is trained by iterative feed forward of the training examples through the network and by calculating the error with respect to the desired outcome. The error and its gradient are then evaluated at the last layer of the network and back-propagated through all the layers to adjust all the weights. Repeating this process until decreasing the error to a certain limit is called the gradient descent algorithm [47]. We use a widely applied variation of the algorithm where on each iteration, the error and gradients are calculated using a randomly selected set of training examples usually called a mini-batch [47].

3.3 Case study

In this section, a track from the Dutch railway network is considered to illustrate the capabilities of the proposed methodology. Track availability can be affected by rail surface defects. Among all types of rail surface defects, like rail corrugation, head checks, shatter cracking, vertical splits, head horizontal splits and wheel burns, squats play an important role in having a significant impact on the health condition of the track. Therefore, our main focus is on detecting the squats in this case study.

We select a sample from these data that contains recordings over a track in the North of The Netherlands from Zwolle to Groningen corresponding to approximately 300,000 captured frames. Two successive measurements of the same location along the track are matched together using the available time and geographic data. In total 4220 samples are labeled and used for training and testing of the neural network model. Out of the total set of samples, 3170 are normal rail samples and roughly 1000 are squats. The proposed DCNN architecture for analyzing this amount of image frames is presented in Figure 3.2. Initially the input images are down-scaled to 375×275 pixels and converted into gray-scale. The sequence of three fully-connected layers translates the extracted high-level features from the previous layers, into 3 classes representing the normal rail, trivial defects (seed squats), and squats.

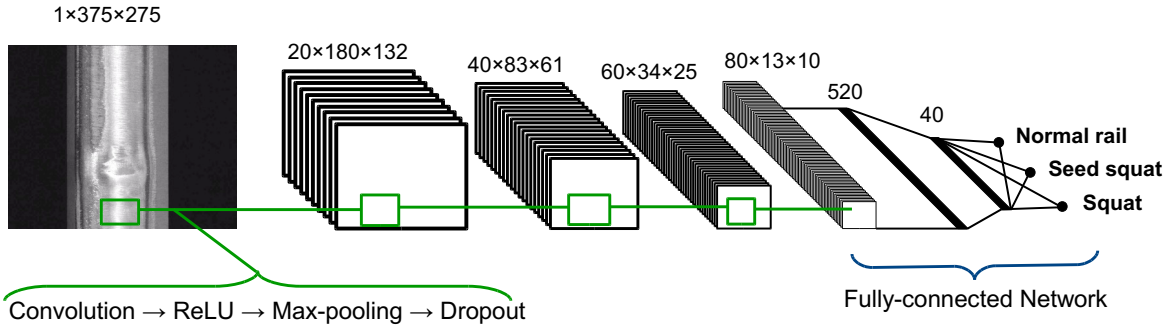


Figure 3.2 Architecture of the proposed DCNN model

Trivial defects appear in the form of spots or small damages to the rail head, while squats are usually defects that are fully grown indentations and deformations of the rail surface. The normal class includes all other components such as plain rails, switches, welds, possible non-defect contaminations, etc. To train the network, a set of manually labeled examples is collected from several locations along the measured track and is compiled into a training set for each one of the 3 classes. The network is trained once and then is used for multiple time predictions. The training time is 40 hours per 1500 examples. Once the network is trained, it is used to find squats in the large pool of previously unseen samples (prediction). These samples are collected from other monitoring sessions. Unlike the training time, the prediction time is insignificant (30 seconds per 15000 examples). The prediction result then has an average binary accuracy of 96.9 % (squat vs. normal) when training on 80% of the labelled dataset and testing on the remaining 20%. By putting a high acceptance threshold on the network output response, we opt to detect the correct cases of squats, trivial defects and the normal cases. Hence, after training and testing, we use the model to predict the severity of squats from the large amounts of available unlabelled data, from which we choose 109 detected squats for manual measurement of visual lengths in the track Zwolle-Groningen. Then, the samples are used in the next step where the growth of visual lengths is considered as described in Section 3.2.2.

Here, squats with a visual length below 15 mm are considered as light squats, in which cracks have not appeared yet (surface initiation is assumed, and we cannot see beneath the surface from the image). Squats with visual length ranging from 15 to 30 mm are considered to be at the medium stage of growth. The medium squats evolve to severe squats when the network of cracks spreads further. Figure 3.3 shows reference photos of squats ranging from light to severe together with crack evolution. After repeated train passes, light squats will evolve into medium or severe squats. Once the squat is severe, the squat will evolve into a defect with surface-initiated cracks growing along the depth beneath the rail surface [49].

After detecting squats by image processing, we apply the approach as described in Section 3.2.2 for this particular case to construct severity function. From real data of visual length, we estimate $F^{h}_{S,approx.}$ from (3.3).

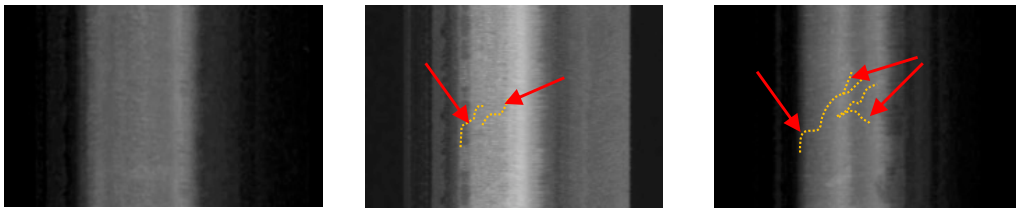


Figure 3.3 A sample of squats in different classes of severity, the red arrows show the evolution of the crack when it gets severe.

Figure 3.4. shows the relation between two consecutive measurements of visual length for a fixed value of MGT-step (m = 1). Relying on the physical understanding of how a squat grows, we fit a polynomial regression model of degree 3, using the least-absolute residual method [50], to represent the stochasticity of the growth. The residual plot together with the R-square value of 0.9778 determines how well the polynomial model fits the data. We consider the fit model as an average growth scenario, and the 3-sigma control limits as slow and fast growth scenarios. We use the estimated function of Figure 3.4 for 8-step ahead prediction, and consider a fixed MGT increment of 3.01 in each step. As a result, a model-based prediction function for the visual lengths versus MGT is depicted in Figure 3.5, considering the three scenarios of average (a), fast (b), and slow (c).

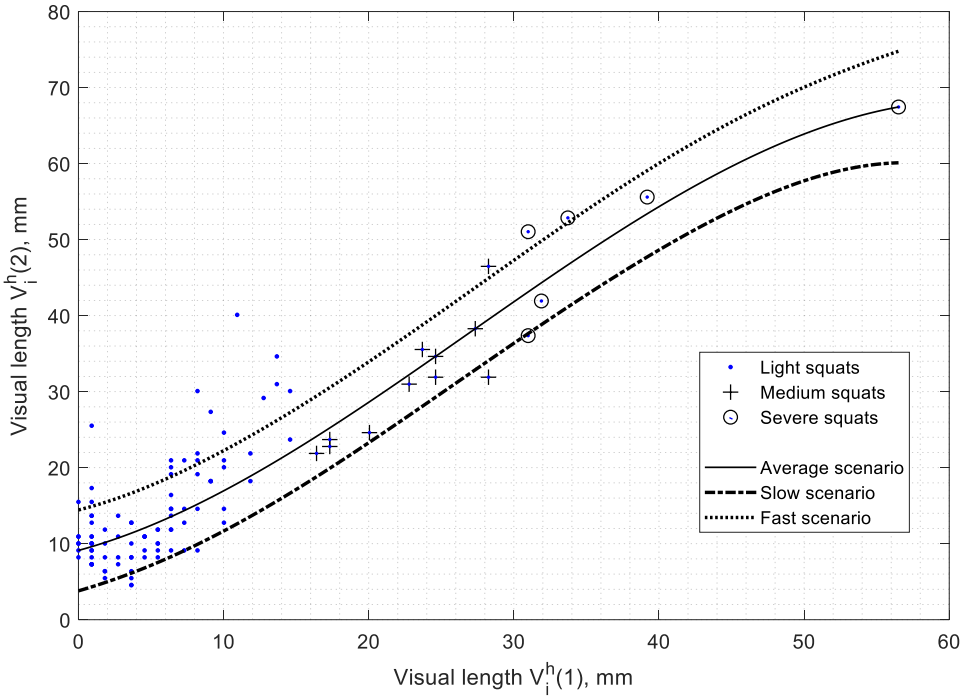


Figure 3.4 Estimation of the visual length of the squats for m = 1, and based on real data

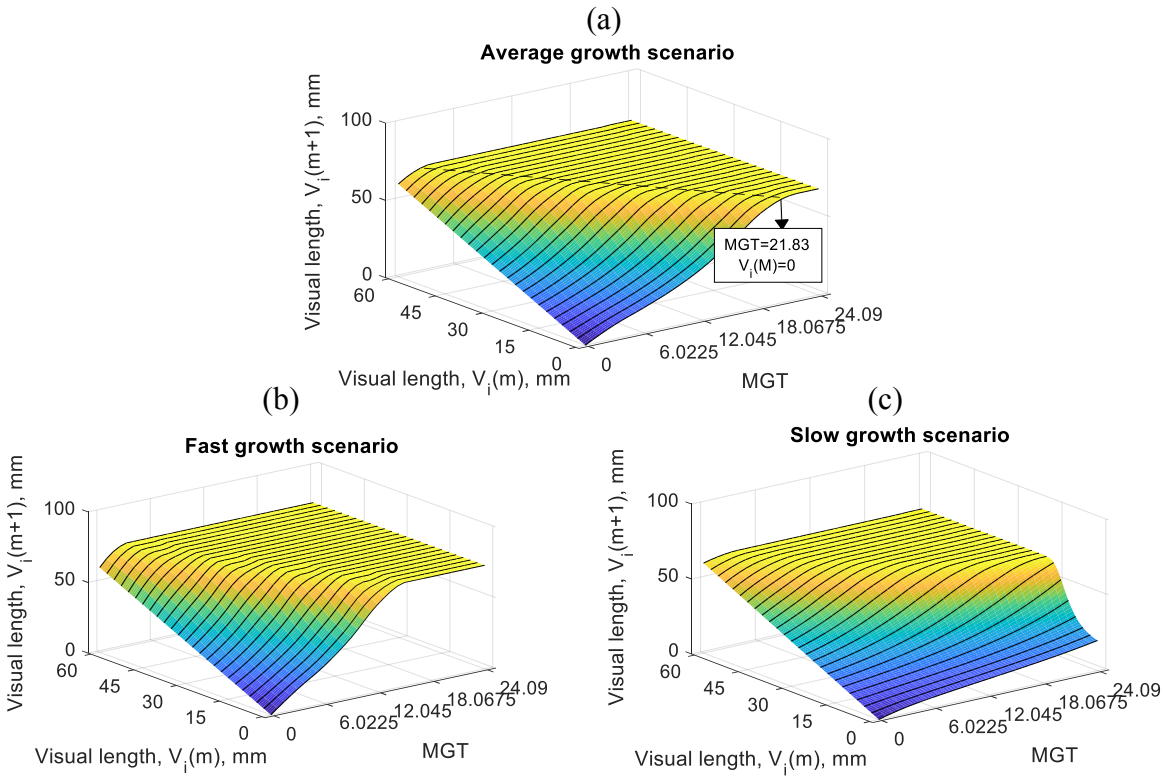


Figure 3.5 Growth of squat visual length over MGT for the following model-based growth scenarios: (a) average, (b) fast, (c) slow, the dotted line depicts an upper bound of squat visual length.

The dotted line shows the upper bound of the estimation for visual length, i.e. it is very rare to observe a squat with a length over the upper bound in reality. Assuming $V_i^h(m) = 0$, the visual length at MGT step $m+1$ at the fast scenario, reaches the upper bound with a MGT $(\mathrm{MGT}_{h_1} = 15.06)$, lower than at the average scenario $(\mathrm{MGT}_{h_3} = 21.83)$ and at the slow scenario. It means that the degradation process in the fast scenario is more accelerated than in the average and slow scenarios as the traffic load on rail increases. We estimate the crack growth function, $F_C^h(\cdot)$ by relying on ultrasonic measurement data. The model-based relation between the crack growth length and MGT is shown in Figure 3.6. In addition, three different scenarios are considered to capture the crack growth dynamics, including the average scenario, the slow scenario, and the fast scenario. As seen in the figure, at the fast scenario, crack propagation of the squat at a given MGT is significantly faster than squats that are at the average and slow scenario. For example, at MGT = 10.36, it is estimated that the crack length of a squat grows 1 mm at the slow scenario, 2 mm at the average scenario, and 8 mm at the fast scenario. We can assess the risk of rail failure considering any of the different scenarios of crack growth length.

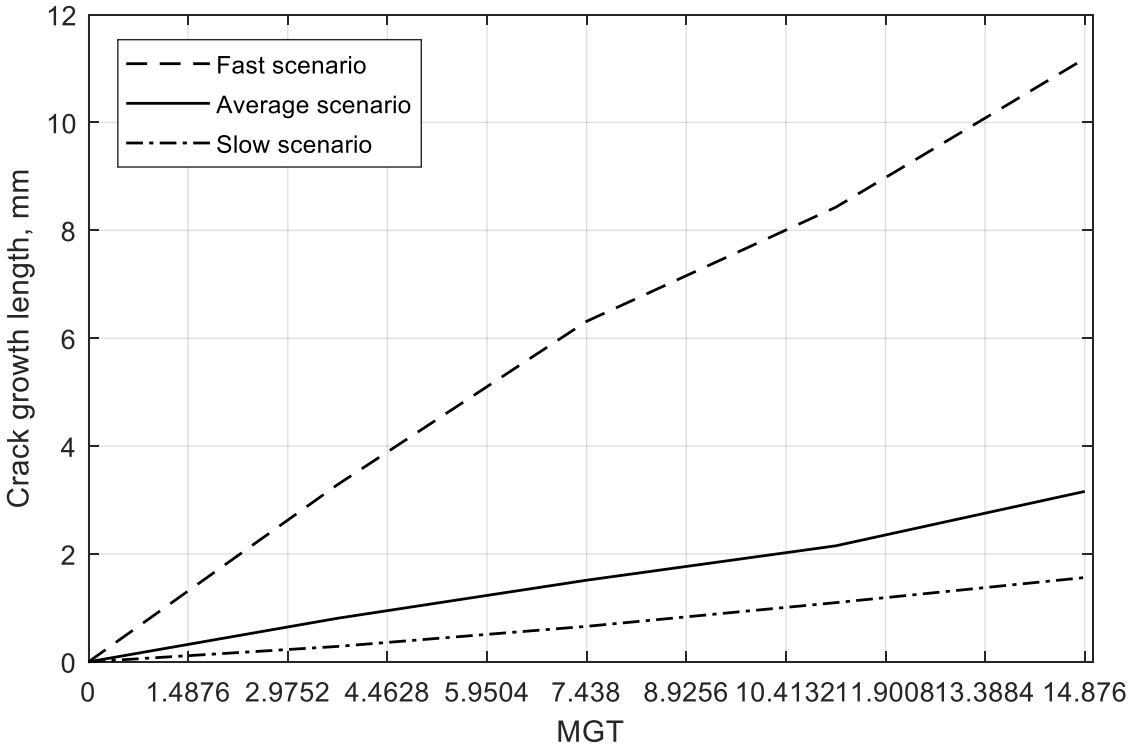


Figure 3.6 A model-based relation between crack growth length and MGT

In the failure probability model, we consider that a rail is prone to fail when a squat reaches a crack length of 9 mm as the threshold for rail safety criticality. The crack length of each squat is measured to see how it has grown over MGT, and how many cracks have reached a length of 9 mm or even more. We use normal priors for the regression parameters (a,b). Relying on the data for the crack growth length, the parameters are estimated by a slice sampling algorithm considering 1000 samples. Respectively, Figure 3.7 and Figure 3.8 shows how the mean of the parameter *a* and *b* varies over the samples and converges to a constant value. As seen in the figures, the posterior means of parameters converge to a stationary status after the first 50 samples.

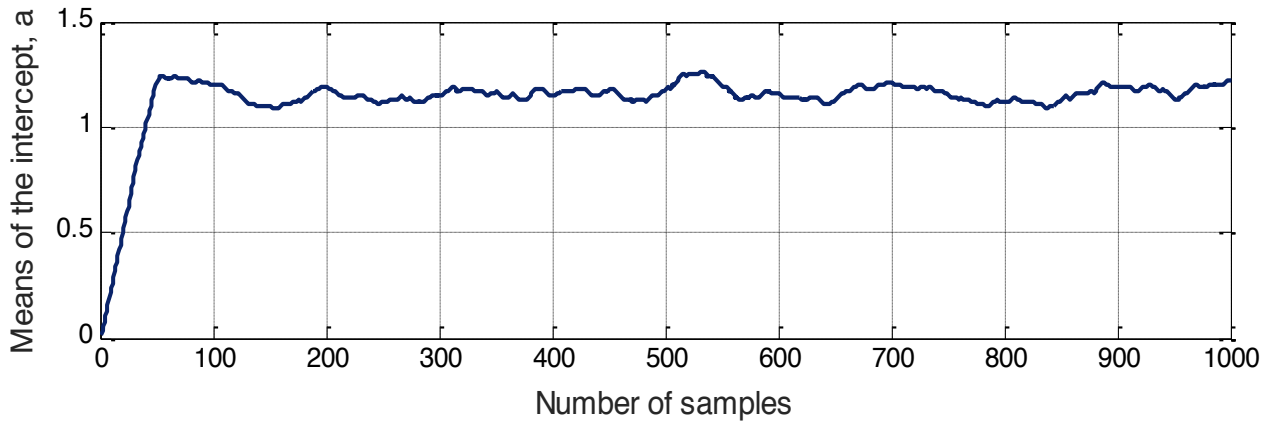


Figure 3.7 Posterior distributions of regression parameter *a*

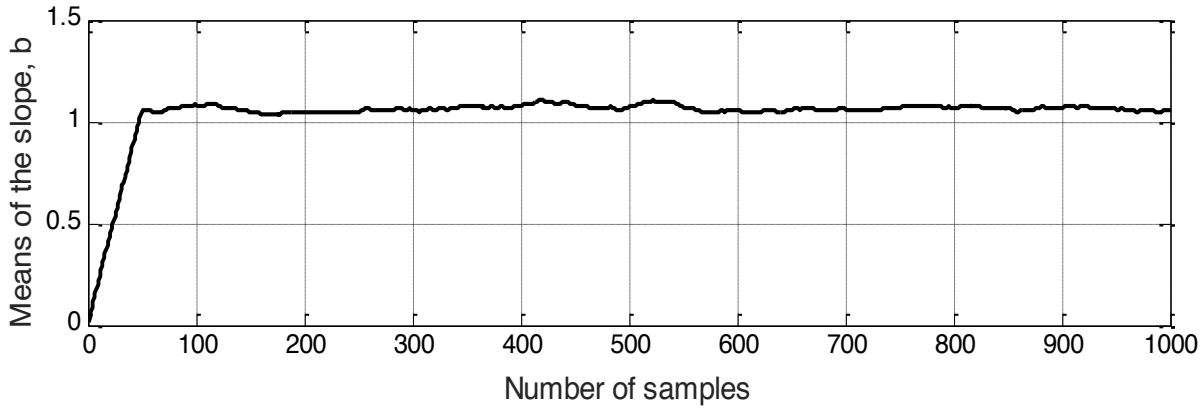


Figure 3.8 Posterior distributions of regression parameter *b*

3.4 Results and discussion

For a detected squat with measured visual lengths in one MGT step, we estimate the risk of rail failure as follows: From the model in Figure 3.5, we estimate the MGT for the visual lengths in two consecutive measurements. Then, from the model in Figure 3.6, we find the crack growth length for the estimated MGT. Finally, we estimate the failure probability from the crack growth length in Figure 3.9. The failure probability plot represents how probable a squat fails in the next MGT step when the crack growth length is given. As an example, if the crack length of a squat increases 6 mm for MGT =7.04, the probability that the squat could lead to a rail break is roughly 0.82. In Figure 3.10, a sample of 5 squats is visualized, and the estimates of failure probability from the given visual lengths are presented.

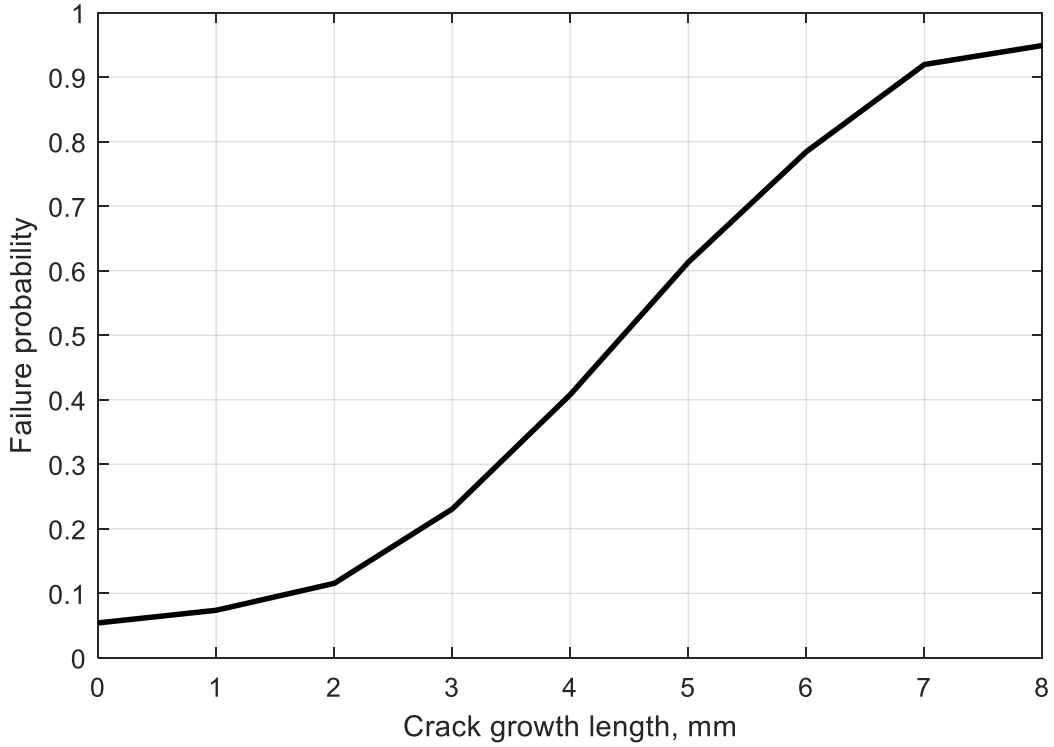


Figure 3.9 Probability of rail failure based on the growth of crack length

For instance, the squat with $V_1 = 42$ mm and $V_2 = 57$ mm will cause a rail break with a probability of 28.9% in the next MGT step, if no maintenance action is operated. However, no serious failure threatens the squat at the early stage and the failure probability is then almost 10% (see the squat with 16 mm in visual length). In Table 3.1, more samples of squats are presented. The table includes 64 samples of squats with their measurements of visual length for two MGT steps. As expected, the squat at the severe stage will be prone to a rail break if no operation is carried out on the rail within a given MGT step. For example, there is a 53% chance of failure for the 64th squat in which the crack growth length is 4.10 mm within the given MGT step. The estimated risk values for the squats at the late stage indicate the need for immediate rail replacements. For the squats at early stage, a grinding operation is suggested to postpone rail failure by treating the squats.

Table 3.1. Failure risk estimation for a sample of squats, detected on the track Zwolle-Groningen

Squat	V_1	V_2	ΔL	F_{Prob}	Squat	V_1	V_2	ΔL	F_{Prob}
1	3.65	4.56	0.02	0.055	33	16.41	21.87	0.48	0.079
2	8.20	9.11	0.03	0.056	34	10.03	14.58	0.42	0.076
3	3.65	5.47	0.05	0.057	35	6.38	19.14	0.46	0.078
4	7.29	9.11	0.05	0.057	36	8.20	21.87	0.48	0.079
5	3.65	6.38	0.08	0.058	37	17.32	23.70	0.55	0.083
6	5.47	8.20	0.09	0.059	38	7.29	20.96	0.49	0.080
7	6.38	9.11	0.09	0.059	39	6.38	20.96	0.53	0.082
8	4.56	8.20	0.10	0.060	40	9.11	27.34	0.63	0.087
9	5.47	9.11	0.11	0.060	41	11.85	18.23	0.60	0.085
10	2.73	7.29	0.13	0.061	42	8.20	30.08	0.78	0.095
11	3.65	8.20	0.13	0.061	43	14.58	23.70	0.77	0.094
12	4.56	9.11	0.14	0.061	44	28.25	31.90	0.95	0.104
13	2.73	8.20	0.15	0.062	45	11.85	21.87	0.90	0.101
14	5.47	10.03	0.15	0.062	46	10.03	20.96	0.94	0.103
15	6.38	11.85	0.17	0.063	47	14.58	30.08	1.17	0.122
16	7.29	12.76	0.19	0.064	48	30.99	37.37	1.55	0.156
17	3.65	10.03	0.19	0.064	49	13.67	30.99	1.31	0.134
18	4.56	10.94	0.19	0.064	50	12.76	29.16	1.29	0.133
19	5.47	11.85	0.21	0.065	51	10.03	24.61	1.20	0.125
20	8.20	14.58	0.21	0.065	52	20.05	24.61	1.48	0.151
21	10.03	12.76	0.25	0.067	53	13.67	34.63	1.48	0.151
22	2.73	10.94	0.24	0.067	54	24.61	31.90	1.91	0.190
23	6.38	13.67	0.24	0.067	55	31.90	41.92	2.23	0.231
24	7.29	14.58	0.24	0.067	56	10.94	40.10	1.95	0.194
25	6.38	14.58	0.27	0.068	57	22.78	30.99	2.35	0.248
26	3.65	12.76	0.27	0.068	58	24.61	34.63	2.56	0.277
27	9.11	18.23	0.29	0.069	59	27.34	38.28	2.62	0.286
28	2.73	13.67	0.34	0.072	60	39.19	55.59	3.05	0.348
29	6.38	16.41	0.34	0.072	61	23.70	35.54	3.09	0.355
30	8.20	19.14	0.38	0.074	62	33.72	52.86	3.69	0.461
31	17.32	22.78	0.46	0.078	63	28.25	46.48	3.87	0.493
32	8.20	20.96	0.44	0.077	64	30.99	51.04	4.10	0.532

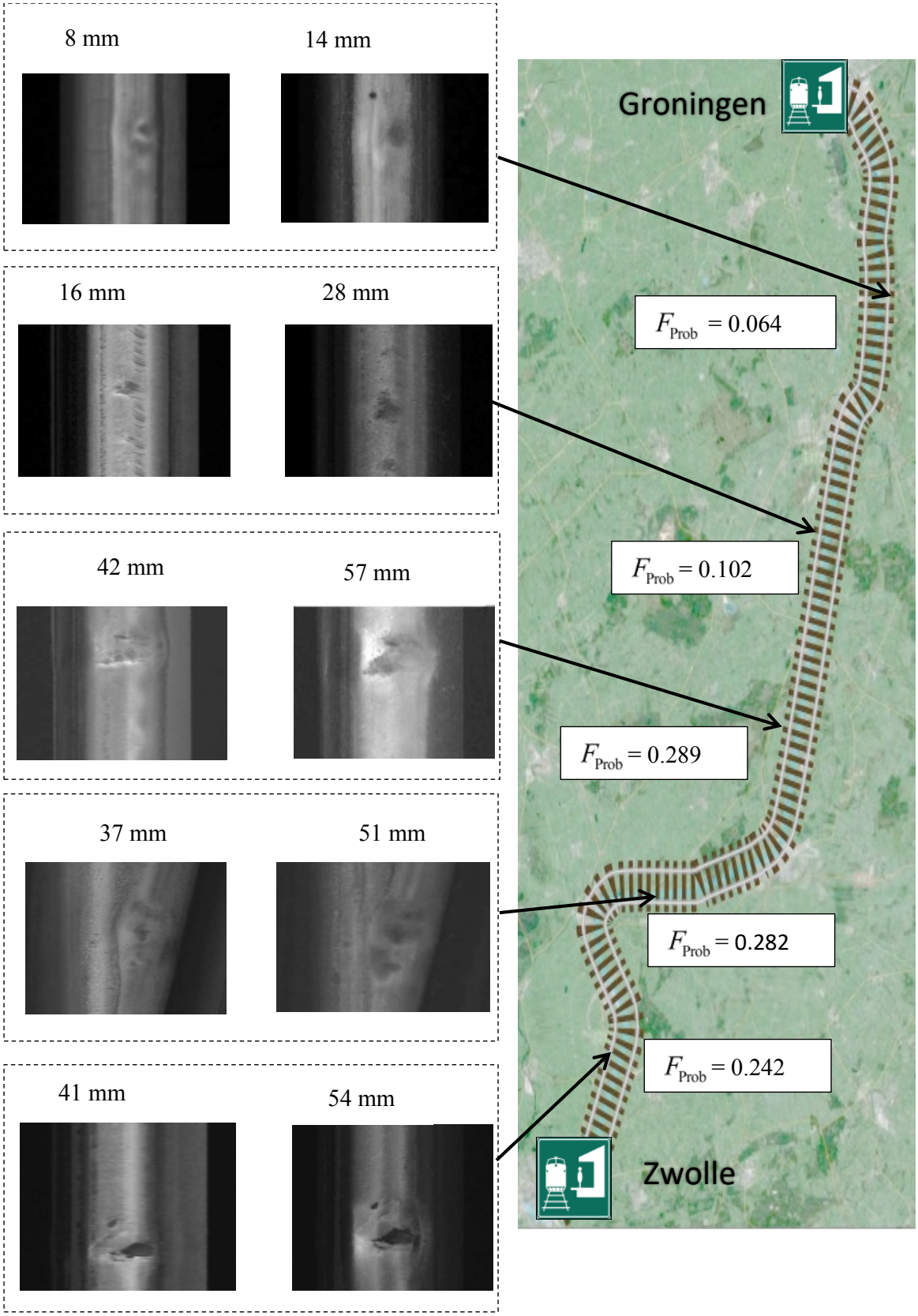


Figure 3.10 A sample of failure risk estimates for 5 squats over the track

3.5 Conclusions

In this chapter, we present a methodology for the risk assessment of rail failure for a type of rail surface defects called squats. A big data analysis approach is used to automatically detect squats from rail images. The visual lengths of squats are measured in order to use them in the severity analysis model, which captures the growth of visual length over MGT increments. In addition, due to the influence of crack growth on estimation of the failure risk, a crack growth analysis based on MGT has been performed. At the end, a Bayesian model is employed to estimate the failure probability. By relying on the estimated failure risk, the infrastructure manager is able to take actions at the right time and the right place in order to prevent unexpected consequences induced by rail breaks. While this chapter is focused on the analysis of squats, the results can also be applicable for the analysis of other types of rail defects such as head checks and corrugations.

3.6 References

[1] Greenberg, M., Lioy, P., Ozbas, B., Mantell, N., Isukapalli, S., Lahr, M., Altiok, T., Bober, J., Lacy, C., Lowrie, K., & Mayer, H. (2013). Passenger rail security, planning, and resilience: Application of network, plume, and economic simulation models as decision support tools. *Risk Analysis*, 33(11), 1969-1986.

[2] Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137-144.

[3] McGhee, M. J., Galloway, G., Catterson, V., Brown, B., & Harrison, E. (2014). Prognostic modelling of valve degradation within power stations. *Annual Conference of the Prognostics and Health Management Society 2014 (PHM)*. 29 September- 02 October, Texas, USA.

[4] She, S., Ma, C., & Wu, D. D. (2010). General Probability-Time Tradeoff and Intertemporal Risk-Value Model. *Risk Analysis*, *30*(3), 421-431.

[5] Wu, D., & Birge, J. R. (2016). Risk intelligence in big data era: A review and introduction to special issue. *IEEE Transactions on Cybernetics*, *46*(8), 1718-1720.

[6] Choi, T. M., Chan, H. K., & Yue, X. (2017). Recent development in big data analytics for business operations and risk management. *IEEE Transactions on Cybernetics*, *47*(1), 81-92.

[7] Hofstetter, P., & Hammitt, J. K. (2002). Selecting Human Health Metrics for Environmental Decision-Support Tools. *Risk Analysis*, *22*(5), 965-983.

[8] Thekdi, S. A., & Lambert, J. H. (2012). Decision analysis and risk models for land development affecting infrastructure systems. *Risk Analysis*, *32*(7), 1253-1269.

[9] Zhou, Q., Lambert, J. H., Karvetski, C. W., Keisler, J. M., & Linkov, I. (2012). Flood protection diversification to reduce probabilities of extreme losses. *Risk Analysis*, *32*(11), 1873-1887.

[10] Faghih-Roohi, S., Ong, Y. S., Asian, S., & Zhang, A. N. (2016). Dynamic conditional value-at-risk model for routing and scheduling of hazardous material transportation networks. *Annals of Operations Research*, *247*(2), 715-734.

[11] Siddiqui, A., & Verma, M. (2013). An expected consequence approach to route choice in the maritime transportation of crude oil. *Risk Analysis*, *33*(11), 2041-2055.

[12] Paltrinieri, N., Bonvicini, S., Spadoni, G., & Cozzani, V. (2012). Cost-Benefit Analysis of Passive Fire Protections in Road LPG Transportation. *Risk Analysis*, *32*(2), 200-219.

[13] Wu, D. D., Chen, S. H., & Olson, D. L. (2014). Business intelligence in risk management: Some recent progresses. *Information Sciences*, *256*, 1-7.

[14] Wu, D. D., Kefan, X., Gang, C., & Ping, G. (2010). A risk analysis model in concurrent engineering product development. *Risk Analysis*, *30*(9), 1440-1453.

[15] Fumeo, E., Oneto, L., & Anguita, D. (2015). Condition based maintenance in railway transportation systems based on big data streaming analysis. *Procedia Computer Science*, 53, 437-446.

[16] Zoeteman, A., Dollevoet, R., & Li, Z. (2014). Dutch research results on wheel/rail interface management: 2001–2013 and beyond. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *228*(6), 642-651.

[17] Liu, X., Barkan, C., & Saat, M. (2011). Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. *Transportation Research Record: Journal of the Transportation Research Board*, (2261), 178-185.

[18] Liu, X., Saat, M., & Barkan, C. (2012). Analysis of causes of major train derailment and their effect on accident rates. *Transportation Research Record: Journal of the Transportation Research Board*, (2289), 154-163.

[19] Sandström, J., & Ekberg, A. (2009). Predicting crack growth and risks of rail breaks due to wheel flat impacts in heavy haul operations. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *223*(2), 153-161.

[20] Molodova, M., Li, Z., Núñez, A., & Dollevoet, R. (2014). Automatic detection of squats in railway infrastructure. *IEEE Transactions on Intelligent Transportation Systems*, *15*(5), 1980-1990.

[21] Fan, Y., Dixon, S., Edwards, R. S., & Jian, X. (2007). Ultrasonic surface wave propagation and interaction with surface defects on rail track head. *Ndt & E International*, *40*(6), 471-477.

[22] Song, Z., Yamada, T., Shitara, H., & Takemura, Y. (2011). Detection of damage and crack in railhead by using eddy current testing. *Journal of Electromagnetic Analysis and Applications*, *3*(12), 546.

[23] Mariani, S., Nguyen, T., Phillips, R. R., Kijanka, P., Lanza di Scalea, F., Staszewski, W. J., Fateh, M., & Carr, G. (2013). Noncontact ultrasonic guided wave inspection of rails. *Structural Health Monitoring*, *12*(5-6), 539-548.

[24] Li, Y., Trinh, H., Haas, N., Otto, C., & Pankanti, S. (2014). Rail component detection, optimization, and assessment for automatic rail track inspection. *IEEE Transactions on Intelligent Transportation Systems*, *15*(2), 760-770.

[25] Li, Q., & Ren, S. (2012). A visual detection system for rail surface defects. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *42*(6), 1531-1542.

[26] Li, Q., & Ren, S. (2012). A real-time visual inspection system for discrete surface defects of rail heads. *IEEE Transactions on Instrumentation and Measurement*, *61*(8), 2189-2199.

[27] Skakun, S., Kussul, N., Shelestov, A., & Kussul, O. (2014). Flood hazard and flood risk assessment using a time series of satellite images: A case study in Namibia. *Risk Analysis*, *34*(8), 1521-1537.

[28] Sharkey, J., Scarfe, L., Santeramo, I., Garcia-Finana, M., Park, B. K., Poptani, H., Wilm, B., Taylor, A., Murray, P. (2016). Imaging technologies for monitoring the safety, efficacy and mechanisms of action of cell-based regenerative medicine therapies in models of kidney disease. *European Journal of Pharmacology*, *790*, 74-82.

[29] Singh, V. (2016). Higher pollution episode detection using image classification techniques. *Environmental Modeling & Assessment*, *21*(5), 591-601.

[30] MoradiAmin, M., Memari, A., Samadzadehaghdam, N., Kermani, S., & Talebi, A. (2016). Computer aided detection and classification of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis. *Microscopy Research and Technique*, *79*(10), 908-916.

[31] Javaid, M., Javid, M., Rehman, M. Z. U., & Shah, S. I. A. (2016). A novel approach to CAD system for the detection of lung nodules in CT images. *Computer Methods and Programs in Biomedicine*, *135*, 125-139.

[32] Jamshidi, A., Núñez, A., Dollevoet, R., & Li, Z. (2017). Robust and predictive fuzzy key performance indicators for condition-based treatment of squats in railway infrastructures. *Journal of Infrastructure Systems*, *23*(3), 04017006.

[33] Jamshidi, A., Roohi, S. F., Núñez, A., Babuska, R., De Schutter, B., Dollevoet, R., & Li, Z. (2016). Probabilistic defect-based risk assessment approach for rail failures in railway infrastructure. *IFAC-PapersOnLine*, *49*(3), 73-77.

[34] Xu, D., & Zhao, W. (2005). Reliability prediction using multivariate degradation data. *Proceedings of the IEEE in Reliability and Maintainability Symposium*, Alexandria, 24-27 January, VA, USA.

[35] Li, L., Wang, J., Leung, H. & Jiang, C. (2010). Assessment of catastrophic risk using Bayesian network constructed from domain knowledge and spatial data. *Risk Analysis*, 30(7),1157-1175.

[36] Droguett, E. L., & Mosleh, A. (2008). Bayesian methodology for model uncertainty using model performance data. *Risk Analysis*, *28*(5), 1457-1476.

[37] Kazemi, R., & Mosleh, A. (2012). Improving default risk prediction using Bayesian model uncertainty techniques. *Risk Analysis*, *32*(11), 1888-1900.

[38] Faghih-Roohi, S., Xie, M., & Ng, K. M. (2014). Accident risk assessment in marine transportation via Markov modelling and Markov chain Monte Carlo simulation. *Ocean Engineering*, *91*, 363-370.

[39] Marques, R., Bouville, C., Ribardière, M., Santos, L. P., & Bouatouch, K. (2013). A spherical Gaussian framework for Bayesian Monte Carlo rendering of glossy surfaces. *IEEE Transactions on Visualization and Computer Graphics*, *19*(10), 1619-1632.

[40] Gilks, W. R. (2005). Markov chain monte carlo. *John Wiley & Sons, Ltd*.

[41] Neal, R. M. (2003). Slice sampling. *Annals of Statistics*, 705-741.

[42] Faghih-Roohi, S., Hajizadeh, S., Núñez A, Babuska, R., De Schutter, B. (2016). Deep convolutional neural networks for detection of rail surface defects. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)*. 24-29 July, Vancouver, Canada.

[43] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Image net classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097-1105.

[44] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

[45] Sainath, T. N., Kingsbury, B., Saon, G., Soltau, H., Mohamed, A. R., Dahl, G., Ramabhadran, B. (2015). Deep convolutional neural networks for large-scale speech tasks. *Neural Networks*, 64, 39-48.

[46] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.

[47] Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, 2(1), 1-127.

[48] Li, Z., Dollevoet, R., Molodova, M., & Zhao, X. (2011). Squat growth–some observations and the validation of numerical predictions. *Wear*, 271(1-2),148–157.

[49] Bassett, J. R. G., & Koenker, R. (1978). Asymptotic theory of least absolute error regression. *Journal of the American Statistical Association*, 73(363), 618-622.

A rail maintenance decision support approach using big data analysis

This chapter corresponds to the reference: *A. Jamshidi, S. Hajizadeh, Z. Su, M. Naeimi, A. Núñez, R. Dollevoet, B. De Schutter and Z. Li, "A decision support approach for condition-based maintenance of rails based on big data analysis". Transportation Research Part C: Emerging Technologies, Volume 95, October 2018, Pages: 185-206. DOI: 10.1016/j.trc.2018.07.007*

4.1 Introduction

The increase of train traffic and axle loads affect the health condition of railway infrastructure. Hence, efficient infrastructure monitoring and maintenance is among the major concerns of infrastructure managers in order to improve the performance of railway operations [1]. As such, infrastructure health condition should be monitored and considered in the decision making of maintenance. Effective management of infrastructure health condition is crucial to guarantee the desired asset quality level [2]; [3]. It also plays an important role in meeting the demand when the infrastructure is upgraded e.g. when increasing traffic capacity, the maintenance regime should be adapted to avoid compromising safety and infrastructure health requirements. To keep the infrastructure system working at an effective level, a condition-based maintenance system is required not only to consider the actual heath condition but also its evolution during the maintenance decision horizon [4].

Condition-based monitoring is used in railway infrastructures to estimate the actual health condition of the assets, so that degradation processes can be effectively controlled. It helps to keep the infrastructure manager continually informed of the estimated health of the railway infrastructure. Condition-based monitoring is supposed to collect information that will allow an effective operation by reducing maintenance cost, eliminating unnecessary operations and focusing on places where the problems are located and where they will be in the coming period. Furthermore, the enhancement in usage of the railway infrastructure needs a systematic monitoring plan to keep the trains running safely by considering all related data influencing the health condition. The data for a typical railway infrastructure includes a large amount of frequent measurements from the monitoring systems of the assets involved in the railway operations. To ensure the performance level, a huge amount of data should be collected, transmitted, processed and properly stored so that it can be used as historical information. This whole process reflects a transition from raw infrastructure data into actionable maintenance knowledge. Therefore, the database constituted from continuous monitoring will become larger and larger over time and applying big data analysis approaches is inevitable [5]. The classic methods fail to scale up to the huge volumes of the rail data. Thus, in order to design proper maintenance plans in railways, it is necessary to explore and analyse the growing amount of data and to extract what is useful information of it. To do so, different sensors can be used to collect the data obtained in railway track monitoring at different times, environmental conditions and frequencies. These data can include different characteristics: (1) discrete or continuous, (2) spatial or temporal, (3) signal and images among others [6]; [7]; [8]; [9]; [10]. In condition-based maintenance for railways, the monitoring data are mostly collected periodically with regular sampling intervals. For some critical assets, the monitoring can be adapted to other possible needs including a continuous measurement. The basic concept is to follow correctly the degradation of the infrastructure, in particular for critical infrastructure like rails. This chapter focuses on rail condition monitoring which have a critical role in the network performance [11]; [8]. A notable amount of the maintenance budget should be assigned therefore for the rail in an intensive railway network like the Dutch railway network [12]. Due to the fact that rail infrastructure contributes hugely to the failures occurred in the network, it is important to assess the rail health condition. The analysis of the health condition analysis involves the detection of severe rail defects which play key role in causing rail break [13]; [14]; [15].

Rolling contact fatigue (RCF) affects the health condition of the rail due to the contact in the interface between wheel and rail [16]. RCF is a generic term describing a range of rail surface defects and has been an interesting challenging research topic, in particular the influence of RCF on maintenance decision making [17]. Moreover, its influence is related to other factors including traffic type, train speed, traffic load, rail/wheel profile, train characteristics and maintenance policy [18]. Once RCF appears, it induces considerable dynamical forces on the rail surface, and subsequently cracks are initiated and propagated from the surface [19]; [20]. The most important cause of defect appearance is the large number of trains passing over rail critical components, most significantly at welds, joints, and switches [21]. Early detection of surface defects is important to mitigate induced maintenance costs as well as unforeseeable consequences of rail breaks. There are different methods to diagnose the condition of rail defects, including ultrasonic measurements [22], eddy current testing [23] and guided-wave based monitoring [24].

In this chapter, the focus is on a type of rail surface defect called squat. The costs for treating these defects in the Dutch railway network are considerably high (more than 5000 euro/km per year) [21]. The maintenance of squats should be different according to their severity. For late-stage squats, a rail replacement plan is a proper decision while for the light squats, grinding a thin layer from the rail surface is the most effective solution. Hence, when all residual damages are removed, grinding is effective and the rail will be turned to a healthy condition. To optimally plan grinding

operations, condition-based maintenance relying on early detection of the squats is required. Although a defect detection method could give an indication of the health of the rail, the infrastructure manager requires prior knowledge to (1) be aware of all influential factors, (2) analyse interdependency between the rail observations and the influential factors and (3) obtain a future view of the track condition. Hence, by having knowledge about the track characteristics, potential risks about the rail can be anticipated due to the effect of the influential factors on the defect appearance and consequently on the rail health condition. Therefore, an analysis of influential factors should be taken into account to give at the most a proper prospect of the infrastructure health condition.

Mixed Integer Linear Programming (MILP) is a common approach for track maintenance scheduling. A MILP model is developed in [25] for optimal condition-based preventive maintenance for a single track divided into multiple segments, considering various economic and technical factors such as train speed limit and track quality. The optimal planning of routine maintenance activities and projects like grinding to minimize maintenance costs and track possession time for a single track is formulated as an MILP problem in [26]. The optimal scheduling of rail, sleeper and ballast renewal at a network level is formulated as an MILP problem in [27] to minimize the expected life-cycle cost and track unavailability. In [28], the optimal clustering of track maintenance jobs into projects to minimize total maintenance costs for a network of track is recast as a Vehicle Routing Problem (VRP). The track maintenance problem considering different priorities for each section in the railway network is formulated as a VRP with customer costs in [29]. A time-space network model is developed in [30] for the optimal scheduling of capital maintenance projects like rail replacement. A metaheuristic based on simulated annealing is developed in [31] to determine the optimal tamping length of a tamping machine, minimizing the associated logistic costs and fixed machine costs. In this chapter, we use a simplified MILP model to optimize the rail grinding decision plan into clusters that can be related to the actual condition of the rail. The proposed MILP model eases implementing the condition-based maintenance strategy, reaching an effective maintenance plan in terms of rail health condition and also reduces the high cost of track maintenance activities. In this chapter, we propose a condition-based maintenance methodology taking both the observations and the prior knowledge of the track into account. The idea is to find interdependency between defect status and all major influential factors of the track prior knowledge. The defect status is defined in terms of number and severity of the defects. We investigate the interdependencies between the influential factors and the defect appearance by studying the track characteristics. Once the interdependency is studied, a set of rules is generated to connect rail conditions to the influential factors. The results then indicate which pieces of the rail are prone to be defective. The infrastructure manager is then able to propose maintenance planning according to the critical pieces of railway track. The methodology uses big data analytics, with real-life data measured from a Dutch railway track, using Axle Box Acceleration (ABA) measurements and rail video images [21]; [32]; [33].

Figure 4.1 shows the flowchart of the proposed methodology divided into five steps. The major contribution of the chapter is in proposing a methodology that combines different methods for rail defect detection and also maintenance intervention optimization. For instance, in Step 5 we have adapted the model proposed in [34] to fit the new formulation proposed in this chapter in Step 1. The integrated methodology is aimed to be implemented by the infra-managers; thus, simplicity and coherence between steps are very important to guarantee real-life implementation. In Step 1, the most important defects are detected using the ABA signals and rail video images considering the track position and the severity. A list of defects is assumed to be provided by the measurement systems. In Step 2, major influential factors, γ_j, are presented to give context on the prior knowledge of the track in the segment j. Step 3 presents the interdependency analysis between the

influential factors and the observations obtained from Step 1. The aim is to investigate on how the influential factors are related to the observations. In Step 4, a decision system is proposed using a fuzzy inference system. A set of possible condition rules, $r = r_1, r_2, ..., r_R$, is generated to build up the inference system estimating the rail health condition relying on Step 3. Finally, in Step 5, maintenance planning is proposed based on the infrastructure health condition. For this purpose, the track can be divided to N_s rail segments (or clusters), $j = j_1, j_2, ..., j_{N_s}$ to facilitate the maintenance decisions for infrastructure manager, particularly in long tracks. The infrastructure manager then gets informed of each segment's status within a maintenance time horizon. Once one rail segment requires maintenance operation for grinding, the corresponding maintenance intervention is suggested. Therefore, the methodology covers two major parts of rail maintenance framework which are condition monitoring and maintenance plan.

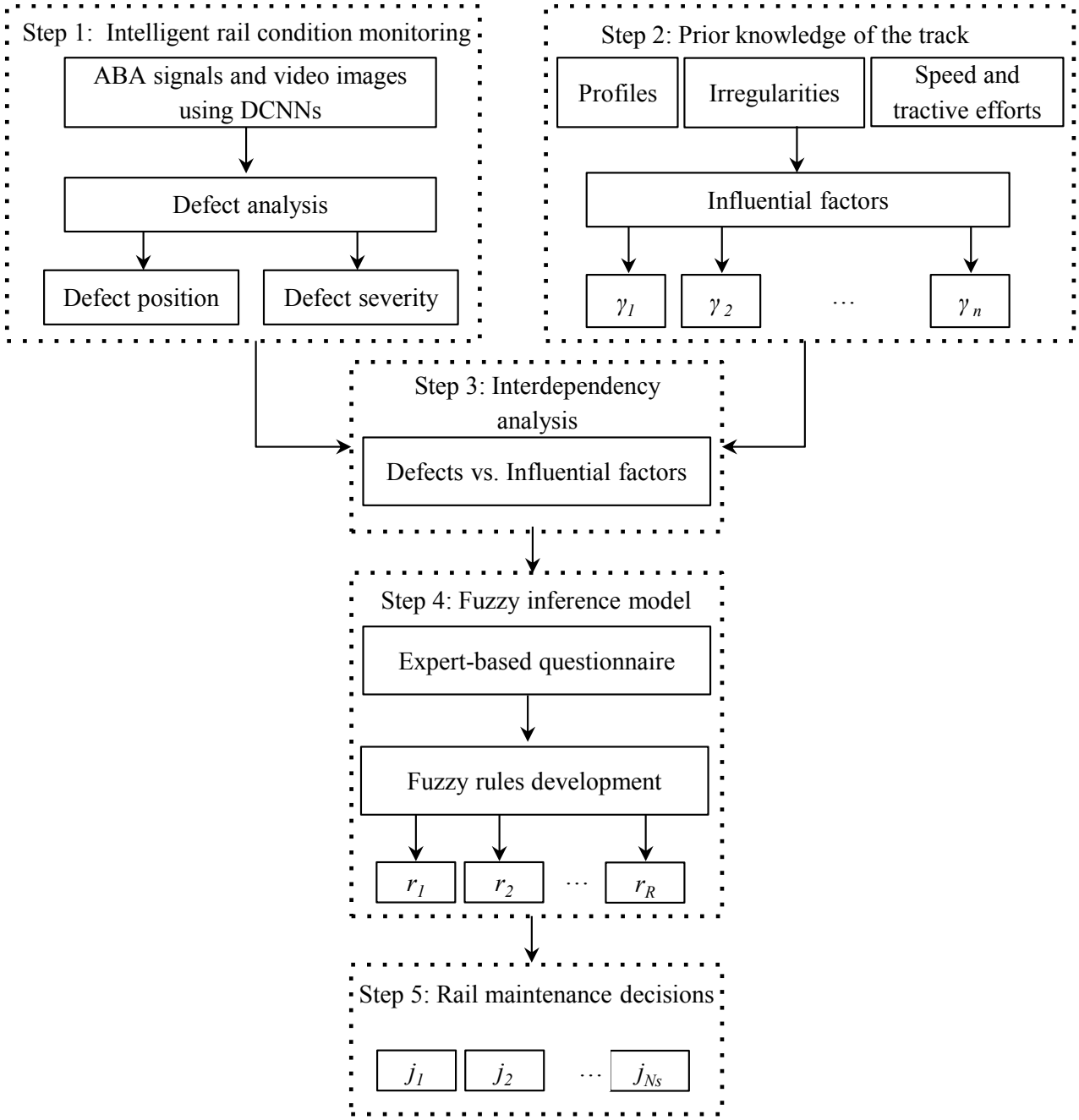


Figure 4.1 Flowchart of the proposed methodology

4.2 Step 1: Intelligent rail condition monitoring

In this chapter, we require a technology that can detect defects in an early stage. Hence, we consider to use ABA measurements [35]. To enhance the visualization, ABA measurements are combined with rail image videos [36]; [37]. In our case study, the ABA measurement and rail video images are used to study rail surface defects; specifically squats, as they are costly for the Dutch railway network. A global scheme of the measurement systems is given in Figure 4.2.

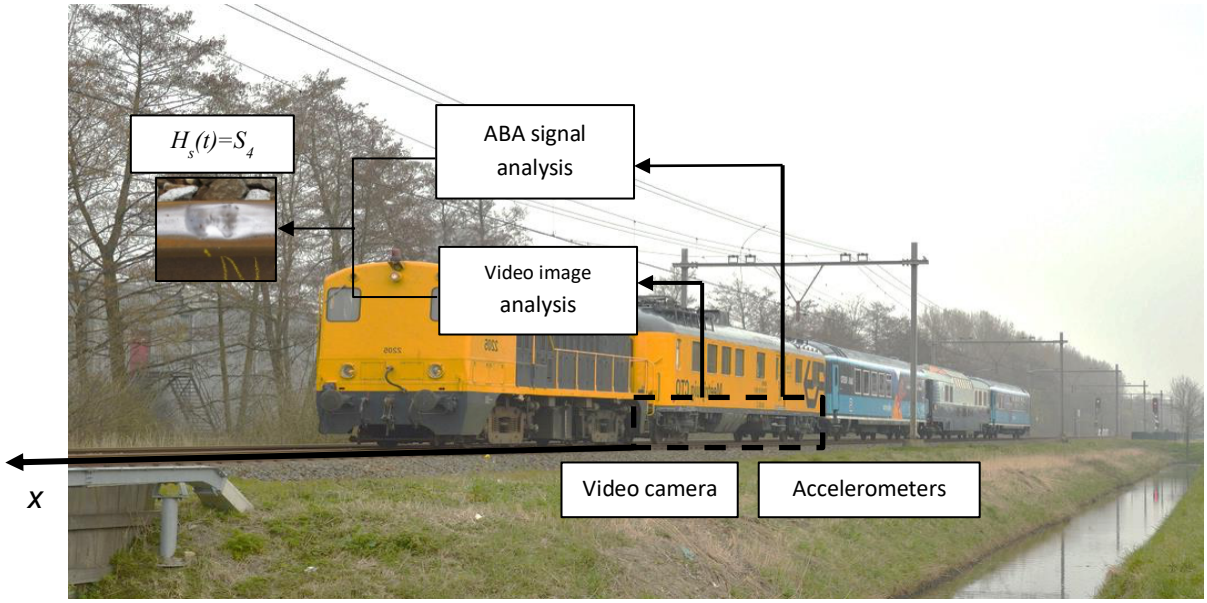


Figure 4.2 Defect severity analysis via ABA signals and image data using on-board train measurement. In the scheme, a severe squat is shown (S_4). The actual measurements were obtained from two different trains: CTO Train for the ABA signals, and Inspectation measurement train for the video images.

In [32] the feasibility of early-stage squat detection using an ABA system is shown. The ABA system can be employed to detect a range of surface defects, most importantly, squats, corrugations, and damaged welds, insulated joints and switches. The ABA system can be embedded on in-service operational trains. Four channels are assigned for the ABA measurement including left rail and right rail, and horizontal and vertical accelerations to capture all the ABA signals. A set of cameras with high frame rate are provided to capture high resolution images. The cameras are installed on the train measurement. A length of 15 cm of the track can be covered by each image.

Deep convolutional neural networks (DCNNs) have been applied for different problems in the area of classification due to its algorithmic advances [38]; [39]. We use a DCNN model in order to automatically estimate from the ABA signals the defect severity throughout tracks based on a big data analysis. For training the DCNN, based on previous results [36]; [34], we obtain a set of labelled images with their severity. The labels used from the image samples are on a scale from 0 to above 4 according to the severity level of the defects visible in the squats found by rail images analysis. Non-defect track images are assigned a value of zero and defects are assigned from 1 and above. The severity of the squat s can be used to represent the health condition of the rail, $H_S(t)$, at the time instant t of the measurement as follows

$$H_s(t) = \begin{cases} S_1 & \text{if} \quad 0 < L_s(t) \le 1 \\ S_2 & \text{if} \quad 1 < L_s(t) \le 2 \\ S_3 & \text{if} \quad 2 < L_s(t) \le 3 \\ S_4 & \text{if} \quad 3 < L_s(t) \le 4 \\ S_5 & \text{if} \quad 4 < L_s(t) \le 5 \end{cases} \tag{4.1}$$

where $L_s(t)$ is the measured level of severity, S_1 refers to a seed squat, S_2 is a light squat, S_3 is a moderate squat, S_4 is a severe squat, and S_5 is a squat with risk of rail break. The images and their severity are matched with their corresponding ABA signal. To do so, a window of the ABA samples is defined with a length of 3036 samples, covering full responses to local defects. This facilitates matching the signals with the video frames. Figure 4.3 shows two samples of image data used for the severity analysis associated with the corresponding ABA signals. The labelled data is therefore split into two parts for training and testing. To keep consistency in the defect detection, the labelled samples are collected from different locations over the measured track and they cover all the types of squats. They are compiled into a training set for each of the classes. The dataset was obtained by manual labelling of the images by an expert. The labelled sample defects are then divided into a training set and a testing set. The sample size was 125 squats. The distribution of the squat classes in terms of severity set is 70 samples for S_1, 8 for S_2, 6 for S_3, 8 for S_4, and 33 for S_5. 75% of the data is assigned for training and 25% for validating of the network performance. The samples of the labelled images are composed of 125 different squats collected from different locations of the track. We train a convolutional neural network regression model using the samples. The average binary accuracy (defect vs. non-defect) of the network on all tested samples is taken into account. Although putting a high acceptance threshold on the network output response might increase the rate of false positive detection, we use the threshold to detect the correct classes of the defects, seed (trivial) defects, and the normal classes. Once the DCNN for the image data is trained, defects in the large pool of previously unseen samples can be found.

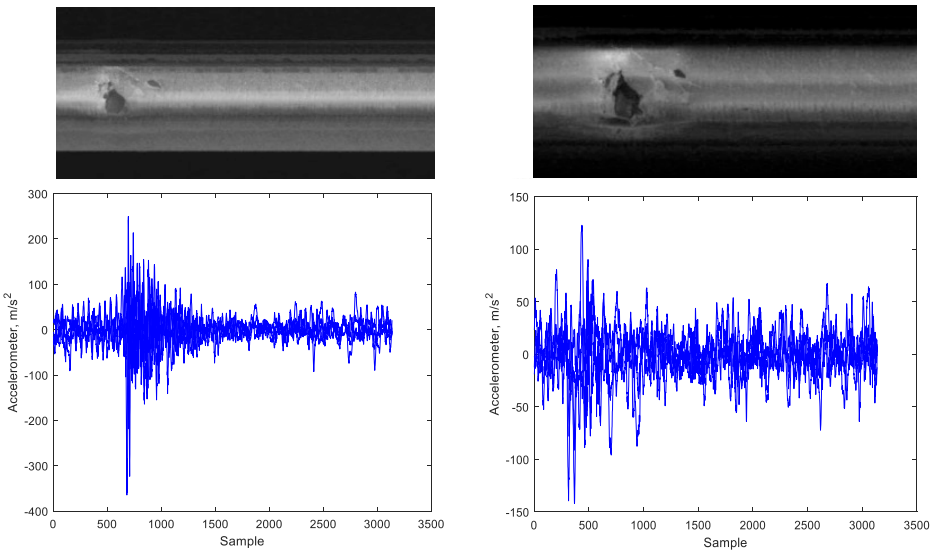


Figure 4.3 ABA signals including acceleration matched with rail video frames

Using a set of convolutional layers, the defect features are included in the DCNN model as filters to recognize distinguishing features and create a feature map. We use Rectified Linear Unit (ReLU) as activation function with max-pooling layers in order to down sample the outcome of

each layer [40]. The convolutional and pooling layer are finally attached to a sequence of three fully-connected layers to get class predictions (see in Figure 4.4).

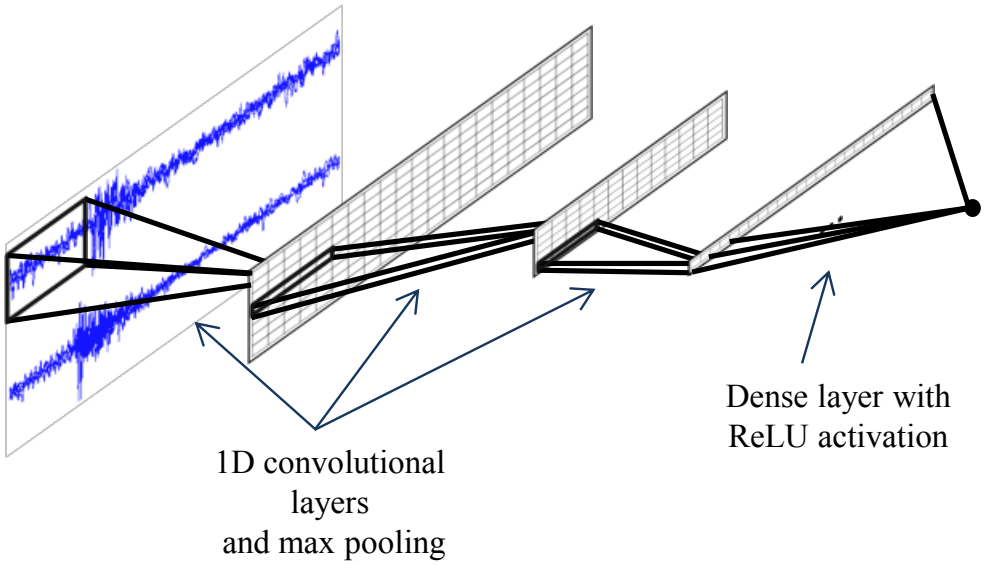


Figure 4.4 The prediction model consists of one-dimensional convolutional layers each followed by a max-pooling operation, and a ReLU activated dense layer on top, which results in the final scalar estimation of the severity.

The separating rail observations (detecting squats using DCNNs) from track characteristics (determined by influential factors) is one of the major contributions of this chapter. On the one hand, the DCNN is used to estimate the severity of the defects according to the ABA and image sources. This just gives the defect analysis (the rail observation in the Step 1) and not the rail health condition. On the other hand, track prior knowledge containing the influential factors can impact the rail health condition (Step 2) as those factors affect the quality of rail use over time (rail degradation). Thus, influential factors are collected to contribute the track characteristics for the estimation of rail health condition. For instance, a piece of rail positioned on a rail curve can get degraded faster than the same rail piece on a straight rail. To include track characteristics effects, the interdependency between the rail observations (the DCNNs) and the track prior knowledge is investigated in Step 3.

4.3 Step 2: Prior knowledge of the track

General characteristics of the railway track system can have a large influence in the initiation and growth of the rail defects. A list of some generic track characteristics that are potentially relevant to the appearance of rail defects are discussed next. The idea in this chapter is to take seven factors into account as “general characteristics of track” as according to the literature survey, they are proved to be significantly influencing in the initiation and growth of the rail defects. In particular, we classified the seven influential factors based on Step 2 into three categories (1) track profiles, (2) track irregularities and (3) operational speed profile and tractive efforts. However, there are other factors that can affect the track. As an example, train traffic can be influential and has an important role in the actual rail health condition. In this chapter, we assume that the influence of the traffic tonnage, which increases the amount of contact force between wheel-rail, can be seen in the defect severities (the rail observation). Furthermore, tonnage will be an influential factor when predicting defect evolution over time. During the same time period, the rail defects in segments with a higher tonnage evolved faster than the defects in segments with a low tonnage. Additionally, observations indicate that a higher number of defects will be found in tracks with a higher tonnage. These are

two possible ways to include the effect of the traffic tonnage in the proposed approach: (1) Indirectly via the effect of the tonnage in the rail observation. Condition monitoring measurements will automatically update both the appearance of new defects and the severity of the defects. (2) Directly via the inclusion of tonnage as influential factor. This case is most suitable when the infrastructure manager wants to predict the evolution of the defects; as tonnage will indicate how fast detected defects will evolve. In this chapter, we do not include the prediction of the defect evolution, so in this case the indirect method via rail observation is conducted. Part of the future research is to consider the effect of tonnage within a predictive approach. We employ various sources of information to obtain the prior knowledge of track using a big data analysis.

4.3.1 Track profiles

Deviations of the track alignments (vertical, lateral, etc.) with respect to the nominal alignment can lead to track irregularities [41]; [42]. [43] analyse the contact between wheel and rail conditions in the curved railway track to consider the influence of track profile on the initiation and growth of rail defects. [44] reviews the research on squats and squat-type defects. According to [44], squats appear mostly on straight and gentle curves, and those defects barely occur on sharp curves. Likewise, [35] report that squats in the Netherlands occur normally on straight tracks and gentle curves. On the contrary, another rail surface defect, called head checks, shows up mostly on the curve tracks with radii no more than 3000 m (sharp curves) [45]. In this chapter, the horizontal curvature of the track is taken into account. Furthermore, the rail segments are defined based on the rail curvature. In this way, only one influential factor for the horizontal curvature is considered for one segment. Vertical profile is ignored as those changes in the Dutch railway network are small.

4.3.2 Track irregularities

The track geometry changes from the design geometry due to trains passing over the track. More passing trains could worsen the track geometry condition. In the literature, the irregularity amplitude and wavelengths are mostly used as the controlling factors of the track quality. The limits for those controlling factors are typically analysed using measurements and dynamic simulations. The presence of track irregularities was found to have an influential effect on RCF defect appearance [46]. Track geometry problems are widely explained as one of the influential factors considering wheel-rail interactions, maintenance planning, and life of railway tracks. Irregularities have an impact on ride comfort and traffic safety level. All those influences are therefore very critical in railway dynamics. Nonetheless, the critical level is directly related to track usage. In the literature, there are also different studies about the influence of track geometry on the track condition and the track degradation. Thus, by considering the significant contribution of the track geometry on the track condition and then subsequent maintenance plans, control of track irregularities plays an important role on facilitating condition-based maintenance planning [47]; [48]. An infrastructure manager requires to employ the geometry measurements for the maintenance planning [49]; [50]. A maintained track geometry considerably contributes not only to train safety but also track health condition. Furthermore, track geometry monitoring could help to prolong the effective track life time by managing the track degradation, the track health condition and subsequently cost of the maintenance operations [42].

The literature surveys show that measurement data has been used to develop statistical modelling of railway track irregularities in the last three decades. Track safety and ride comfort are among the first track irregularities analysis using field data. [51] discusses the impact of track quality on track maintenance decisions and performance-based analysis of track geometry using a statistical model for a long track. The paper develops a degradation-based track condition model to

explain interaction between rail defects and performance indicators. A similar investigation has been carried out using linear models to capture the track response to a train load in terms of track irregularities and potential appearance of rail defects [52]. [53] use a comprehensive model to predict the track quality for maintenance operations. They employed multiple data of traffic and train speed, track structure, maintenance time slots to determine the track quality at two consecutive time periods. In this chapter, based on the available data, we select three sets of irregularity-related influential factors including (1) the vehicle effect, which is a signal indicating the train ride quality based on several geometry measurements and operating trains characteristics, (2) track geometry, which is an indicator estimated based on a combination of different track geometry measurements such as horizontal alignment, the vertical alignment and cant differences, and (3) track super elevation which is the difference between the designed cant and the measured cant.

4.3.3 Operational speed profile and tractive efforts

Tractive effort and curving in the track are found to be potentially responsible for RCF-type rail damages [54]. The review of the squat defects in [44] reveals that these defects can come with driving traction i.e. locomotives, compared to curving traction. Observations by [55] show the relationship between braking and squat occurrence in the Dutch railway network. The authors conclude that the traction performance of the rolling stock has a large influence on the initiation and growth of squats. They found many squats at pieces of a track where the gradient of the speed was the highest and the speed was low. Moreover, the low speed was also influential, as more use of the Anti-lock Brake System (ABS) system happens at lower speeds. Tractive and braking efforts, which differ by the types of locomotives, can also influence on the occurrence of RCF defects. A comparison is made between different locomotives in different operational situations in Australia to investigate the initiation and development of squats in the rail head [56]. [57] finds that the most frequent kilometre positions for the squat are locations with low-speed running associated with high wheel slip and low adhesion. [57] investigated the traction characteristics of the typical traction motors to find the potential link between the generation of defects and the rolling stock type. In this chapter, the speed profile of the typical rolling stock is investigated to determine its potential correlations with the defect occurrence. The related effects are considered in this chapter including: (1) train speed profile, which is the speed of the measurement train in km/h, (2) train acceleration profile, which is the acceleration of the measurement train in m/s^2 and (3) rail head wear, which estimates the difference between the measured height of the railhead and the nominal height of a new rail railhead in mm. The measurements were obtained with tacho signals, accelerometers and scanning laser sensors mounted on the measurement train.

4.4 Step 3: Interdependency analysis

According to the track prior knowledge explained in Step 2, those track factors that are observed to be influential on rail conditions in the Dutch railway network are considered. We use the data available in the Dutch railway infrastructure monitoring system (BBMS) to acquire the signals of the influential factors. In this chapter, we use both dynamic and static measurements to obtain the influential factors. After processing the measurements, the influential factors are calculated for a single measurement campaign. Part of the further research includes the use of historical measurements to study the evolution of the influential factors over time. Hence, seven signals are chosen as influential factors which might significantly affect the rail condition containing (1) train speed profile, (2) train acceleration profile (3) rail head wear, (4) track horizontal curvature, (5) vehicle effect, (6) track geometry parameter and (7) track super-elevation. In Figure 4.5, a map is

employed to show track including all the seven influential factors. The data are captured over the whole track to analyse the dynamics of the track influential factors. Hence, on the one hand we have a set of data over the track representing the track knowledge and on the other hand, squats are detected along the track with their severity and location using the ABA signals and the image data. The interdependency is defined by investigating how to match the location and the severity of a certain defect with the signals of the track influential factors. To do so, the track is partitioned into different segments and the interdependency is investigated per segment.

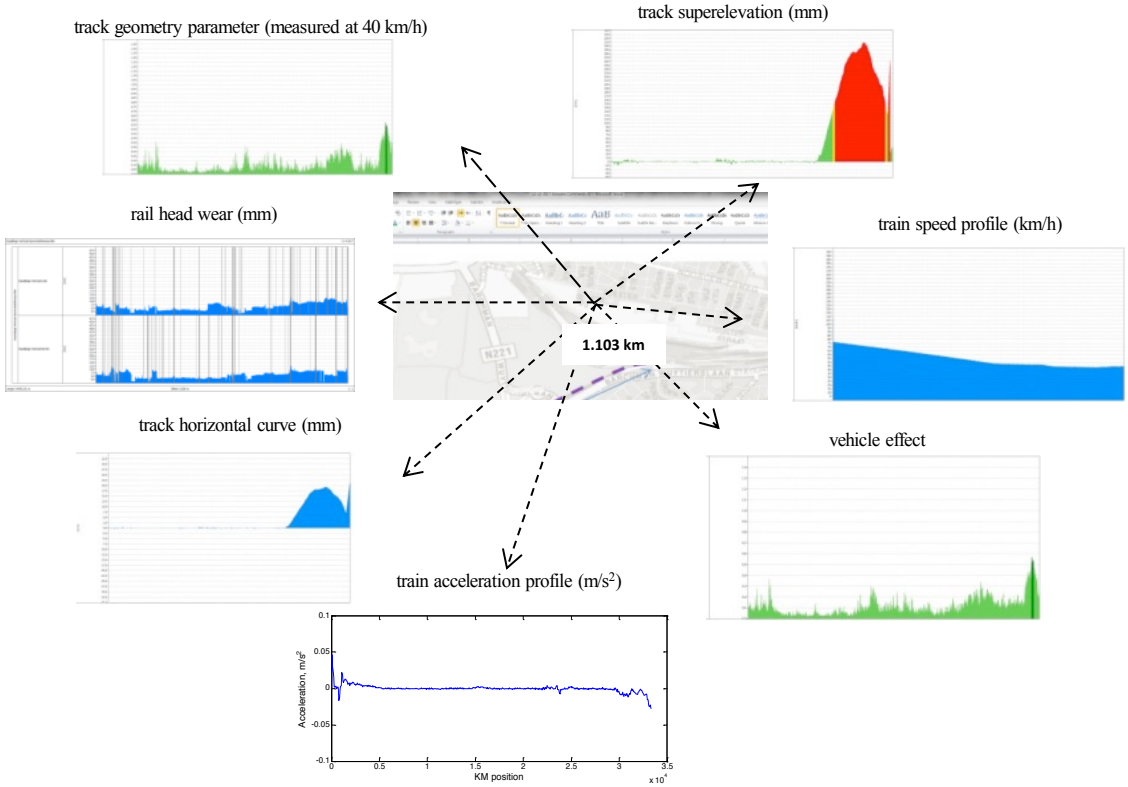


Figure 4.5 A schematic of the GIS map of the track influential factors in a piece of the track shown by the red line.

To numerically represent the severity of a segment, we consider the average of the severities of all the squats that are located in segment *j*:

$$\aleph_j(t) = \left(\frac{\sum_{\text{s in segment j}} H_s(t)}{\sum_{\text{s in segment j}} \delta_s(t)} \right) \tag{4.2}$$

where $H_s(t)$ is the severity of the squat *s* provided by the ABA detection algorithm for the measurement time *t*. The function $\delta_s(t)$ equals to 1 when *s* is a squat, and equals 0 otherwise. Regarding the processing of the datasets, once all the data sets (signals) over the track are acquired, the signals are processed according to equations (3) and (4). First, signals are normalized using (3), and then the influential factor is obtained by the average of the signal as in (4). The influential factor is then a "representative" value of the measured signal for that segment. So the signals should all be normalized between L_{int} and L_{end} which are respectively the

upper bound and lower bound of the interval selected for the normalization. The function can be expressed as:

$$\gamma_{j,Nor}^{k}(x,t)=\frac{\left(\gamma_{j}^{k}(x,t)-\gamma_{j,\min}^{k}(t)\right)\left(L_{\text{end}}-L_{\text{int}}\right)}{\gamma_{j,\max}^{k}(t)-\gamma_{j,\min}^{k}(t)}+L_{\text{int}} \tag{4.3}$$

where $\gamma_{j}^{k}(x,t)$ is the data for the *k-th* influential factor at the location *x* and time instant *t*, $\gamma_{j,\min}^{k}(t)$ and $\gamma_{j,\max}^{k}(t)$ are minimum and maximum values of the signal at the segment *j*. The variable *t* is defined in case we have different measurements over different periods; then, we can define all the influential factors over time. By considering $x_{j,avg}^{k}$ as the location where average value of the data occurs (as the representation of each segment), the data value for the segment *j* is calculated according to:

$$\gamma_{j}^{k}(t)=\gamma_{j,Nor}^{k}(x_{j,avg}^{k},t) \tag{4.4}$$

where $\gamma_{j}^{k}(t)$ is the influential factor for the segment *j* and the time step *t*. By having a matrix containing $\gamma_{J}^{k}(t)=\left[\left(\gamma_{1}(t)\right)^{T},...,\left(\gamma_{N_s}(t)\right)^{T}\right]$ where N_s is the total number of segments and $\gamma_{j}(t)=[\gamma_{j}^{1}(t),...,\gamma_{j}^{n}(t)]$, a clustering model is provided. Due to the simplicity of a method called Fuzzy C-Means, in this chapter we use this method for the clustering problem [58]. Just for illustration, three clusters are defined over the influential factors. The membership degree to the cluster determines how much a segment belongs to the cluster. The track is partitioned into 15 segments. Figure 4.6 shows a schematic view of the clusters. As seen in the figure, segment 5 is highlighted by a rectangle indicating a high membership degree of the cluster 2 in the segment indicated by an arrow. Rail segment 4 has the higher membership to cluster 1; however, it does not belong to the cluster 1 as much as segments 1, 2, 3, 6, 14 and 15, which they all have membership values near to one. The results are used in order to obtain rail health condition decision rules.

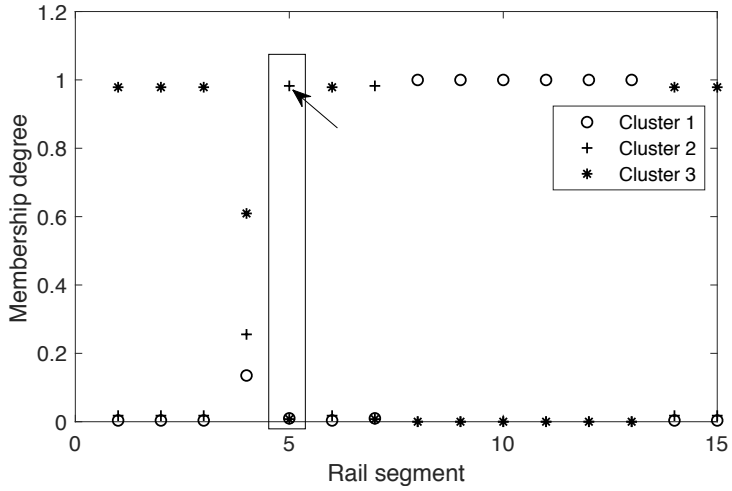


Figure 4.6 Membership degree for all the segments based on their influential factors. Highlighted by a rectangle, indicating that rail segment number 5 belongs highly to the cluster 2 (membership degree almost one, indicated by the arrow).

In this chapter, five levels are defined including very low (L_1), low (L_2), medium (L_3), high (L_4) and very high (L_5) to represent the interdependency between the defect severity and the influential factors (For simplicity and interpretability of the data, linguistic terms such as very low, low, medium, high and very high can be defined). We rely on expert opinion to select the proper level.

4.5 Step 4: Fuzzy inference model

In this chapter, a fuzzy inference system is used to develop rules about rail health condition based on the influential factors $\gamma_j^k(t)$. The Mamdani fuzzy system approach is considered due to its interpretability and simplicity [59]; [60]. To explicitly express the inference system, the Mamdani inference per rule can be defined as follows:

$$Y_j^m(t) = f_{\text{Mamdani}}\left(\gamma_j^1(t), \gamma_j^2(t), ..., \gamma_j^k(t), ..., \gamma_j^n(t)\right) \tag{4.5}$$

where $Y_j^m(t)$ is the rail health condition in section j and $\gamma_j^k(t)$ the influential factor k in section j. Figure 4.7 shows the architecture of the inference model. In the first layer is to use the values of input variables, e.g. $\gamma_j^k(t)$. The membership degrees of the inputs to the fuzzy values are obtained in layer 2 and employed to compute the rule truth values in layer 3. At the layer 4, according to the rule truth values, the rail health condition of each rule in the segment j is estimated.

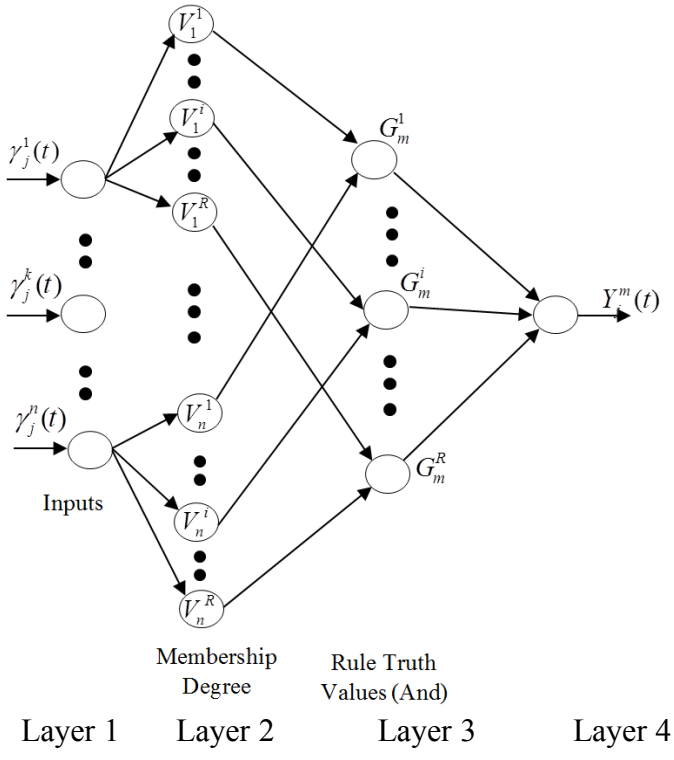


Figure 4.7 Generic structure of the fuzzy decision model to compute the rail health condition

The *R* fuzzy if-then rules are generated based on Equation (4.5) to capture combinations of the influential factors. The purpose is to assign a membership degree to each $\gamma_j^k(t)$. Gaussian membership functions are used to fuzzify the inputs. In this chapter, the Gaussian type of membership function is used as it is smooth and nonzero at all points [61]; [62]. The Gaussian membership function is based on two parameters and can be represented as:

$$Gaussian\left(x;\ c,\ \sigma\right) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \tag{4.6}$$

where for each membership function, c and σ are the parameters of the membership function. The parameters are tuned so that every membership function has around 30 percent overlapping with the neighbouring functions. The rule r_i can be expressed as:

$$r_i: \text{If } \gamma_j^1(t) \text{ is } V_1^i \text{ and .... } \gamma_j^k(t) \text{ is } V_k^i \text{ and .... and } \gamma_j^n(t) \text{ is } V_n^i \text{ then } Y_j^m(t) \text{ is } G_m^i \tag{4.7}$$

where V_k^i is the fuzzy set related to input variable $\gamma_j^k(t)$ and G_m^i is the fuzzy set of the rail health condition selected based on the expert judgment for rule r_i. The minimum of the fuzzified input values is given as the rule truth value of each rule. The fuzzy set of the output is obtained by the Mamdani union operator over all the rules. To defuzzify the output, the centre of gravity approach is applied so to obtain a crisp value. The fuzzy inference system (Mamdani) is to map the inputs (the influential factors) to the output (the rail health condition) using a set of fuzzy rules. Thus, the fuzzy rules are components of the fuzzy inference system. To set the fuzzy rules, a questionnaire is provided to systematically analyse the combinations of possible inputs. As the judgment relies on the expert knowledge, it is prone to bias. Thus, the investigation is used to support the experts on the validation of the judgements. The inclusion of the investigation results in the questionnaire, helps the expert to visualize the effect of $\gamma_j^k(t)$ over the segment *j* on the actual rail health conditions. Furthermore, as the questionnaire will lead to a model of the rail condition using the knowledge of expert, the expert qualified to fill out the questionnaire is a rail maintenance engineer or a rail inspection expert. The expert should have experience with both rail monitoring and rail maintenance. By using the proposed methodology, the infrastructure management company will benefit from systematically keeping the knowledge of rail experts in the company. So, in case a rail expert is not available, the railway company can still use the previously developed rules or update them according to new infrastructure requirements. In the questionnaire, two options are given including “influential” “non-influential”. Then, the experts are asked to rank between 1 and 2 the effect of the combination of influential factors into the health conditions of the rail. A major contribution of the fuzzy system is to include non-crisp values (fuzzy values) in the output (the rail health condition). Although a binary approach is used for the questionnaire, (1) we can capture the fuzzy dynamics on the rail health condition and (2) we cover all the rule combinations. Otherwise and with using five-level ranking, number of the rules created would be too much time consuming for the experts whereas some of those rules would be useless in the decision making. Moreover, the five-level ranking is used to just improve the visualization quality of the interdependency analysis. The questionnaire is converted into a fuzzy inference system, where the rules are given by the options of the questionnaire (two possible fuzzy sets per influential factor) and the output fuzzy sets of each rule are given by the answers of the experts (three possible fuzzy sets).

4.6 Step 5: Rail maintenance decisions

After estimating the rail health condition for each segment, the entire rail can be evaluated according to the estimated health condition. The aim is to find the most critical pieces of the track for the condition-based grinding planning. Squats can be considerably treated by grinding completely when they are at an early stage of growth or effectively keep at safe level (to avoid having disastrous consequences) when they are severe. In this chapter, a clustering method is proposed to grind the most critical pieces of the track efficiently based on defined maintenance time slots determined by the infrastructure manager. As different tracks have different maintenance time slots, it is important to consider the available time slots to carry out the grinding operation. In the Dutch railway network, the time slots vary from one railway station to the next railway station. This means that not all segments of a long track that include different railway stations have the same maintenance time available for doing grinding. Hence, the grinding planning can be formulated as Figure 4.8. As depicted in Figure 4.8, if maintenance time is still available after the grinding, the clustering approach can be applied to the next critical track segment ranked second by the expert system to effectively utilize the whole available maintenance time slot. Hence, the infrastructure manager makes sure that the maintenance time is fully used to avoid inducing extra maintenance costs. The clustering approach strives to cover as many severe defects using as few clusters as possible within the limited maintenance time slot, which usually takes 8-10 hours at night in the Dutch railway network (this depends on the type of operations, and it could change per day, week and year). The proposed clustering approach assigns a defect, e.g. a squat, to a cluster. The model includes not only the squat position, but also the squat severities acquired by the ABA system measurement and rail image data. The proposed grinding model is elaborated in the previous work of the authors [34]. Table 4.1 presents the notations used in the model.

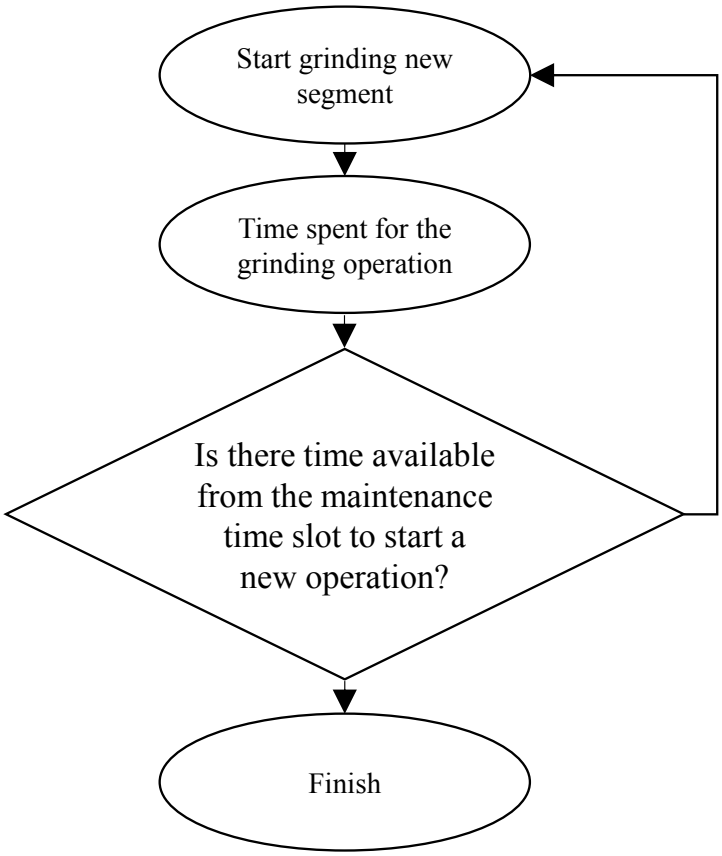


Figure 4.8 The proposed simplified grinding planning scheme.

Table 4.1 A list of the notations of the clustering model

Notations	*Definitions*
N_c	*Number of clusters*
N_d	*Number of squats*
$\underline{\xi}$, $\overline{\xi}$	*Track starting and ending positions, respectively*
Ω	*Defect severity*
T_s	*Setup time for a grinding maintenance operation.*
T_t	*Duration of maintenance time slot*
d_g^{start} *and* d_g^{end}	*Track starting and ending positions of the g-th cluster*
Δd_{min} *and* Δd_{max}	*Minimum and maximum size of each cluster*
v_G^{on}	*Grinding machine speed running over the track*
v_G^{off}	*Driving speed of the grinding machine when the machine is off*
T_G^{on}	*Time needed to switch from grinding to driving*
T_G^{off}	*Time needed to switch from driving to grinding*

We call $\left[\underline{\xi},\overline{\xi}\right]$ the physical range, and clusters located within the physical range are called active clusters. Also, the setup time, T_s, typically includes the machine travelling time, preparation time and finishing time for a maintenance operation. The start and end locations of the *g*-th cluster are the decision variables of the clustering problem. Thus, the grinding model can be formulated as the following non-smooth optimization problem [34]. The indicator function I_σ takes value 1 if the statement σ is true, and 0 otherwise. The first term in the objective function (4.8) rewards the squats covered by a cluster depending on their severities, while the second term serves to maximize the number of non-active clusters, i.e. minimize the number of active clusters. The second term in (4.8) counts the number of clusters outside the physical range, i.e. non-active squats. As the total number of available clusters *Nc* is fixed, maximizing the number of non-active clusters is equivalent to minimizing the number of active clusters. The active cluster is defined via the kilometre positions of its start and end points within the physical range $\left[\underline{\xi},\overline{\xi}\right]$. A non-active cluster is outside the physical range and has no physical meaning. We use the idea of non-active cluster to be able to have idle clusters. Also, X_l is the kilometre location of the *l*-th squat. Constraints (4.11)-(4.12) set the distance range of the clusters. Note that the upper bound ξ_{max} is set as indicated to allow the situation of non-active cluster, i.e. all clusters are located outside the physical range $\left[\underline{\xi},\overline{\xi}\right]$. The term ε in (4.12) is included to avoid the overlapping of clusters. To determine ε, we suggest just to take a tiny positive value (like ε =0.001 m). When ε is high, the distance between clusters will be higher and interesting rail pieces might not get covered by a cluster. Constraint (4.13) restricts the size of each cluster.

$$\max_{\{d_g^{start},d_g^{end}\}_{g=1}^{N_c}} \sum_{g=1}^{N_c}\sum_{l=1}^{N_d} \omega_l I_{d_g^{start}\le X_l\le d_g^{end}} + \sum_{g=1}^{N_c} I_{d_g^{end}>\overline{\xi}} \tag{4.8}$$

subject to

$$I_{d_g^{start} \le X_l \le d_g^{end}} = \begin{cases} 1 & if \quad d_g^{start} \le X_l \le d_g^{end} \\ 0 & otherwise \end{cases} \tag{4.9}$$

$$I_{d_g^{end} > \bar{\xi}} = \begin{cases} 1 & if \quad d_g^{end} > \bar{\xi} \\ 0 & otherwise \end{cases} \tag{4.10}$$

$$d_1^{start} \le \underline{\xi} \tag{4.11}$$

$$d_{N_c}^{end} \le \bar{\xi} + 2N_c(\Delta d_{\min} + \varepsilon) \tag{4.12}$$

$$\Delta d_{\min} \le d_g^{end} - d_g^{start} \le \Delta d_{\max} \qquad \forall g \in \{1,...,N_c\} \tag{4.13}$$

$$d_{g+1}^{start} - d_g^{end} \ge \varepsilon \qquad \forall g \in \{1,...,N_c - 1\} \tag{4.14}$$

$$d_g^{start} \le \bar{\xi} \Rightarrow d_g^{end} \le \bar{\xi} \qquad \forall g \in \{1,...,N_c\} \tag{4.15}$$

$$\sum_{g=1}^{N_c} I_{d_g^{end} \le \bar{\xi}} \cdot \left(\frac{d_g^{end} - d_g^{start}}{v_G^{on}} + T_g^{on} + T_g^{off} \right) + \sum_{g=1}^{N_c - 1} I_{d_g^{end} \le \bar{\xi}} \cdot \frac{d_{g+1}^{start} - d_g^{end}}{v_G^{off}} \le T_t - T_s \tag{4.16}$$

The minimal and maximal size of a cluster is indeed determined by operational considerations of the grinding machine. The minimal size of a cluster is usually set to be the shortest length that the grinding machine can manipulate. The maximal size of a cluster should be less than the length of the rail considered. Constraint (4.13) restricts the size of each cluster. Constraint (4.14) ensures that the clusters are not overlapping, where the small positive parameter ε is the minimum distance between two clusters. So, there may be track sections between clusters that will be not included in the grinding planning. The constraint (4.15) forbids fractional clusters. The fractional cluster means that the start and end points of a cluster must both be inside or outside the physical range. We only allow to use active clusters (start and end points are both inside the physical range) and non-active clusters (start and end points both outside the physical range). The constraint (4.16) is the time limit constraint to ensure that the resulting clusters can be processed within the given maintenance time slot. The left-hand side of constraint (4.16) computes the total maintenance time, including the time to grind the active clusters (first term), the time for the machine to travel between the clusters (second term), and the setup time T_s. Constraint (4.16) guarantees that the total maintenance time to execute the clustering plan is less than the duration of the maintenance time slot T_s. The non-smooth optimization problem (4.8)-(4.16) can either be solved by gradient-free algorithms like pattern search and genetic algorithms, or transformed into an MILP problem following the standard procedure described in [64]. In [34], the clustering method was employed as part of the low-level optimization, in a setup where the decisions are based on prediction including uncertainties via a scenario-based chance-constrained approach.

4.7 Numerical results

The track Amersfoort-Weert in the Netherlands is selected as a case study (nearly 125 kilometers of track). The track passes through Utrecht, Geldermalsen, 's-Hertogenbosch, and Eindhoven to reach the destination (Weert) (Figure 4.9). The whole track is partitioned into 15 segments to take all the signals of the influential factors per segment into account. Also, the definition of the segments is based on track curvature, which means that each curve is included into one segment regardless the segment sizes. The squat problem is aimed in the case study due to the fact that: (1) squats are

one of the most commonly observed defects on rails, (2) squat-related costs are more than 5000 euro/km per year in The Netherlands. Although the rail grinding helps to treat all type of rail defects, e.g. corrugation, head checks and wheel burn, the optimal maintenance decisions proposed in the current chapter focus on the squat problem and for the other rail defect types, it is crucial to take the effect of those defects in the maintenance decisions into account. For the estimation of the actual rail conditions, as explained in Section 4.2, the images are analysed using image processing to detect the ones including squats.

Figure 4.9 Schematic track map between two stations, Amersfoort and Weert.

The rail image analysis is defined based on the images as input. The images are rescaled to 375×275 pixels. The three joined layers changes the high-level features through the previous layers, into three different classes defining the normal rail, rail with trivial defects, and rail with severe squats. To avoid mismatching between rail images and the ABA signal, first the frames of video images are pre-processed to eliminate the overlaps between two consecutive video images. Then, we align images with the ABA signal using GPS tags and different reference points of the rail infrastructure (such as switches, crossings, joints, etc.). To train the network, a set of manually labelled examples is collected from several locations along the measured track and is compiled into a training set for each one of the 3 classes (normal, trivial defects and squats). Once the network for the image data is trained, it is used to find squats in the large pool of previously unseen samples. By selecting a high value for the accuracy threshold on the output, the true cases of defects as well as the normal classes can be detected. The dataset was obtained by manual labelling of the images by an expert. The samples are chosen from different parts of the track, and they cover all the types of squats. The labelled sample defects are then divided into a training set and a testing set. We train a convolutional neural network regression model using the samples. Figure 4.10 shows the mean absolute error as a function of the training epoch of the network for both training and validation data. The convolutional neural network helps to distinguish between normal rail with abnormal pieces which contain squats. Figure 4.11 shows the comparative predictions and ground truth values for all samples, in the test set. Thus, although we believe that the number of samples is limited, as the samples were picked up from different locations and vary from light to severe squats, the dataset covers all the interesting cases.

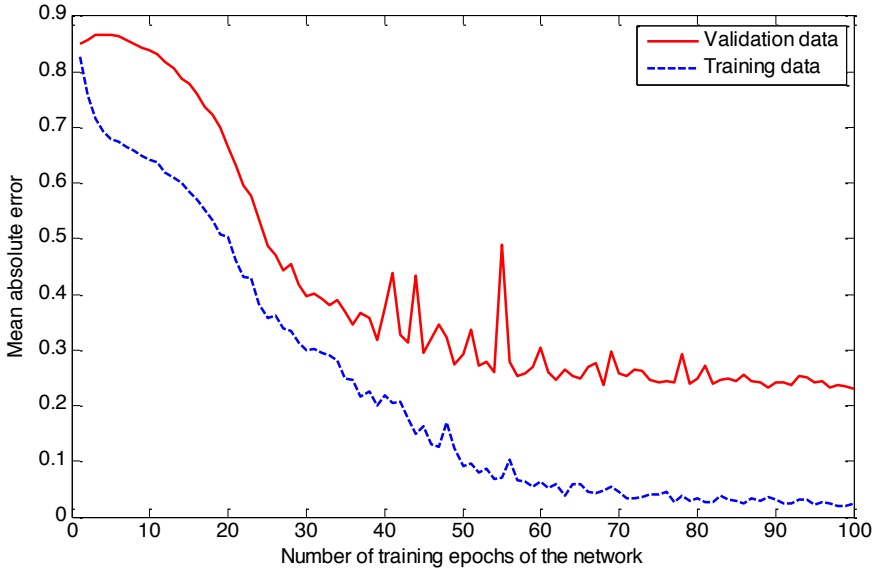


Figure 4.10 Mean absolute error (MAE) of the ground truth severities and the predictions. The network is trained using 75% of the data and is validated on the remaining 25%.

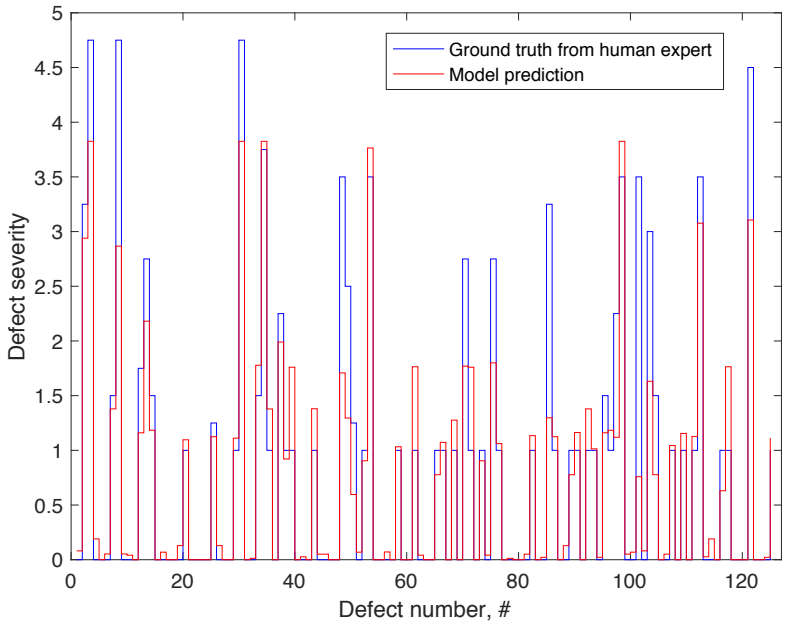


Figure 4.11 Ground truth values provided by a human expert by estimating the defect severities from defect images vs prediction of the severity level from the ABA signal.

Finally, the trained model is used with the new samples provided from the target track and finally predictions based purely on ABA are calculated. Figure 4.12 shows a sample plot of the results by the detection algorithm, which are used as the rail actual health condition, and shows the position of the defects and their severity.

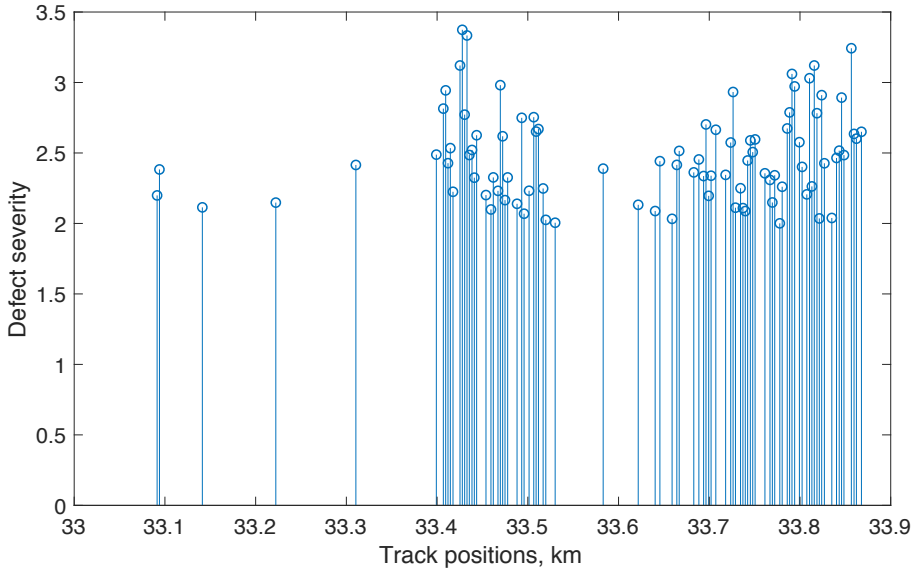


Figure 4.12 A sample of defect locations versus defect severity between kilometre 33 and 33.9 in the track Amersfoort-Weert.

The training time is approximately 41 hours for 1500 samples. A large pool of unseen rail images is employed to predict the trained network with much less computational time compared to the training time (30 seconds for 15000 samples). The prediction results show average accuracy of 97 % when training data covers 80% of the labelled dataset while 20% of the dataset is assigned for testing.

In Figure 4.13, to perform the interdependency analysis we have compared the severity of a segment (rail defect problems) with each influential factor (track characteristic). This information can be used to guide the design of fuzzy rules created from interview with experts about the relation between health condition and influential factors. Based on the interdependencies, a set of fuzzy rules are defined to estimate the health condition based on the influential factors as obtained in Equation (4.3). All the rules are given with the same weight. Moreover, all the input variables are combined through the rules. In this chapter, 127 fuzzy if-then rules are generated in order to meet the possible interdependencies. Furthermore, based on the fuzzy rules, the sensitivities of the health condition to the influential factors are captured as shown in Figure 4.14. This figure presents how the influential factors model the rail health condition, varying from fully healthy (severity equal to zero) to completely unhealthy (severity equal to one), while all the other influential factors are assumed to be fully healthy (equal to zero). Three plots are used to show the sensitivity.

Variation of the inputs of an expert in the questionnaire can lead to different final maintenance decision results. Several experts are asked to fill out the questionnaire so that variations cause by single expert are reduced. Among all the influential factors, train speed has the highest effect on the grinding decision and superelevation has less influence. An increase of 20% in the train speed related influential factor gives an 8% increase on the rail health condition, whereas an increase of 20% in the superelevation related influential factor gives 5%. A misestimating of 20% in a single factor gives at most 8% difference in final results error in the case of changing train speed related influential factor with superelevation related influential factor.

As an example, in Figure 4.14(a), the effect of two input variables namely train speed, $\gamma_j^1(t)$, and train acceleration profile, $\gamma_j^2(t)$, respectively, is presented. As shown in the figure, the train speed

changes over the track affect more the rail health condition in comparison with train acceleration profile. This is a good indicator of the importance of the train speed for the maintenance decisions. Figure 4.14(b) depicts the influence of the speed profile versus supperelevation, $\gamma_j^7(t)$. The plot shows that the rail health condition cannot get excited by the influence of the supperelevation as same as the effect of the train speed profile. In Figure 4.14(c), vehicle effect, $\gamma_j^6(t)$, is compared with the supperelevation $\gamma_j^7(t)$. As can be seen in the plot, both factors are not as influential as train speed and train acceleration on the rail health condition. However, the vehicle effect can influence more in comparison with the supperelevation.

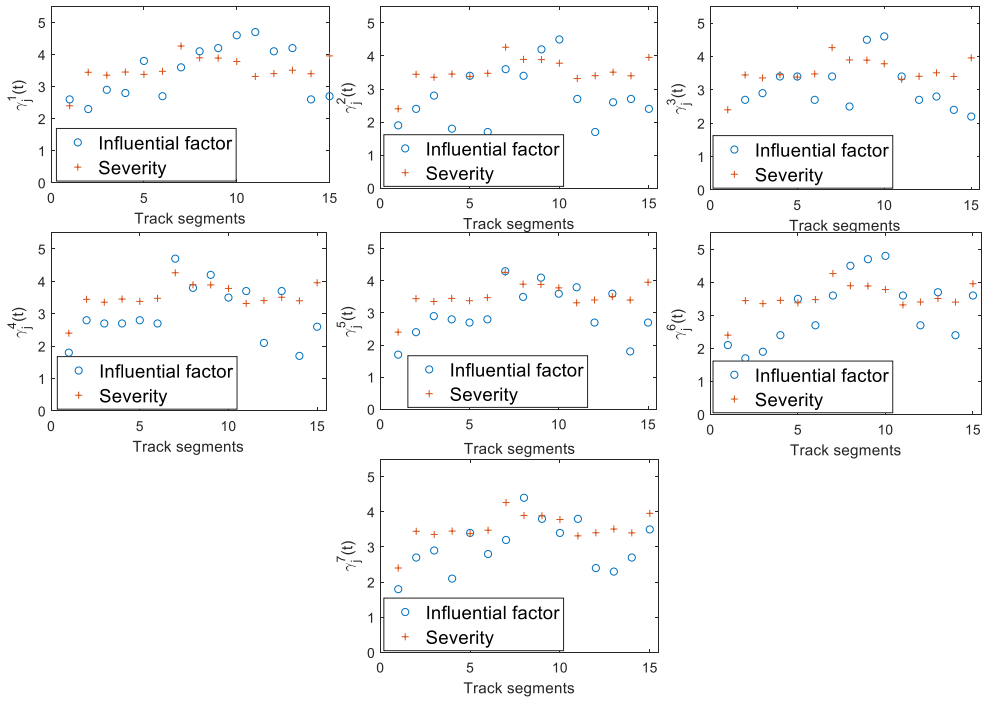


Figure 4.13 Interdependency analysis between defect severity and track influential factors over 15 track segments where, $\gamma_j^1(t), \gamma_j^2(t), \gamma_j^3(t), \gamma_j^4(t), \gamma_j^5(t), \gamma_j^6(t)$ and $\gamma_j^7(t)$ are train speed profile (m/s), train acceleration profile (m/s2), track horizontal curve (mm), track geometry parameter (measured at 40 km/h), rail head wear (mm), vehicle effect, track superelevation (mm), respectively.

However, the vehicle effect can influence the health conditions more in comparison with the superelevation. Therefore, an increase in the most influencing factors, i. e. $\gamma_j^1(t)$ and $\gamma_j^2(t)$ can increase the criticality of the segment up to requiring maintenance. If this criticality goes beyond the given rail health conditions of other rail segments of the track, then the grinding decision changes directly. Therefore, the infrastructure manager should take the segments with higher train speed profile and train acceleration into account in the maintenance plan.

Figure 4.15 also indicates that j_7, j_8, j_9, j_{10} obtain the highest values of the rail health condition. It means that those segments have a critical health condition compared to other rail segments. These segments highlighted by the red line in the figure belong to the track between railway stations Geldermalsen and 's Hertogenbosch. Furthermore, the rail actual condition (rail observation) is depicted in Figure 4.15.

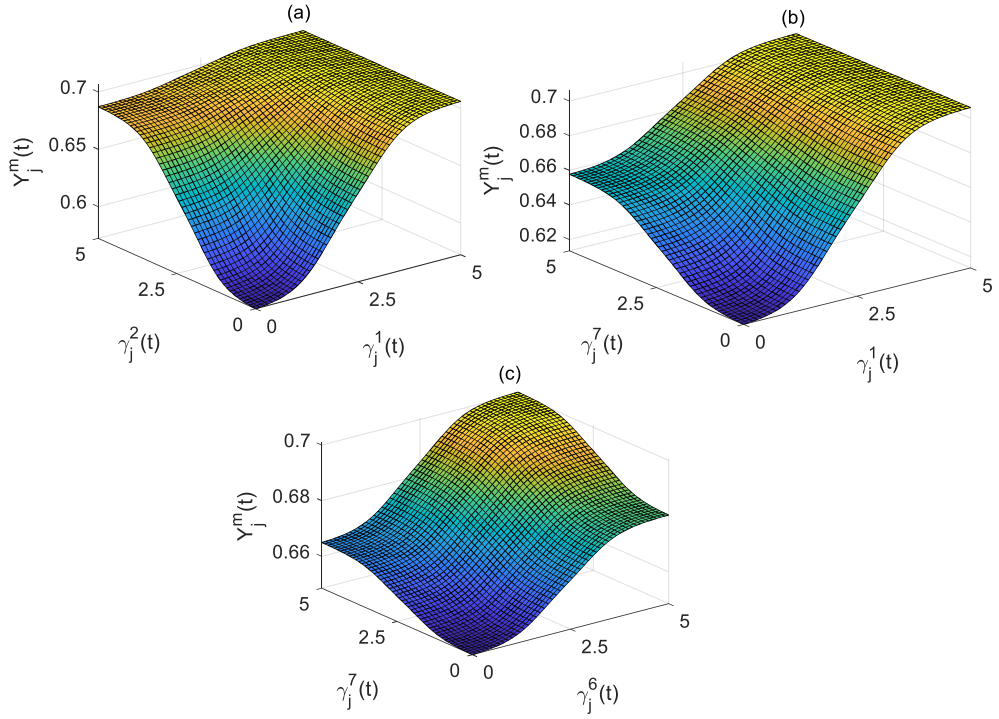


Figure 4.14 Examples of how the fuzzy rail condition rules are based according to the interdependency analysis

Table 4.2 Calculated influential factors per segment and estimated rail health condition using the proposed fuzzy inference system.

Segments	$\gamma_j^1(t)$	$\gamma_j^2(t)$	$\gamma_j^3(t)$	$\gamma_j^4(t)$	$\gamma_j^5(t)$	$\gamma_j^6(t)$	$\gamma_j^7(t)$	$Y_j^m(t)$
1	2.6	1.9	1.1	1.8	1.7	2.1	1.8	0.6507
2	2.3	2.4	2.7	2.8	2.4	1.7	2.7	0.6522
3	2.9	2.8	2.9	2.7	2.9	1.9	2.9	0.6521
4	2.8	1.8	3.4	2.7	2.8	2.4	2.1	0.4929
5	3.8	3.4	3.4	2.8	2.7	3.5	3.4	0.6683
6	2.7	1.7	2.7	2.7	2.8	2.7	2.8	0.6481
7	3.6	3.6	3.4	4.7	4.3	3.6	3.2	0.6957
8	4.1	3.4	2.5	3.8	3.5	4.5	4.4	0.6982
9	4.2	4.2	4.5	4.2	4.1	4.7	3.8	0.6938
10	4.6	4.5	4.6	3.5	3.6	4.8	3.4	0.6949
11	4.7	2.7	3.4	3.7	3.8	3.6	3.8	0.6721
12	4.1	1.7	2.7	2.1	2.7	2.7	2.4	0.6803
13	4.2	2.6	2.8	3.7	3.6	3.7	2.3	0.6721
14	2.6	2.7	2.4	1.7	1.8	2.4	2.7	0.6489
15	2.7	2.4	2.2	2.6	2.7	3.6	3.5	0.6849

Relying on the fuzzy model, the rail health condition is estimated. Each segment is evaluated based on the health condition as shown in Figure 4.15 and Table 4.2. Table 4.2 presents the results of the case study. Given the influential factors, the rail health condition based on the fuzzy inference system is estimated. Although some rules might not be needed as they might never apply in practice,

we aimed to develop a questionnaire that captures all the possibilities to have a full coverage of inputs. Using the proposed inference system, any rail segment can be evaluated with given influential factors. Table 4.2 gives an example on how the fuzzy inference system performs. The influential factors are obtained from rail field measurements and the last column is calculated using the fuzzy rules.

The figure shows the number of squats over the full track from Amersfoort to Weert. The defects are detected based on the proposed detection model described in the Section 4.2. The segments with the most severe squats are distinguished by two different arrows in Figure 4.16. As seen in the figure, the segments 7 and 10 include the highest number in squat numbers. Thus, the segments 7, 8, 9 and 10 are selected as the critical segments to be maintained. Depending on available maintenance time slot, the track can be ground. If after grinding the above-mentioned segments, there is time to maintenance the rest of the rail, the segment 13 and 15 are candidates to be maintained (marked by black arrows). From the figure, j_{13} has more squats than j_{15}, but according to Figure 4.15, j_{13} is more critical in terms of the health condition and also is shorter in length, which increases the squats density distributed on the track. Therefore, segment 15 is chosen as the alternative option.

Once the critical segments are determined, the optimal clustering model is used to cover squats subject to the time limit imposed by the maintenance slot. The proposed clustering model is able to treat the most important squats. In this chapter, the length of the maintenance time slot is set to 8 hours. This time includes the setup time which covers the cost associated with transportation, machinery, personnel, etc.

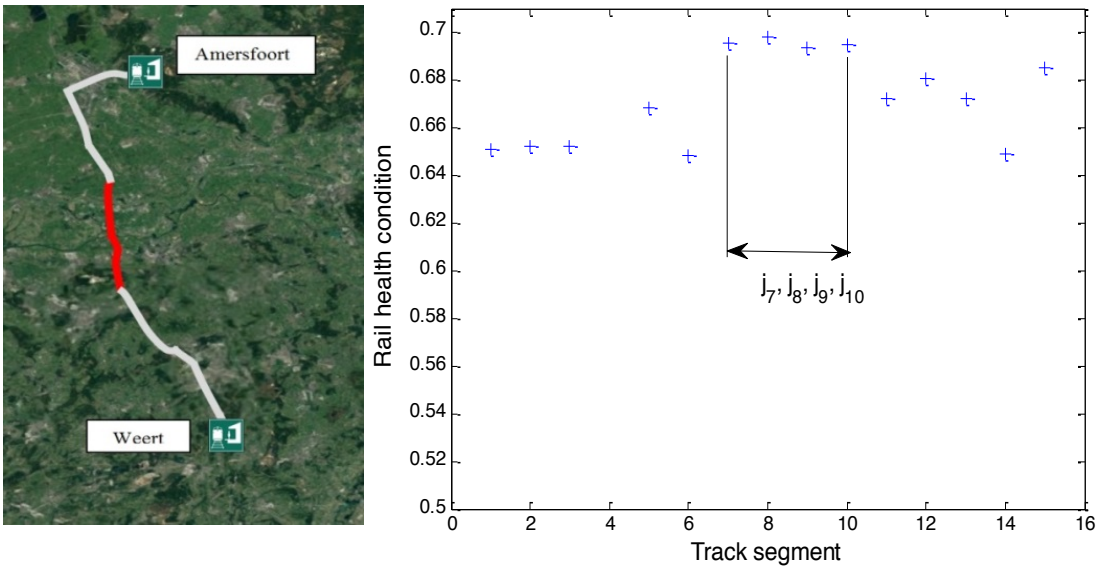


Figure 4.15 Rail health condition over the track segment showing the most critical pieces of the track highlighted with red line on the map

The most relevant squats are covered by a cluster, as the clustering model penalizes a squat outside any cluster by its severity. Hence, the most important squats are treated by grinding, even when the maintenance slot is not long enough. This is not normally the case for cyclic grinding in which the grinding machine starts grinding from the start point going towards to the end of the track without any guarantee to capture the most important squats.

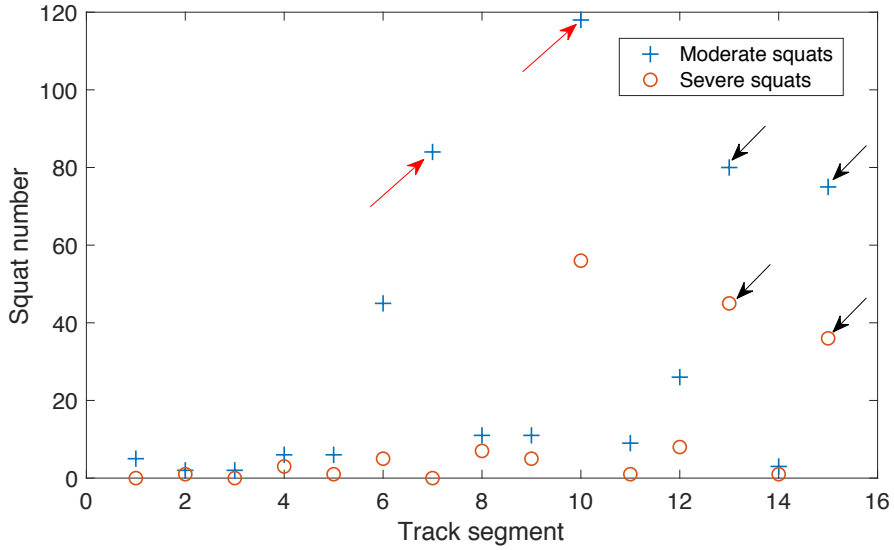


Figure 4.16 Number of moderate and severe squats within different segments. Red arrows show the target segments including most severe squats and black arrows indicate the alternative options for grinding in case the maintenance time slot remains after maintaining the target segments.

Figure 4.17 shows the clustering result between the stations Geldermalsen and 's Hertogenbosch covering the critical segments, i.e. j_7, j_8, j_9, j_{10}. The target track is around 20 km as shown in the x-axis of the figure. According to the proposed detection model, the squat severity is estimated as indicated in the y-axis. The grinding model proposes two clusters within the maintenance time slot capturing the most severe squats by considering density of the squats. In this way, the grinding machine starts grinding from the beginning of the track to reach the kilometer 52.42, then the machine stops working to drive to cluster two (the transfer time is supposed negligible) which starts at the track position 60.51 km until the end of the track. The squats remained not treated, will be maintained in the next maintenance operation, which typically is performed every 6 months in the Dutch railway network. Moreover, the number of the defects between (remaining track piece) 52.46 km and 60.51 km is much less (43 defects and average severity 2.10) than in the first cluster between 46 km until 52.4 km (77 defects and average severity 2.15) and the second cluster between 60.51 km until 66 km (187 defects and average severity 2.25). Thus, although we have defects between 52.46 and 60.51, by considering (1) the maintenance time slot limitation and (2) maintenance priority of the segments 7 and 10 (Figure 4.17) in terms of higher value of health condition, the defects between 52.42 and 60.51 remain with no maintenance intervention until the next maintenance time slot. Without the proposed clustering model, the grinding machine will not be able to capture the most important squats, either at the beginning of the track or at the end of the track. Some severe squats would therefore remain untreated, which would increase maintenance costs and the probability of rail failure.

The cost to employ the grinding machine is 35k euro for one night considering 10 hours. Note that 10 hours is fixed meaning that for shorter maintenance time slots (time<10 hour) the cost is the same. Thus, the infrastructure manager will be charged the same amount of money, although the machine is used for less than 10 hours. Thus, in case there is 2 hour more free-traffic time after 8 hours, then the infrastructure manager has the chance to fill all the available time to keep the grinding machine running. In that case, according to the proposed methodology, the grinding machine can be transferred to the segment 10, j_{10}, which has a more critical health condition compared to the rest of the target track. Then, the infrastructure manager can ensure that traffic-free hours have efficiently used to treat all the most important squats over a long track.

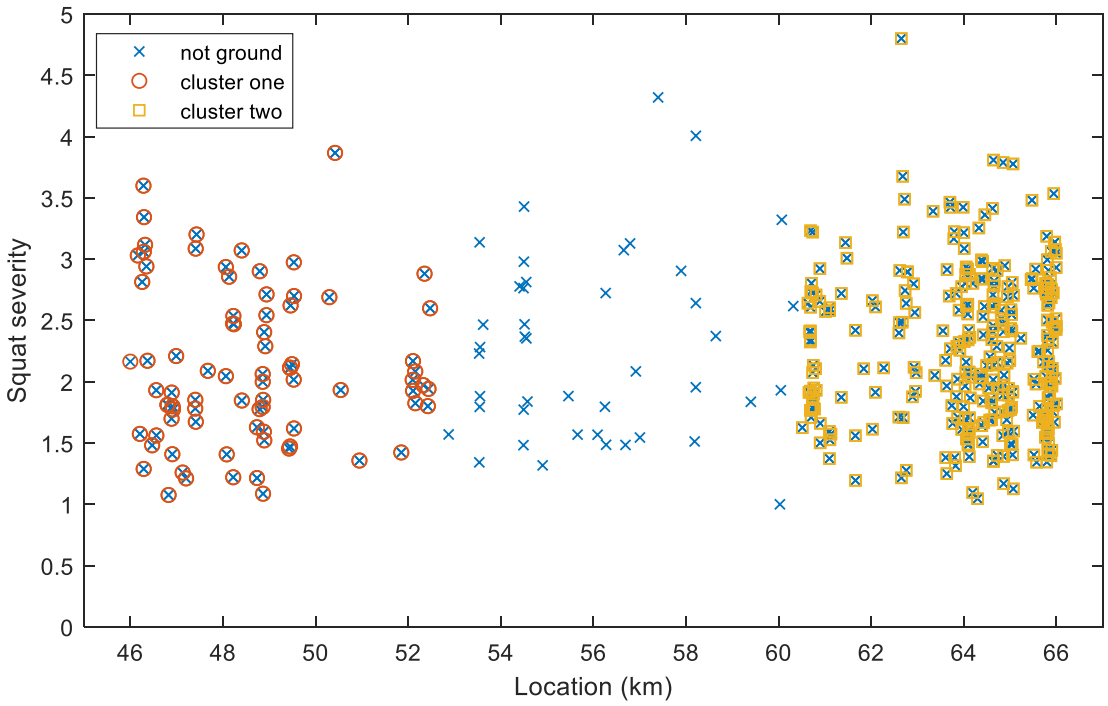


Figure 4.17 Result of the grinding model that determines the optimal clustering of the target track. A squat is marked by either a colored square or a colored circle if they are covered by a cluster.

4.8 Conclusion

In this chapter, we propose an integrated approach for maintenance decision system of the railway infrastructures. The methodology includes infrastructure condition monitoring and maintenance decision making. The proposed approach is applied to the condition-based treatment of squats, with big data information coming from a track in the Dutch railway network. The algorithm makes use of both ABA signals and rail video images, which contribute a huge amount of data. The use of both rail data sources reduces the detection error of the surface defects. Moreover, we have used the track characteristics of the Dutch railway network, enabling the infrastructure manager to interconnect the track influential factors with the actual rail health condition. We therefore investigated how to define a list of decision actions to support the decisions regarding the maintenance plan by analysing the above-mentioned interdependency. The results propose a maintenance decision approach based on the actual condition of the rails but together with the insights resulting from the influential factors. We proposed a partitioning of 15 different segments for a track that can be considered quite long (105 km). In the further research the interdependency analysis can be conducted at a more detailed level, for instance at every kilometre or even at meter of track. In future research, based on the influential factors it will be possible to anticipate much better the rail condition, so a predictive maintenance could be achieved. The maintenance decision system is proposed using a clustering model to perform grinding over the critical pieces of the rail. The results include the most severe squats covered by the maintenance clusters. Thus, although not all the squats are treated, the infrastructure manager can make sure that there is considerably less safety risk or high maintenance cost until the next rail measurement campaign. That is the reason we consider the maintenance time slot as a constraint, to include possible practical limitations as well. Different pass numbers of the grinding machine, resulting in

different grinding depths, have an impact on the rail defect risk after grinding. Different pass numbers also lead to different grinding speeds. The current clustering model considers only one grinding depth, meaning one fixed pass number and grinding speed. In the future work, we will consider flexible pass number of the grinding machine to obtain more efficient clustering plans.

While this chapter is focused on the analysis of squats, the results are applicable to the analysis of other types of rail defects like corrugations, damaged insulated joints, welds and other types of RCF defects. To apply the proposed methodology to all those defects, the infrastructure manager will need to analyse the rail observations in terms of that specific type of defect versus the track characteristics to define the list of decision rules. The methodology for the design of the rules is flexible, so they can be adapted to different railway networks. The maintenance operations could be different from one type of defect to another, but the general methodology can be adapted, as far as the defects can be grouped into different clusters. In addition, the proposed methodology can be linked to a rail maintenance cost analysis to reduce life cycle cost (LCC). Also, by having different measurement sets of rail data, a prediction model of how the defects can grow over time could be added to the methodology, correlated to the influential factors. This will help the infrastructure manager to predict the rail health condition in advance and also to prolong the maintenance decision time horizon. Another topic for further research is to evaluate the methodology at different regions to investigate the influence of exogenous factors like environmental factors to the decision rules and consequently the maintenance decision rules.

Acknowledgement

This research is part of the NWO/ProRail project "Multi-party risk management and key performance indicator design at the whole system level (PYRAMIDS)", project code 438-12-300, and the STW/ProRail project "Advanced monitoring of intelligent rail infrastructure (ADMIRE)", project 12235, which are partly funded by the Ministry of Economic Affairs. The authors also would like to thank INSPECTATION for providing us with image data and technical support.

4.9 References

[1] Åhrén, T. and Parida, A., 2009. Maintenance performance indicators (MPIs) for benchmarking the railway infrastructure: a case study. Benchmarking: An International Journal, 16(2), pp.247-258.

[2] Parida, A. and Chattopadhyay, G., 2007. Development of a multi-criteria hierarchical framework for maintenance performance measurement (MPM). Journal of Quality in maintenance Engineering, 13(3), pp.241-258.

[3] Gandomi, A. and Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), pp.137-144.

[4] Jamshidi, A., Núñez, A., Dollevoet, R. and Li, Z., 2017. Robust and predictive fuzzy key performance indicators for condition-based treatment of squats in railway infrastructures. Journal of Infrastructure Systems, 23(3), p.04017006.

[5] Fumeo, E., Oneto, L. and Anguita, D., 2015. Condition based maintenance in railway transportation systems based on big data streaming analysis. Procedia Computer Science, 53, pp.437-446.

[6] Lasisi, A. and Attoh-Okine, N., 2018. Principal components analysis and track quality index: A machine learning approach. Transportation Research Part C: Emerging Technologies, 91, pp.230-248.

[7] Attoh-Okine, N.O., 2017. Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering. John Wiley & Sons.

[8] He, Q., Kamarianakis, Y., Jintanakul, K. and Wynter, L., 2013. Incident duration prediction with hybrid tree-based quantile regression. In Advances in Dynamic Network Modeling in Complex Transportation Systems (pp. 287-305). Springer, New York, NY.

[9] Liu, X. and Dick, C.T., 2016. Risk-based optimization of rail defect inspection frequency for petroleum crude oil transportation. Transportation Research Record: Journal of the Transportation Research Board, (2545), pp.27-35.

[10] Ghofrani, F., He, Q., Goverde, R.M. and Liu, X., 2018. Recent applications of big data analytics in railway transportation systems: A survey. Transportation Research Part C: Emerging Technologies, 90, pp.226-246.

[11] He, Q., Li, H., Bhattacharjya, D., Parikh, D.P. and Hampapur, A., 2015. Track geometry defect rectification based on track deterioration modelling and derailment risk assessment. Journal of the Operational Research Society, 66(3), pp.392-404.

[12] Zoeteman, A., Dollevoet, R., and Li, Z., 2014. Dutch research results on wheel/rail interface management: 2001–2013 and beyond. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 228(6), 642-651.

[13] Liu, X., Barkan, C. and Saat, M., 2011. Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transportation Research Record: Journal of the Transportation Research Board, (2261), pp.178-185.

[14] Liu, X., Saat, M. and Barkan, C., 2012. Analysis of causes of major train derailment and their effect on accident rates. Transportation Research Record: Journal of the Transportation Research Board, (2289), pp.154-163.

[15] Islam, D.M.Z., Laparidou, K. and Burgess, A., 2016. Cost effective future derailment mitigation techniques for rail freight traffic management in Europe. Transportation Research Part C: Emerging Technologies, 70, pp.185-196.

[16] Makino, T., Neishi, Y., Shiozawa, D., Kikuchi, S., Okada, S., Kajiwara, K. and Nakai, Y., 2016. Effect of defect shape on rolling contact fatigue crack initiation and propagation in high strength steel. International Journal of Fatigue, 92, pp.507-516.

[17] Sciammarella, C.A., Chen, R.J.S., Gallo, P., Berto, F. and Lamberti, L., 2016. Experimental evaluation of rolling contact fatigue in railroad wheels. International Journal of Fatigue, 91, pp.158-170.

[18] Popović, Z., Radović, V., Lazarević, L., Vukadinović, V. and Tepić, G., 2013. Rail inspection of RCF defects. Metalurgija, 52(4), pp.537-540.

[19] Zhuang, L., Wang, L., Zhang, Z. and Tsui, K.L., 2018. Automated vision inspection of rail surface cracks: A double-layer data-driven framework. Transportation Research Part C: Emerging Technologies, 92, pp.258-277.

[20] Makino, T., Kato, T. and Hirakawa, K., 2012. The effect of slip ratio on the rolling contact fatigue property of railway wheel steel. International Journal of Fatigue, 36(1), pp.68-79.

[21] Molodova, M., Li, Z., Núñez, A. and Dollevoet, R., 2014. Automatic detection of squats in railway infrastructure. IEEE Transactions on Intelligent Transportation Systems, 15(5), pp.1980-1990.

[22] Fan, Y., Dixon, S., Edwards, R.S. and Jian, X., 2007. Ultrasonic surface wave propagation and interaction with surface defects on rail track head. Ndt & E International, 40(6), pp.471-477.

[23] Song, Z., Yamada, T., Shitara, H. and Takemura, Y., 2011. Detection of damage and crack in railhead by using eddy current testing. Journal of Electromagnetic Analysis and Applications, 3(12), p.546- 550.

[24] Mariani, S., Nguyen, T., Phillips, R.R., Kijanka, P., Lanza di Scalea, F., Staszewski, W.J., Fateh, M. and Carr, G., 2013. Noncontact ultrasonic guided wave inspection of rails. Structural Health Monitoring, 12(5-6), pp.539-548.

[25] Wen, M., Li, R. and Salling, K.B., 2016. Optimization of preventive condition-based tamping for railway tracks. European Journal of Operational Research, 252(2), pp.455-465.

[26] Budai, G., Huisman, D. and Dekker, R., 2006. Scheduling preventive railway maintenance activities. Journal of the Operational Research Society, 57(9), pp.1035-1044.

[27] Caetano, L.F. and Teixeira, P.F., 2016. Strategic model to optimize railway-track renewal operations at a network level. Journal of Infrastructure Systems, 22(2), p.04016002.

[28] Peng, F. and Ouyang, Y., 2012. Track maintenance production team scheduling in railroad networks. Transportation Research Part B: Methodological, 46(10), pp.1474-1488.

[29] He, Q., Li, H., Bhattacharjya, D., Parikh, D.P. and Hampapur, A., 2015. Track geometry defect rectification based on track deterioration modelling and derailment risk assessment. Journal of the Operational Research Society, 66(3), pp.392-404.
[30] Peng, F. and Ouyang, Y., 2014. Optimal clustering of railroad track maintenance jobs. Computer-Aided Civil and Infrastructure Engineering, 29(4), pp.235-247.
[31] Santos, R. and Teixeira, P.F., 2011. Heuristic analysis of the effective range of a track tamping machine. Journal of Infrastructure Systems, 18(4), pp.314-322.
[32] Li, Z., Molodova, M., Núñez, A. and Dollevoet, R., 2015. Improvements in axle box acceleration measurements for the detection of light squats in railway infrastructure. IEEE Transactions on Industrial Electronics, 62(7), pp.4385-4397.
[33] Hajizadeh, S., Li, Z., Dollevoet, R.P. and Tax, D.M., 2014, August. Evaluating Classification Performance with only Positive and Unlabeled Samples. In Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR) (pp. 233-242). Springer, Berlin, Heidelberg.
[34] Su, Z., Jamshidi, A., Núñez, A., Baldi, S. and De Schutter, B., 2017. Multi-level condition-based maintenance planning for railway infrastructures–A scenario-based chance-constrained approach. Transportation Research Part C: Emerging Technologies, 84, pp.92-123.
[35] Li, Z., Zhao, X., Esveld, C., Dollevoet, R. and Molodova, M., 2008. An investigation into the causes of squats—Correlation analysis and numerical modeling. Wear, 265(9-10), pp.1349-1355.
[36] Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., Li, Z. and Schutter, B., 2017. A big data analysis approach for rail failure risk assessment. Risk analysis, 37(8), pp.1495-1507.
[37] Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R. and De Schutter, B., 2016. Deep convolutional neural networks for detection of rail surface defects. International Joint Conference on Neural Networks (IJCNN), 2584-2589. July, Vancouver, Canada.
[38] Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
[39] LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), p.436.
[40] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), pp.1929-1958.
[41] Wang, L., Wang, P., Quan, S. and Chen, R., 2012. The effect of track alignment irregularity on wheel-rail contact geometry relationship in a turnout zone. Journal of Modern Transportation, 20(3), pp.148-152.
[42] Kawaguchi, A., Miwa, M. and Terada, K., 2005. Actual data analysis of alignment irregularity growth and its prediction model. Quarterly report of RTRI, 46(4), pp.262-268.
[43] Mutton, P.J., Epp, C.J. and Dudek, J., 1991. Rolling contact fatigue in railway wheels under high axle loads. In Mechanics and Fatigue in Wheel/Rail Contact, pp. 139-152.
[44] Grassie, S.L., 2012. Squats and squat-type defects in rails: the understanding to date. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 226(3), pp.235-242.
[45] Li, Z., Zhao, X., Esveld, C., Dollevoet, R. and Molodova, M., 2008. An investigation into the causes of squats—Correlation analysis and numerical modeling. Wear, 265(9-10), pp.1349-1355.
[46] Nielsen, J.C., Ekberg, A. and Lundén, R., 2005. Influence of short-pitch wheel/rail corrugation on rolling contact fatigue of railway wheels. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 219(3), pp.177-187.
[47] Andrade, A.R. and Teixeira, P.F., 2011. Uncertainty in rail-track geometry degradation: Lisbon-Oporto line case study. Journal of transportation engineering, 137(3), pp.193-200.
[48] 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, 36(1), pp.1-8.
[49] Veit, P., 2007. Track quality–luxury or necessity. Railway Technical Review Special: Maintenance and Renewal, pp.8-12.
[50] Sharma, S., Cui, Y., He, Q., Mohammadi, R. and Li, Z., 2018. Data-driven optimization of Railway maintenance for track geometry. Transportation Research Part C: Emerging Technologies, 90, pp.34-58.
[51] Hamid, A. and Gross, A., 1981. Track-quality indices and track degradation models for maintenance-of-way planning. Transportation Research Board, 802, pp.2-8.

[52] Corbin, J.C. and Fazio, A.E., 1981. Performance-based track-quality measures and their application to maintenance-of-way planning. Transportation Research Record, (802), pp.19-27.
[53] Bing, A.J. and Gross, A., 1983. Development of railroad track degradation models. Transportation research record, (939), pp. 27-31.
[54] Grassie, S.L. and Elkins, J.A., 2005. Tractive effort, curving and surface damage of rails: Part 1. Forces exerted on the rails. Wear, 258(7-8), pp.1235-1244.
[55] Li, Z., Dollevoet, R., Molodova, M. and Zhao, X., 2011. Squat growth—Some observations and the validation of numerical predictions. Wear, 271(1-2), pp.148-157.
[56] Vo, K.D., Zhu, H.T., Tieu, A.K. and Kosasih, P.B., 2015. Comparisons of stress, heat and wear generated by AC versus DC locomotives under diverse operational conditions. Wear, 328, pp.186-196.
[57] Scott, D., Fletcher, D.I. and Cardwell, B.J., 2014. Simulation study of thermally initiated rail defects. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 228(2), pp.113-127.
[58] Ma, F., Wu, Q., Yan, X., Chu, X. and Zhang, D., 2015. Classification of automatic radar plotting aid targets based on improved fuzzy C-means. Transportation research part c: emerging technologies, 51, pp.180-195.
[59] Camastra, F., Ciaramella, A., Giovannelli, V., Lener, M., Rastelli, V., Staiano, A., Staiano, G. and Starace, A., 2015. A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference. Expert Systems with Applications, 42(3), pp.1710-1716.
[60] Tosun, M., Dincer, K. and Baskaya, S., 2011. Rule-based Mamdani-type fuzzy modelling of thermal performance of multi-layer precast concrete panels used in residential buildings in Turkey. Expert Systems with Applications, 38(5), pp.5553-5560.
[61] Markowski, A.S. and Mannan, M.S., 2009. Fuzzy logic for piping risk assessment (pfLOPA). Journal of loss prevention in the process industries, 22(6), pp.921-927.
[62] Xie, M., 2003. Fundamentals of robotics: linking perception to action, World Scientific Publishing Company, 54.
[63] Bemporad, A. and Morari, M., 1999. Control of systems integrating logic, dynamics, and constraints. Automatica, 35(3), pp.407-427.

Pareto-based maintenance decisions for regional railways

This chapter corresponds to the reference: *A. Núñez, A. Jamshidi, H. Wang, "Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration". IEEE Transactions on Industrial Informatics. Volume 15, Issue 3, May 2018, Pages: 1496-1507. DOI: 10.1109/TII.2018.2847736.*

5.1 Introduction

Regional railway transport has always been challenging because traditional cost-benefit analyses normally suggest that railway transport is not economically feasible from an operations and maintenance perspective. This lack of economic feasibility can be explained by the relatively low demand and the dispersed nature of the population in rural areas. However, from the social perspective, these regional railway lines are crucial for communities because they

provide access to work and services [1]. Moreover, good access to transportation services can help to prevent the de-population of rural areas and reduce the concentration of capital, services and attractions in big cities, which will reduce urban traffic congestion and increase investment in areas such as agriculture and tourism. Hence, regional railways can be seen as a tool to decentralize opportunities over a territory while accounting for societal equality and to include regionalization in decision-making.

The main challenge faced by regional railways is that their operations and maintenance budget is constantly decreasing. Consequently, new technological solutions that can support the decision-making to spend scarce resources in a smart way are needed. To optimize maintenance costs, many intelligent monitoring systems have recently been used for railway infrastructures, which need big data analyses over the data collected [2], [3]. The range of sensing technologies has expanded rapidly, and sensor devices have simultaneously become cheaper. For example, networking technologies, WSNs, smart phones, accelerometers installed on trains, drones and video cameras have all developed and become less expensive [4], [5]. In this chapter, extending the traditional concept of what a train does is proposed, which is transporting people or goods, to include performing monitoring tasks. In this way, the capabilities of trains can be fully used while providing information about the health conditions of the track. Therefore, axle box acceleration (ABA) measurements are used from a regular train in operation to provide information about the health conditions of rails. This information is useful for detecting shortwave defects and corrugation and for determining the conditions of insulated joints and crossings. The health conditions of rails can be estimated by tracing their degradation over time with the proper indicators [6], [7]. Degradation analyses allow infrastructure managers to be aware of critical locations by providing information about when degradation will reach a critical level, and this information can be used to mitigate the risk of a rail break. Rail degradation is mainly related to rail surface defects [8], and it is a stochastic process that changes over time and space. In this chapter, analyzing weld health conditions based on multiple ABA measurements is focused. The detection and assessment of rail surface defects are normally realized by analyzing frequency-based features. Previously, methods such as the power spectrum density [9], short-time Fourier transform [10] and wavelet transform [11] methods have been adopted for detecting defects in wheels, axle bearings and rails. ABA signals from regional railways are particularly noisy. The reason is that the signals can be affected not only by train speed and wheel conditions but also by a less accurate GPS location and track irregularities. To circumvent the usage of a fixed signal, this chapter employs the Hilbert-Huang Transform (HHT) based on an adaptive signal that can be associated with physical excitations [12]. The Hilbert spectrum of ABA is employed as an indicator of the weld health conditions.

The use of ABA for assessing welds is inherently stochastic, so a set of robust and predictive key performance indicators (KPIs) is defined to capture the weld's degradation dynamics during a given maintenance period. Using a scenario-based approach, two objective functions are used, the performance and the number of weld replacements. The latter function is related to maintenance costs. To facilitate decision-making, a methodology is proposed based on multi-objective optimization. Although the proposed methodology is applied to the rail welds, the structure can be generalized for other types of rail surface defects. Figure 5.1 shows the four major steps of the methodology. First, an operational train is used to collect ABA signal data. Then, the health conditions of the rail are estimated by using the Hilbert spectrum of the ABA signal at the rail welds. Next, KPIs are established to characterize the severity of the detected welds. As the KPIs include the uncertainty induced by measurement-related variables,

such as vehicle speed, condition of wheels, etc., a stochastic analysis is presented in the second step to provide robust KPIs. Moreover, different degradation scenarios are introduced to include the predictability of the KPIs. In the last step, a maintenance decisions optimization problem (MDOP) is solved to obtain Pareto-optimum maintenance solutions.

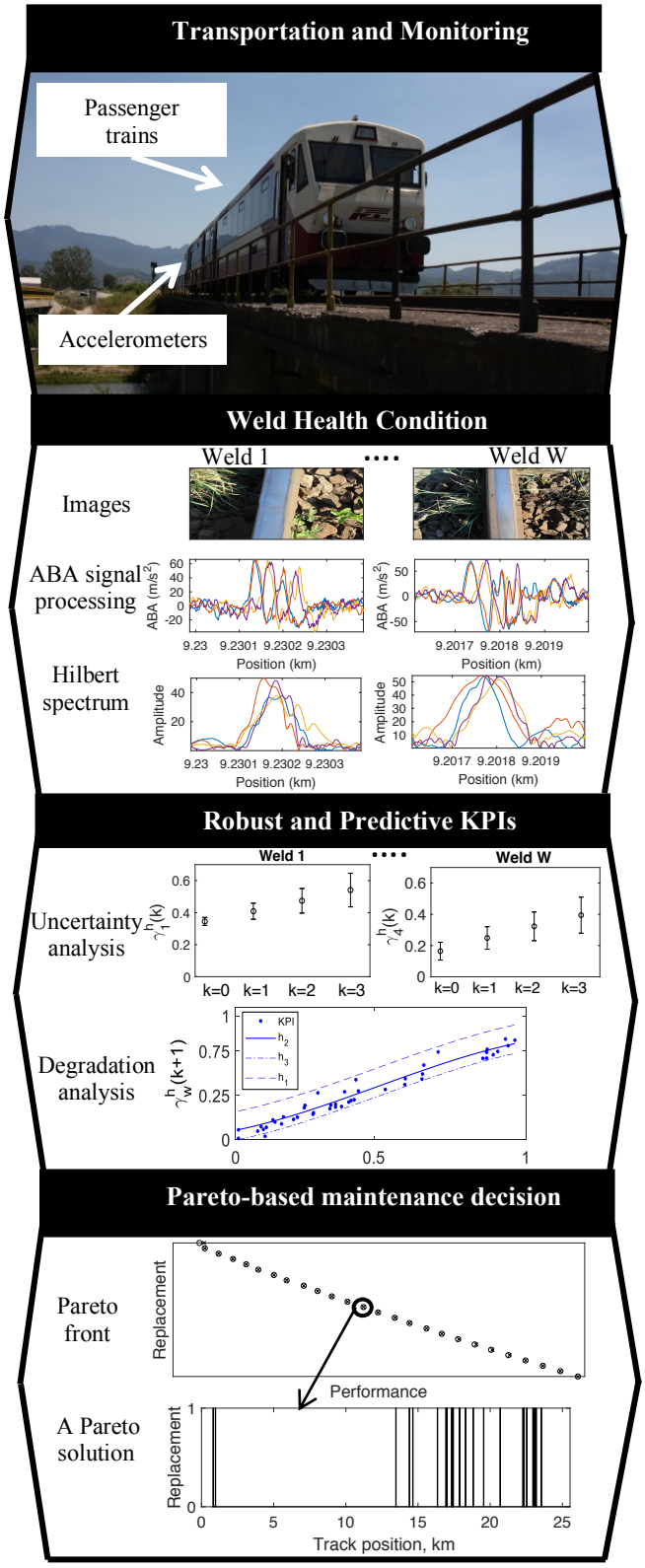


Figure 5.1. A generic flowchart of the proposed methodology

5.2 Methodology

5.2.1 ABA measurement system

There are different methods used to diagnose the condition of rail defects, including ultrasonic measurement [13], eddy current testing [14], and image recognition [15], among other technologies. Each method has different advantages and disadvantages. These methods help to check the crack development underneath rail surface defects (ultrasonic and eddy current) and to visually monitor rails (rail video images). However, these methods are not able to capture the dynamic response of welds under actual operation. By having a dynamic response, valuable information about welds can be acquired. A technology capable of (1) shortwave irregularity detection and (2) the capture of the dynamic response of welds is needed [16]. ABA measurement systems with these capabilities have been reported in different countries: the detection of lateral and vertical track irregularities in Korea [17]; the evaluation of wheel load fluctuation and rolling noise in Japan [18]; the detection of corrugations in Poland [19] and Italy [20]; the detection of rolling contact fatigue defects, damaged welds, and insulated joints in the Netherlands [21], [22]; and the analysis of vertical track geometry in Spain [23]. ABA system implementation has different advantages: (1) ABA is a low-cost measurement system compared to other types of detection methods. (2) The ABA sensors are easy to maintain and (3) ABA can be implemented on in-service operational trains. Moreover, there is (4) the possibility to detect seed rail defects with no need for expensive and complex instrumentation and (5) the ability to estimate the severity level of the dynamic contact force. Additionally, this technology is suitable for regional networks because it is portable and cost-effective. In this chapter, an axle box acceleration (ABA) measurement system is used to detect welds and estimate conditions. In total, 16 accelerometers are mounted on the axle boxes (vertically and longitudinally). For mounting the sensors, small mounting studs have to be glued in dry conditions on each axle box. Furthermore, a GPS antenna for positioning is used and installed on the roof of the train. The positioning system, beside the GPS antenna, uses tacho count which is able to detect pulses at approximately 1 MHz. A schematic view of the ABA system is described in Figure 5.2.

Welds highly excite the vertical acceleration of ABA systems [24]. The location of welds is well known by the infrastructure manager. If not, their locations can easily be obtained by field inspection or video camera systems. However, relating the welds to the ABA signals is not an easy task in the time domain, but it is possible in the frequency domain. To study the relation between the actual welds in the track with the welds detected by the ABA measurement system, a training and validation process is performed. First, a track field inspection is carried out to relate different welds detected by the ABA signals to the actual welds. Based on the examples for training, a validation to other welds is performed by examining similar frequency responses. Finally, predictions are performed for the rest of the track.

5.2.2 Rail welds

Welded rails have considerably improved the problems of rail wear and overloading, which are the main causes of rail breaks. Most field welding in the railway industry is carried out using the aluminothermic technique. Such welds are primarily associated with rail replacement, the installation of insulated rail joints and track construction activities. However, rail welds are typically subject to complex loading and high stresses at the rail head as train wheels pass and they are exposed to cyclic fatigue loading [25]. Thus, the rails with damaged welds produce

vibrations that cause noise and ride discomfort. Over time, they are a safety concern because they contribute to less stable support for moving trains. Crack propagation due to brittle fractures is one of the main factors that lead to a severely defective weld [26]. In regional networks, brittle fractures accelerate the further degradation of defects due to poor maintenance operations.

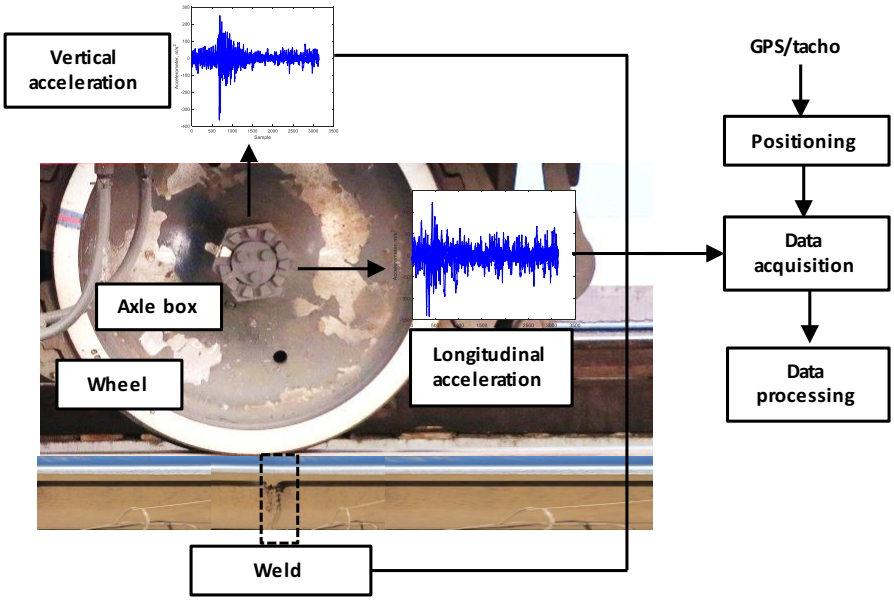


Figure 5.2. A side view of the ABA system

5.2.3 Hilbert spectrum of the ABA signal

The Hilbert spectrum is obtained by combining empirical mode decomposition (EMD) with the Hilbert transform [27], namely, the HHT. EMD is a data-driven algorithm that adaptively decomposes a signal $\{\tau(t), t \in [0,T]\}$ into a number N_{IMF} of intrinsic mode functions (IMFs) $\nu_\ell(t)$, $\ell = 1,2,...,N_{IMF}$ and a residual $r(t)$. The IMFs are selected using a sifting process that only terminates when a stopping criterion has been satisfied. In detail, the EMD algorithm can be described by Algorithm 1. The decomposition starts with the original signal $\tau(t)$. The upper and lower envelopes, $e_\nu(t)$ and $e_l(t)$, of the signal are obtained by connecting all the signal maxima and minima, respectively, using spline interpolation. Then, the average envelope signal $e_m(t)$ can be computed to obtain the estimated IMF $\nu_\ell^{(i)}(t)$ on line 9. At this point, the estimated IMF will be checked by the IMF termination criterion.

$$S(i) = \sum_{t=0}^{T} \frac{|\nu_\ell^{(i)}(t) - \nu_\ell^{(i-1)}(t)|^2}{|\nu_\ell^{(i-1)}(t)|^2} < \varepsilon \tag{5.1}$$

where ε is a positive number that typically ranges from 0.2 to 0.3. If the criterion is not fulfilled, the sifting process continues with the estimated IMF; otherwise, the ℓth IMF is obtained as the estimated IMF, with $\nu_\ell(t) = \nu_\ell^{(i)}(t)$, and the global termination criterion will be checked to see

if all the IMFs are obtained. The global termination criterion judges if the updated estimated IMF $v_j^{(i)}(t)$ on line 13 has only one pair of extrema. If not, the sifting process goes on; otherwise, the signal residual $r(t)$ can be finally obtained.

Algorithm 1: General steps of EMD

Input: $\tau(t)$ (original signal)

Output: $v_\ell(t)$ (IMFs), $r(t)$ (residual)

1 $\ell \leftarrow 1$;
2 $i \leftarrow 1$;
3 $v_\ell^{(i)}(t) \leftarrow \tau(t)$;
4 **while** *global termination criterion not fulfilled* **do**
5 **while** *IMF termination criterion not fulfilled* **do**
6 $[e_v(t), e_l(t)] \leftarrow FindEnvelopes(v_\ell^{(i)}(t))$;
7 $e_m(t) \leftarrow [e_u(t) + e_l(t)]/2$;
8 $i \leftarrow i+1$;
9 $v_\ell^{(i)}(t) \leftarrow v_\ell^{(i-1)}(t) - e_m(t)$;
10 $v_\ell(t) \leftarrow v_\ell^{(i)}(t)$;
11 $\ell \leftarrow \ell+1$;
12 $i \leftarrow 1$;
13 $v_\ell^{(i)}(t) \leftarrow \tau(t) - \sum_{\kappa=1}^{\ell-1} v_\kappa(t)$;
14 $r(t) \leftarrow v_\ell^{(i)}(t)$;
15 **return** $v_\ell(t), r(t)$;

As a result, the sum of all IMFs and the residual can be used to reconstruct the original signal as follows:

$$\tau(t) = \sum_{i=1}^{N_{IMF}} v_\ell(t) + r(t) \tag{5.2}$$

The extracted IMFs are considered as inherent signal modes contained in the original signal that can be associated with a physical meaning or cause. The iterative sifting process automatically decomposes the original signal into IMFs that possess different frequency bands decreasing from $\ell = 1$ to $\ell = N_{IMF}$. Thus, EMD can be applied to ABA signals to extract signal features by dividing the full measurement frequency band into several sub-bands of interest. The next step of HHT is to compute the instantaneous frequency of the extracted IMFs. Concretely, the analytic form of IMFs is first computed using the Hilbert transform:

$$z_\ell(t) = v_\ell(t) + \mathrm{i} \cdot H[v_\ell(t)] = a_\ell(t) \mathrm{e}^{\mathrm{i} \cdot \theta_\ell(t)} \tag{5.3}$$

where $H[d_\ell(t)]$ denotes the Hilbert transform of IMF $\nu_\ell(t)$. The following relationship is established:

$$\begin{cases} a_\ell(t) = \sqrt{\nu_\ell^{\,2}(t) + H[\nu_\ell(t)]^2} \\ \theta_\ell(t) = \arctan\left(\dfrac{H[\nu_\ell(t)]}{\nu_\ell(t)}\right) \end{cases} \tag{5.4}$$

Therefore, the instantaneous frequency $\omega_\ell(t)$ of IMF $\nu_\ell(t)$ can be obtained as

$$\omega_\ell(t) = \frac{\mathrm{d}\theta_\ell(t)}{\mathrm{d}t} \tag{5.5}$$

The instantaneous frequency represents the degree of frequency variation in the IMF $\nu_\ell(t)$ at time t. This feature is similar to the variations of the wavelet coefficient that reflect the signal variations within a certain frequency band. However, in theory, when using the IMFs that are extracted from the signal itself instead of the signals reconstructed from a mother wavelet, the frequency variations are more closely related to the physical nature of the signals, namely, the excitations of ABA at the welds. Finally, the Hilbert spectrum of the signal $\tau(t)$ is defined as a function of instantaneous frequency and time as follows:

$$S(\omega,t) = \mathrm{Re}\left[\sum_{\ell=1}^{N_{IMF}} \tau_\ell(t)\mathrm{e}^{\mathrm{i}\cdot\int\omega_\ell(t)dt}\right] \tag{5.6}$$

where Re denotes the real part of the operator for a complex signal. The Hilbert spectrum of a single IMF $\nu_\ell(t)$ represents the energy variations of the instantaneous frequency with time, and it is given by

$$S_\ell(\omega,t) = \mathrm{Re}\left[a_\ell(t)\mathrm{e}^{\mathrm{i}\cdot\int\omega_\ell(t)dt}\right] \tag{5.7}$$

The Hilbert spectrum can provide indicators for many aspects of application in the sense of a power spectrum. Since the IMFs have different frequency bands that will be inherited by their Hilbert spectra, the employment of Hilbert spectra becomes crucial to determine on which frequency ranges the indicator is based. Thus, a linear combination of the Hilbert spectra of IMFs is proposed as the KPI for rail defect detection as follows:

$$\gamma(t) = \sum_{\ell=1}^{N_{IMF}} c_\ell \, \mathrm{Re}[a_\ell(t) \mathrm{e}^{\mathrm{i}\cdot\int \omega_\ell(t)dt}] \tag{5.8}$$

where c_ℓ ($c_\ell \geq 0$) is the weight coefficient for the spectrum of the ℓ-th IMF with $\sum c_\ell = 1$. For ABA signals measured under different conditions or for the detection of different types of defects with varying frequency responses, the distribution of weight coefficients varies depending on the indicative frequency band.

5.2.4 Description of the KPIs

When performing multiple ABA measurements, the data related to a given weld are subject to a different set of stochasticity. The most influencing uncertainties are related to the way the wheel hits the defect. Sources of stochasticity come from sensors, train speed, wheel quality, the track and the welds. There are various strategies to include uncertainties in the optimization [28]. A scenario-based approach is used to cope with the uncertainties related to the multiple measurements, $m = 1,..,M$, and robust KPIs are defined. Three different severity scenarios are proposed: slow, average and fast severity scenarios. The KPI for the average severity scenario, h2, can be expressed as

$$\gamma_w^{h_2}(k) = \frac{1}{M}\sum_{m=1}^{M} \gamma_w^m(k) \tag{5.9}$$

where $\gamma_w^m(k)$ is the KPI using measurement m at the weld w (located at position x_w). The other two KPIs are defined as

$$\gamma_w^{h_1}(k) = \max\left(\gamma_w^m(k)\right),\ m = 1,2,...,\ M \tag{5.10}$$

$$\gamma_w^{h_3}(k) = \min\left(\gamma_w^m(k)\right),\ m = 1,2,...,\ M \tag{5.11}$$

where $\gamma_w^{h_1}(k)$ and $\gamma_w^{h_3}(k)$ are the fast and slow severity scenarios for the weld w, respectively.

5.2.5 Degradation model

The generic degradation process of a rail infrastructure component is stochastic. The degradation can be stopped by applying maintenance actions. In this chapter, we consider only replacement. The maintenance actions per weld are binary: the weld is replaced or not replaced. For a set of welds, the optimization variable at time step k is $u(k) = [u_1(k),...,u_W(k)]^T \in U$, where W is the number of welds. The following model is proposed to describe the effects of the maintenance actions at a weld located at *x:*

$$\gamma_w^h(k+1) = F_S\left(\gamma_w^h(k), u_w(k)\right) = \begin{cases} F_S^0\left(\gamma_w^h(k)\right) & \text{if } u_w(k)=0 \quad \text{no maintenance} \\ F_S^1\left(\gamma_w^h(k)\right) & \text{if } u_w(k)=1 \quad \text{replacement} \end{cases} \tag{5.12}$$

For other types of defects, such as squats or corrugation, (5.11) should be extended to include grinding or other maintenance actions. To quantify the effects of maintenance on the degradation over time, predictions over the complete maintenance horizon should be conducted [29] for the three degradation scenarios, namely, $h = h_1, h_2, h_3$, by solving (5.12) recursively

$$\gamma_w^h(k+t) = F_S\left(\gamma_w^h(k+t-1), u_w(k+t-1)\right), t = 1,...,N_p \tag{5.13}$$

The predicted health condition of the weld *w* could be considered as an interval prediction, $[\gamma_w^{h_3}(k+t), \gamma_w^{h_1}(k+t)]$, and is obtained from the KPIs measured from the ABA signal.

5.2.6 Optimization of rail maintenance decisions

Because the objective functions of maintenance operations are usually conflicting, i.e., a solution that optimizes one objective may not optimize others, the use of multi-objective optimization is considered. When there are multiple objectives that vary over time, optimization by tuning the weights [30] will work. A maintenance decisions optimization problem (MDOP) is formalized so that the trade-off between Pareto optimal solutions is acquired for the infrastructure manager. The proposed MDOP can be expressed as

$$\min\left\{J_1\left(u(k)\right),...,J_G\left(u(k)\right)\right\} \text{ subject to } u(k) \in U \tag{5.14}$$

Let us consider minimization of the $G \geq 2$ objective functions simultaneously. The solution to 5.14 is known as the Pareto optimal set. A solution $u_p(k)$ is said to be Pareto optimal if and only if another solution $u(k)$ does not exist, such that

$$J_g\left(u(k)\right) \leq J_g\left(u_p(k)\right) \tag{5.15}$$

$$J_g\left(u(k)\right) < J_g\left(u_p(k)\right) \text{ for at least one } g \in \{1,...,G\} \tag{5.16}$$

The set of all objective function values corresponding to the Pareto optimal solutions is known as the Pareto front. In this chapter, two objective functions are considered, the performance and cost functions. For the first objective function, the contribution of the weld *w* to the degradation scenario *h* can be expressed as

$$J_w^h(k) = \sum_{t=1}^{N_p} \alpha_t . \gamma_w^h(k+t) \tag{5.17}$$

where $\gamma_w^h(k)$ is the KPI at the time step k of weld w and under severity scenario h and α_t is a weight applied across the time predictions of the KPIs. According to the predictive and robust KPIs provided with the ABA signals, the objective function J_1 is defined to cover all the welds and all the scenarios over the maintenance period N_p:

$$J_1\left(u\left(k\right)\right)=\frac{\sum_{h\in\{h_1,h_2,h_3\}}\sum_{w=1}^{W}\alpha_h J_w^h\left(k\right)}{\sum_{h\in\{h_1,h_2,h_3\}}\sum_{t=1}^{N_p}\alpha_h\alpha_t} \tag{5.18}$$

where J_1 is the KPI at the time step k and α_h is the weight per scenario. The second function (J_2) is the number of replacements and is directly related to the budget availability and time constraint determined based on the decisions of the infrastructure manager.

$$J_2\left(u\left(k\right)\right)=\sum_{w=1}^{W}\sum_{t=1}^{N_p}\left(1-u_w\left(k+t\right)\right) \tag{5.19}$$

Furthermore, we use a decision window (D_w) based on the given rail replacement length. If more than one weld can be covered by one replacement operation, it will be grouped in advance into a cluster W_i. The replacement operation will then improve the conditions of all the welds in the cluster with only one replacement operation. In this case, the following rule is applied

$$\begin{aligned}\text{if } u_w(k)=1,\ w\in W_i \Rightarrow u_{\varpi}(k)=0,\ \ \forall\,\varpi\in W_i\setminus\{w\}\wedge\\ \gamma_{\varpi}(k+1)=F_s(\gamma_{\varpi}(k),1),\ \ \forall\,\varpi\in W_i\end{aligned} \tag{5.20}$$

where W_i is the i-th cluster of welds. Algorithm 2 shows the general architecture of an MDOP. To solve the MDOP problem, six optimization algorithms are selected and described in Section 5.3 through two case studies.

Algorithm 2: general architecture of an MDOP

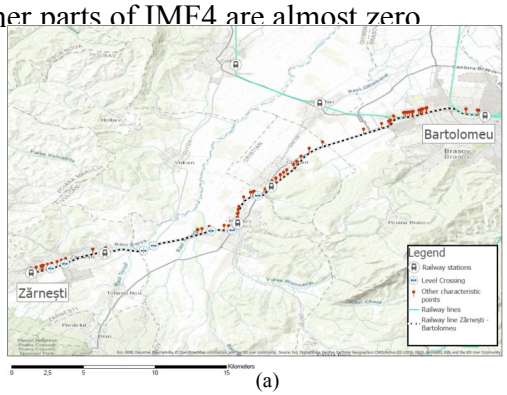
Input: *N (population size), NR (number of reference points and archive size)*
Output: *P (final population)*

```
P ← Random Initialize (N);
R ← Uniform Reference Point(NR);
A ← P;
R′ ← R;
while termination criterion not fulfilled do
    P′ ← Mating Selection (P,R');
    O ← Variation (P',N);
    [A,R′] ← Ref Point Adaption;
    P ← Environmental Selection;
return P;
```

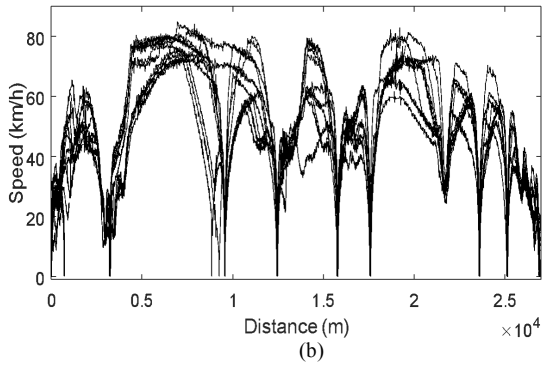
5.3 Numerical results and discussion

In this section, a real-life case study is provided to show the capability of the proposed methodology. The regional track Bartolomeu-Zărneşti in Romania is used as the case study (Figure 5.3). A compactRIO measurement computer is used to collect 16 ABA channels (vertical and longitudinal) and store in the FPGA (field programmable gate array) at 25600 Hz. The tacho and GPS data are also stored in the FPGA. Next, a data block composed of 200 samples is transferred to a computer. The ABA data processing is later performed offline after collecting all the data from the measurement campaign. The measurement data size per kilometer depends on the train speed. The sampling rate is a fixed 25600 Hz. The average data size for the whole track (which covers 26.9525 km) is 273.44 MB/km for one measurement run. For a measurement campaign, ten different measurement runs are normally carried out over the track to assure the repeatability and reduce signal noise. The size of the full data is approximately 73698.916 MB. The weather condition can differently influence rail conditions. High temperatures can increase the risk of incidents related to rail buckles. An excessive amount of snow and ice can cause huge problems for network performance: (1) a frozen catenary might damage the overhead wire and block the train from receiving electrical energy, or (2) a track switch could become dysfunctional when the temperature drop is extreme. Moreover, the welds can be ruptured in a rapid variation of temperature due to thermal expansion and contraction. Although it is an important topic, to the best knowledge of the authors, there is no study on how extreme weather conditions influence the quality of ABA measurement.

In the case study, the ABA system was installed on a passenger train. The speed was varied from 0 km/h at stations, up to 80 km/h. The signals collected at nearly 0 km/h do not contain the necessary excitation for analysis of welds. Signals approximately 70 km/h (higher than 60 km/h) have been selected for the processing. The coverage in the track of the case study is approximately 80% of the infrastructure. For the rail pieces at the train speed below 70 km/h (most of them at stations or near them), quantitative relationships with the signature tunes and maximum ABA should be incorporated using a regression model to make full use of the data collected. Figure 5.4 shows the IMFs decomposed from a piece of the track measured when the train is passing over a weld. The weld located at around the midpoint of the signal segment results in two peaks in ABA. The IMFs numbered from IMF1 to IMF9 show the modes of ABA in different frequency bands in a descending order. The mode that reflects the excitation of the weld can be identified as IMF4, where the two peaks attributed to the weld are adequately preserved, while the other parts of IMF4 are almost zero.



(a)



(b)

Figure 5.3 (a) A map of the railway track with a legend, and (b) the speed profile of the track obtained during various measurement rounds from Brasov to Zărneşti.

Thus, this mode can be considered as the result of EMD separating the vibration response of the weld from other sources of excitation. In this chapter, the frequency band of the IMF4 of the ABA signal is employed as the best representation of the frequency response for welds. As given in (5.8), the weight coefficients c_ℓ are thus determined to be $c_4 = 1$ and $c_{\{\ell:\ell\neq 4\}} = 0$. With this method, all the weld-like impacts will be detected, which are interesting locations for monitoring and rail replacement. Next, the amplitude of the Hilbert spectrum of the decomposed ABA is used to describe the severity of the welds. To outline the next steps of the case study according to the methodology, a four-welds-sample case (Case study one) is provided to describe the details of the methodology. Next, the whole track of the case study is aimed at showing the applicability of the methodology in practice (Case study two). The description of the case studies is given as follows.

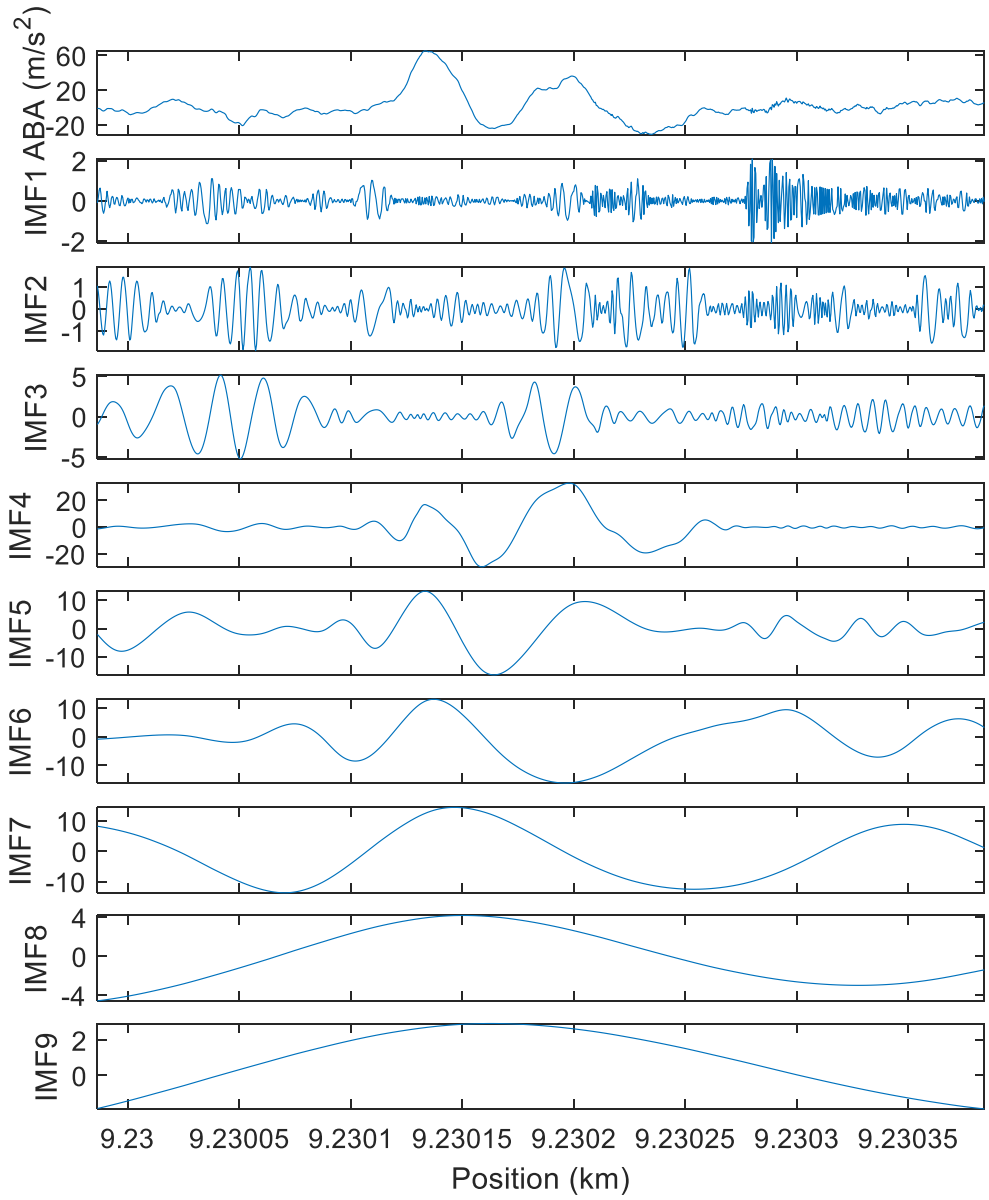


Figure 5.4 A segment of the ABA signal and its IMFs decomposed by EMD.

5.3.1 Case study one

Four welds were selected from the full track of the case study first to give a detailed description of the steps. In Figure 5.5, the ABA signals of four welds are analyzed. As seen in Figure 5.5, four measurements are used to show the impact of the welds on the ABA signals associated with their corresponding IMF4, and the amplitudes of the Hilbert spectrum are presented. All the measurements are obtained from the vertical acceleration signal of the left rail. Moreover, the accuracy of the welds' positioning is weak in the regional network due to the geographic locations. In the case study of this chapter, by defining a set of reference points in the ABA signals, together with tacho and GPS, the positioning of the welds is obtained with errors between 1 to 10 meters. In the welds of Figure 5.5, the data positioning accuracy is approximately 1 meter when compared to ABA peaks due to welds in different measurements.

The obtained amplitudes are used to indicate the severity of the weld damage. In Figure 5.6 (a), the KPIs are estimated, and it is possible to see the degradation behavior of the welds. Relying on the physical understanding of how a surface rail defect grows, a polynomial regression model is fitted using the least-absolute residual method to represent the stochasticity of the growth [31]. Figure 5.6 (b) validates the maintenance decision results. As seen in the ranking, Weld 2 has the highest values in all the scenarios out of the four welds, and Weld 1 has the second highest rank. However, although Weld 3 has higher values than does Weld 4 in both fast and slow scenarios (h1 and h3), the average scenario of Weld 4 shows more criticality in comparison with Weld 3 (higher value). This plot can be used later to determine the extent that the optimization results are in agreement with the expected rankings of the welds. In Case study one, four replacements are possible. In general, the number of replacements should be determined according to the decision of the infrastructure manager. The decision can be explained in terms of (1) budget limit and (2) maintenance time slots. In Case study one, the decision window Dw is shorter than the distance between the welds, so only one weld is replaced by one replacement operation (no clustering needed).

Figure 5.7 shows the Pareto front for four different MDOPs: 1) h1 is a fast severity scenario, 2) h2 is an average severity scenario, 3) h3 is a slow severity scenario. Additionally, αk is an exponentially growing parameter for the global KPI, which is the time indication in the prediction model. In addition, αh is 0.8, 0.6 and 0.4 for the fast, average and slow severity scenarios, respectively. The black arrow in Figure 5.7 when J2=2 (two replacements allowed) shows that once the severity scenario changes from h1 (fast degradation scenario) to h3 (slow degradation scenario), the first objective function (J1) gets a lower value because the degradation predictions are more “optimistic”. The selection of the degradation scenarios will influence the decision-making. A simple calculation shows an average improvement of 49.33% in the rail health condition when two replacements are made rather than zero. Table 5.1 presents the optimal decisions per scenario. The maintenance decision was equal to 1 when a replacement was made (u=1) and 0 when no replacement was made (u=0). From the Table 1, Weld 1 and Weld 2 are the most important welds and are the most common candidates for replacement.

This decision is in agreement with the results shown in Figure 5.5, where the peaks of ABA and the amplitude of the Hilbert spectrum are much higher for these two welds. The maintenance decisions according to the global KPI are also analyzed. Weld 2 is selected when one replacement is allowed ($J_1=3.542, J_2=1$). For two replacements ($J_1=1.793, J_2=2$), Weld 1 and Weld 2 are selected. For three replacements ($J_1=0.8832, J_2=3$), Weld 3 is included in addition to Weld 1 and Weld 2 in slow and fast scenarios, while in an average scenario, Weld 4 is added to the maintenance decision. This finding is in agreement with Figure 5.6 (b). The analysis of the global KPI highlights the influences of Weld 1 and Weld 2 on the replacement decisions.

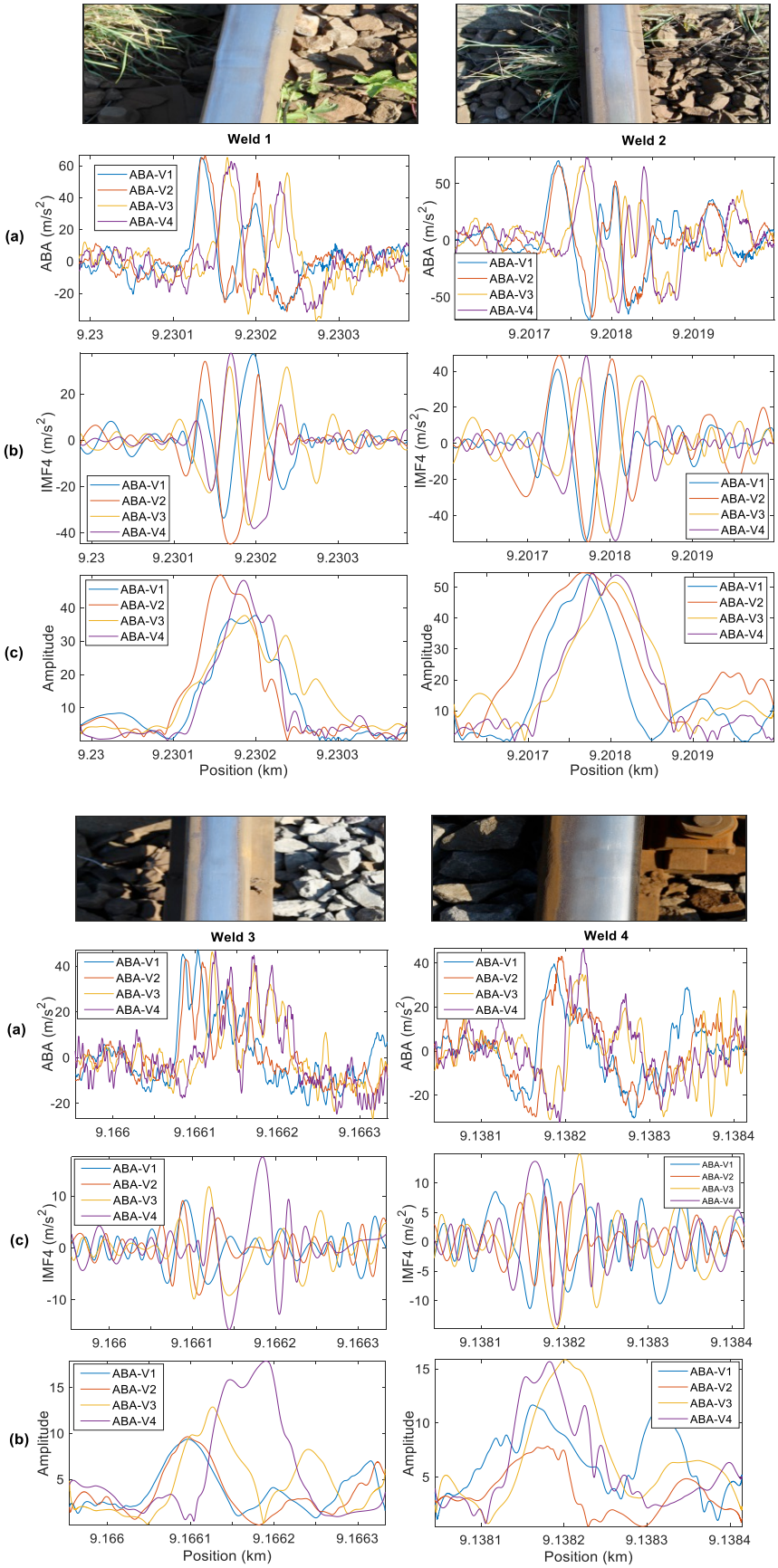
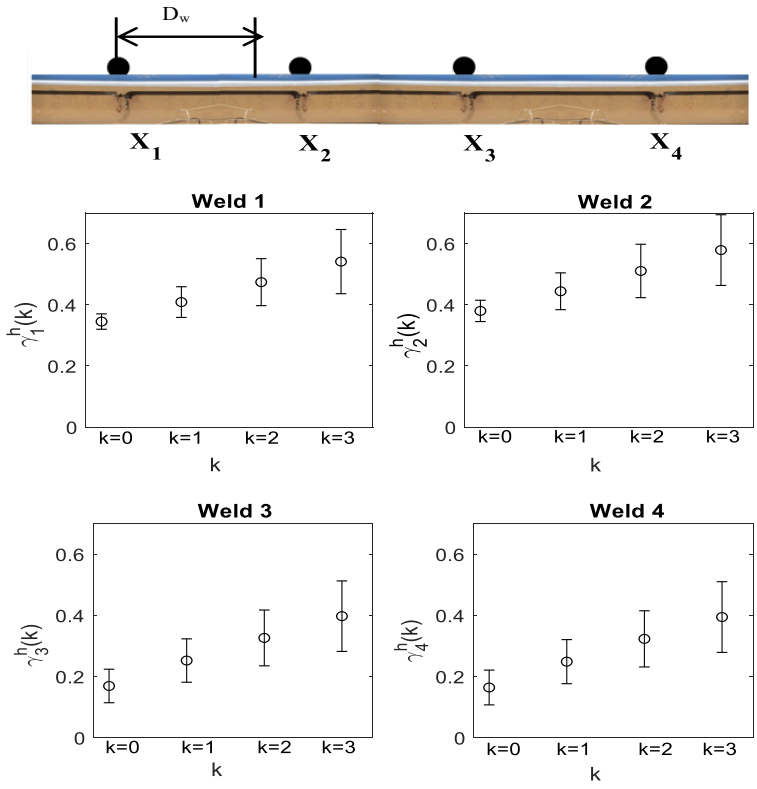
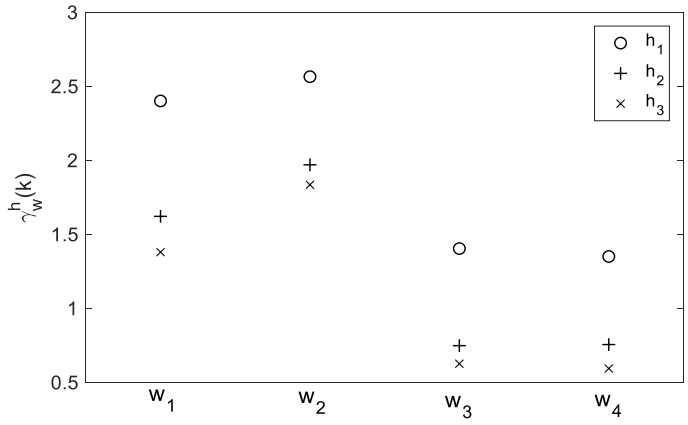


Figure 5.5 (a) Photos of welds for four different measurements with (a) the corresponding ABA signals in the vertical direction, (b) the IMF4 decomposed from the ABA signal, and (c) the Hilbert spectrum amplitude of IMF4.



(a)



(b)

Figure 5.6. (a) Four examples of the KPIs based on four-time-step weld measurements. A schematic view of four welds positioned on a piece of rail. Dw indicates the decision window in meters. (b) J_1 comparison among four welds.

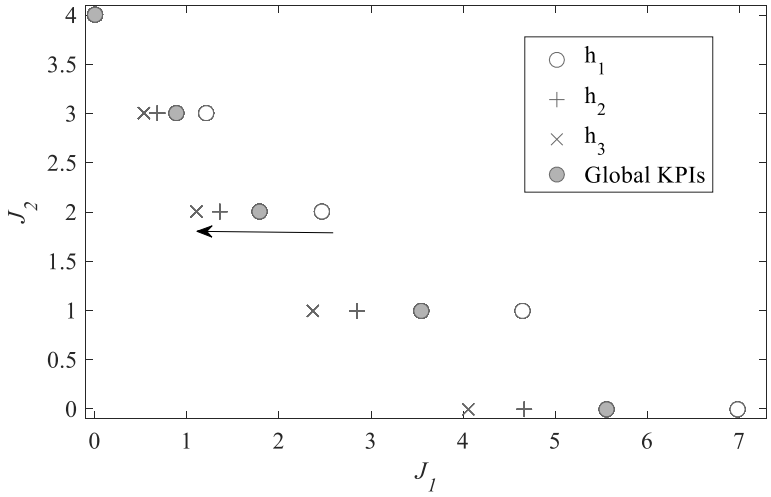


Figure 5.7. The Pareto fronts of the four-welds example. The black arrow depicts the decrease of J_1 as the severity of the scenario changes. The global KPIs result from the methodology considering all scenarios.

Table 5.1 Maintenance decision results for each degradation scenario

J_2	H	J_1	Maintenance decisions			
0	h_1, ○	6.981	0	0	0	0
	h_2, +	4.665	0	0	0	0
	h_3, ×	4.061	0	0	0	0
1	h_1, ○	4.647	0	1	0	0
	h_2, +	2.851	0	1	0	0
	h_3, ×	2.369	0	1	0	0
2	h_1, ○	2.465	1	1	0	0
	h_2, +	1.360	1	1	0	0
	h_3, ×	1.099	1	1	0	0
3	h_1, ○	1.208	1	1	1	0
	h_2, +	0.6764	1	1	0	1
	h_3, ×	0.5339	1	1	1	0
4	h_1, ○	0	1	1	1	1
	h_2, +	0	1	1	1	1
	h_3, ×	0	1	1	1	1

5.3.2 The whole track study

To use the proposed methodology to study the full track from Bartolomeu to Zărneşti, the defect detection algorithm is run for the full track. Over the track, 2128 welds are detected in total. Considering only the welds detected at speeds above 60 km/h, the number of welds needed for optimization is 1849. The detection of welds is presented in Figure 5.8 (a), which depicts their KPIs in terms of the track position in kilometers for a single measurement. Empty severity signals imply that the welds in those locations are not analyzed. In Figure 8 (b), a close view of the track between 1.8 km and 2.4 km is shown. In Figure 5.8 (b), the effects of the decision window *Dw* can be observed. All the damaged welds within the decision window are grouped to indicate the cluster severity. The cluster severity is the summation of the weld severities. In this chapter, a decision window of 15 m is selected, but this window can be modified by the

infrastructure manager if a reasonable alternative is found. Then, the MDOP problem is solved and used by the infrastructure manager to make optimal maintenance decisions. A performance analysis is done to evaluate the reliability of the solutions. Six multi-objective optimization algorithms are compared (1) ARMOEA [32]; (2) NSGA-II [33]; (3) SPEA2 [34]; (4) GrEA [35]; (5) RSEA [36] and VaEA [37]. The reason that those six algorithms are selected is due to their popularity to be used widely in different disciplines. Moreover, some of these algorithms, i.e. GrEA [35] and ARMOEA [32] are of researchers' interests to be employed for multi-objective problems.

Although different MOEAs have been verified on different types of benchmark MOPs, some recent studies have noted that the performance of an MOEA can strongly depend on Pareto front shape of the problems [38]. Hence, some MOEAs are more capable of dealing with regular Pareto fronts, whereas others are specifically tailored for problems with irregular Pareto fronts. For the maintenance optimization problem in this chapter, a binary-based codification is used, i.e., on/off maintenance plan (replacing/not replacing). Considering the nature of a stochastic integer optimization problem, the results of the convergence and diversity of the Pareto fronts are obtained. The algorithm RSEA functions better than do the other algorithms according to Table 5.2. All six algorithms are run using the default setting and 30 independent runs. Populations of 500 solutions and 50000 generations were considered. The maximum number of rail replacements is set to 25 for demonstrations (a reasonable number that should fit the capacities of the replacement operations and budget). The results shown in Figure 5.9 are achieved using the platform PlatEMO [39] and MATLAB 2017b at a desktop computer (2.60 GHz Intel Core i12, 32 GB of RAM). The Pareto fronts are shown in Figure 5.9, and the true Pareto front was obtained from multiple runs over all the algorithms. As shown in Figure 5.9, the RSEA and SPEA2 algorithms approximate the true Pareto front better than the other algorithms. The RSEA converged to the true Pareto front more quickly and with less iteration. To evaluate these algorithms, the metrics that capture the convergence and diversity of the Pareto front approximation delivered by various algorithms can be analyzed. In this chapter, the Normalized Hyper-volume (NHV) is included for performance comparisons. This metric is the only single set quality measure that is known to be strictly monotonic with regard to Pareto dominance; whenever a Pareto set approximation entirely dominates another one, the indicator value of the former approximation will be better. In addition, another metric is considered to include the effects of diversity analysis on the performances of the peer algorithms, namely, the spacing (Sp), which is widely accepted in the literature. The mean and standard deviation (Std.) listed in Table II are statistical results of the same algorithm that was run independently 30 times for the same test problem. The computation time is also calculated for each algorithm. According to the results, although AEROMA gives the fastest runtime, RSEA performs better in terms of convergence and diversity analysis. Figure 5.10 shows the Pareto optimal solution for the maximum number of replacements (25 in total). Figure 5.10 (e) depicts the optimal decision derived by the RSEA. Replacing pieces of the track appears to be critical based on the RSEA (similar to in 10.63 km), but this replacement decision is missing when the results of other algorithms are considered. Moreover, some common rail pieces were found that needed replaced, such as the track pieces between 16.92 km to 18.01 km. The evolution of track deterioration compared to urban railway networks is lower because the railway traffic in regional railways is considerably lower. In the case study line, maintenance plans are organized regularly. A total of 25 locations for the rail replacements are considered to rank as the most important locations. To visualize the maintenance solution obtained by using the RSEA, a map in Figure 5.11 is provided, indicating the 25 locations that are candidates for replacement. Moreover, a zoom-in plot is included to show a piece of the track. The track piece shows seven maintenance solutions labeled with the replacement ranking.

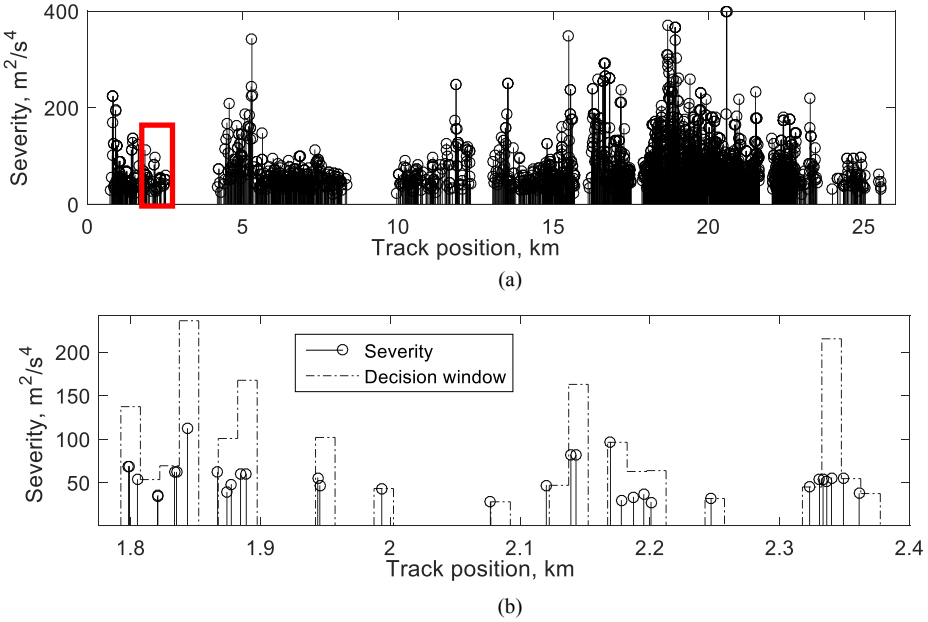


Figure 5.8. (a) Detection results for the track between Bartolomeu and Zărneşti. (b) A piece of the track is highlighted in red to show how a decision window (D_w) is defined over the track.

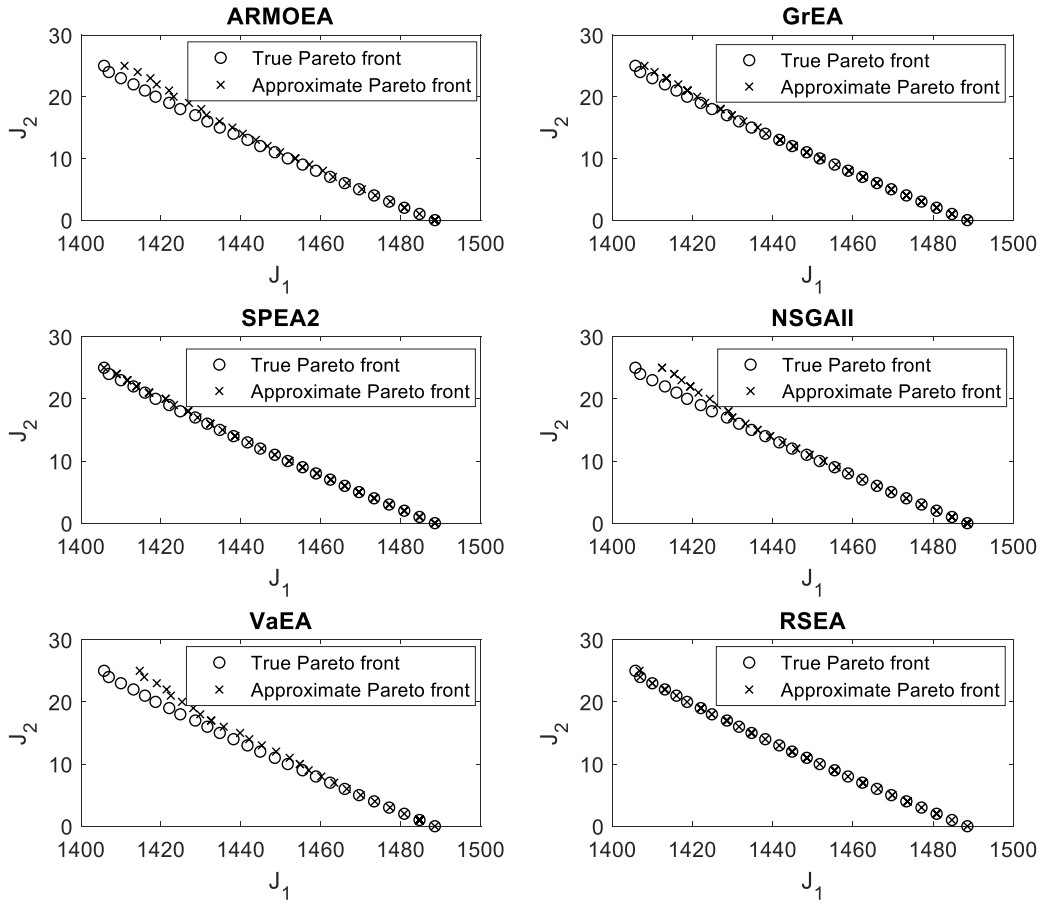


Figure. 5.9 Pareto front results of the three multi-objective optimization methods used to optimize rail maintenance between Bartolomeu and Zărneşti. The black circles show the true Pareto front, while the black cross shows the approximate Pareto front.

Table 5.2 Experimental results of a comparison among the 6 peer algorithms

		NHV	Spacing
ARMOEA	mean	0.87151	0.52538
	Std.	0.00018	0.04732
SPEA2	mean	0.87197	1.05052
	Std.	0.00016	0.15213
NSGAII	mean	0.87187	0.85797
	Std.	0.00016	0.18143
GrEA	mean	0.87203	0.99232
	Std.	0.00014	0.18264
RSEA	mean	0.87229	1.47572
	Std.	0.00012	0.12381
VaEA	mean	0.87127	0.15339
	Std.	0.00023	0.10437

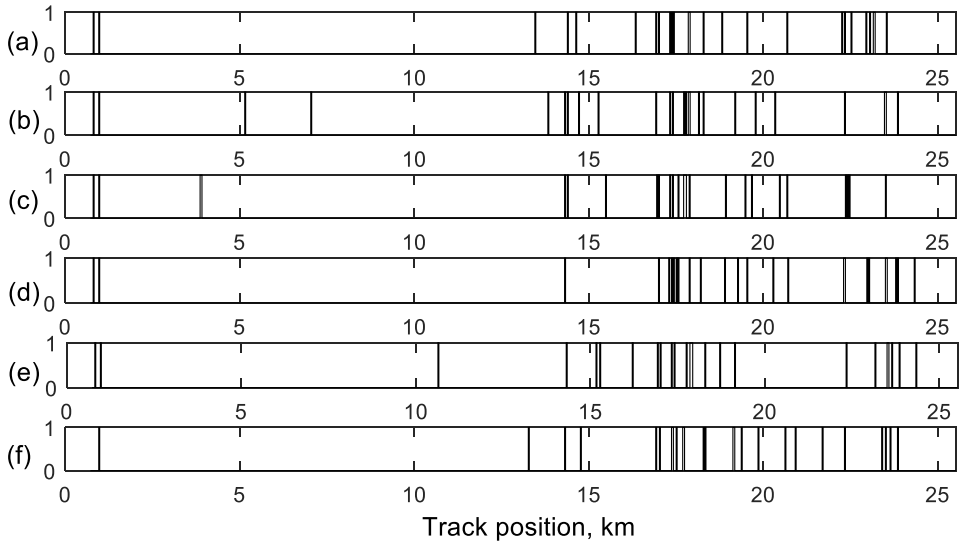


Figure 5.10. Optimal Pareto solutions for the rail maintenance decisions. The x-axis depicts the track distance (km) versus the maintenance decision, where 1 represents replacement and 0 represents no replacement. The following algorithms were used: (a) ARMOEA, (b) NSGAII, (c) GrEA, (d) SPEA2, (e) RSEA, and (f) VaEA.

For example, the location labeled '9' has priority for replacement compared to the other locations with higher numbers. Table 5.3 is presented to show the computational time of (1) data acquisition, (2) the detection algorithm, and (3) the optimization algorithms. For the data acquisition, the data block obtained from the accelerometers is transferred into raw files within 30 seconds. Finally, when a new raw file is created, the old file is transformed into a text file in 10 seconds, and both the raw file and the text file can be stored on a hard drive. Regarding the detection algorithm, the HHT-based detection algorithm requires an average of 2861.630215 seconds to analyze the data measured from the full track, 26.9525 km. Thus, the time needed is approximately 106.17 s/km. Moreover, six algorithms used for the optimization are listed in the table as well. As the processing times for these algorithms are not considerably high, the

computation time difference between algorithms is not of major concern to the infrastructure manager. Furthermore, simplicity and quickness are major indicators of the optimization algorithms to be used in practice. Relying on Table 5.3, the ARMOEA algorithm is fastest compared to the others. However, when considering the simplicity plus the performance of the algorithms, RSEA and SPEA2 are interesting to use in practice as well.

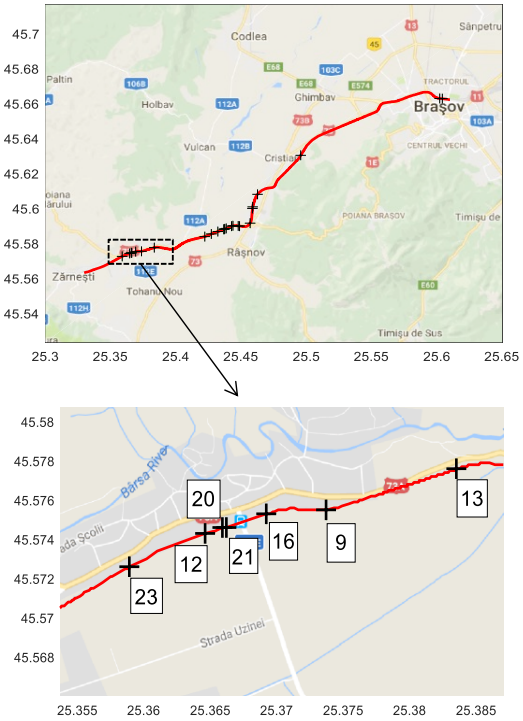


Figure 5.11. Map of the maintenance solutions depicted with "+". A zoom-in plot is attached to show the maintenance priorities with numbers.

Table 5.3 The list of computational times

Transferring time from measurement train to real time computer	30 s/data block		
The HHT-based detection algorithm	106.17 s/km		
The 6 peer algorithms	ARMOEA	mean	24.028 s
		Std.	0.083 s
	SPEA2	mean	30.071 s
		Std.	0.157 s
	NSGAII	mean	25.722 s
		Std.	0.0988 s
	GrEA	mean	24.910 s
		Std.	0.114 s
	RSEA	mean	26.187 s
		Std.	0.090 s
	VaEA	mean	25.446 s
		Std.	0.091 s

5.4 Conclusions

In this chapter, a Pareto-based maintenance decision system is proposed for rail welds. The proposed approach is applied for condition-based maintenance of rail welds based on information from a track in a regional railway network. The Hilbert spectrum approach is used to detect welds based on the ABA. The case study was conducted for track between Bartolomeu and Zărneşti in Romania. To obtain the rail performance, a global KPI was proposed to indicate the conditions of the rail by including degradation scenarios and predictions to provide a clear picture of the network conditions during the maintenance period. The rail replacement number is simply considered as the second objective function, which is related to the budget of the infrastructure manager and the capacity to perform replacement operations during the maintenance period. Six algorithms were run using the same setup to optimize the rail maintenance decisions and evaluate the optimization results. Thus, the infrastructure manager not only receives Pareto optimal solutions but can also compare the results between the different algorithms. Among all six algorithms, the RSEA is the most reliable and has the most similarity to the true Pareto front compared to the other algorithms. Moreover, the coverage and diversity of the algorithms are tested using two performance metrics. The numerical results prove that the RSEA algorithm has good performance. This analysis will drastically reduce the amount of time needed for field inspections. Instead of visiting each weld in the infrastructure, the managers can focus on the welds that provide Pareto optimal solutions based on ABA signals.

In the future, a distributed system could be applied to the proposed framework to include larger sections of tracks. Moreover, by considering the track characteristics such as track geometry, a more elaborate framework for KPIs can be acquired. Another topic for future research is to consider the life-cycle costs as an objective function so that the rail replacement criterion can be reformulated according to different operation costs. Other objectives such as social value and environmental impact can also be included when adequate tools are able to dynamically assess their effects. While this chapter is focused on analyzing rail welds, the methodology can be used to analyze other types of rail defects, such as corrugations, squats and damaged insulated joints. The durability of the ABA measurement system has not been investigated yet, and it will be a topic of future research. In this study, the ABA system was already used for a week during the day in mixed dry and wet conditions, and no technical problems that might affect the results were reported. The effects of high temperature variations, such as snow conditions or high temperatures (at 30°C the rail might have a temperature approximately 70°C), are future research topics.

Acknowledgement

This research was funded by Collaborative Project H2020-MG-2015-2015 GA-398 636237b Needs Tailored Interoperable Railway - NeTIRail-INFRA and also the NWO/ProRail project “Multi-party risk management and key performance indicator design at the whole system level (PYRAMIDS)”, project code 438-12-300, which is partly funded by the Ministry of Economic Affairs. We also thank Jan Moraal, Jurjen Hendriks, Ivan Ramirez, Zili Li and Rolf Dollevoet for their support during measurement campaign, discussions about technical usage of the data and funding. Finally, special thanks to the partners RCCF from Romania for their support during the measurement campaign in the NeTIRail-INFRA project. In this chapter, the authors are users of the software PlatEMO (Evolutionary multi-objective optimization platform).

5.5 References

[1] Andersson, M., Björklund, G., & Haraldsson, M. "Marginal railway track renewal costs: A survival data approach". *Transportation Research Part A: Policy and Practice* 87:68-77, 2016.

[2] Attoh-Okine, N. *Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering*. John Wiley & Sons, 2017.

[3] Hu, H., Tang, B., Gong, X., Wei, W., & Wang, H. "Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks". *IEEE Transactions on Industrial Informatics 13*(4): 2106-2116, 2017.

[4] Hodge, V.J., O'Keefe, S., Weeks, M., & Moulds, A. "Wireless sensor networks for condition monitoring in the railway industry: A survey". *IEEE Transactions on Intelligent Transportation Systems* 16(3):1088-1106, 2015.

[5] Berlin, E., & Van Laerhoven, K. "Sensor networks for railway monitoring: Detecting trains from their distributed vibration footprints". In *Proceedings of the 2013 IEEE International Conference on Distributed Computing in Sensor Systems,* 80-87. IEEE, 2013.

[6] He, Q., Li, H., Bhattacharjya, D., Parikh, D., & Hampapur, A. "Track geometry defect rectification based on track deterioration modelling and derailment risk assessment". *Journal of the Operational Research Society* 66(3): 392-404, 2015.

[7] Wang, J., Liu, X., & Ni, Y.Q. "A bayesian probabilistic approach for acoustic emission-based rail condition assessment". *Computer-Aided Civil and Infrastructure Engineering* 33(1): 21-34, 2018.

[8] Liu, X., Barkan, C., & Saat, M. "Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk." *Transportation Research Record: Journal of the Transportation Research Board* 2261: 178-185, 2011.

[9] Papaelias, M., Amini, A., Huang, Z., Vallely, P., Cardoso, D., & Kerkyras, S. "Online condition monitoring of rolling stock wheels and axle bearings". *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 230(3): 709-723, 2016.

[10] Salvador, P., Naranjo, V., Insa, R., & Teixeira, P. "Axlebox accelerations: their acquisition and time–frequency characterisation for railway track monitoring purposes". *Measurement* 82: 301-312, 2016.

[11] Molodova, M., Li, Z., Núñez, A., & Dollevoet, R. "Automatic detection of squats in railway infrastructure". *IEEE Transactions on Intelligent Transportation Systems* 15(5):1980-1990, 2014.

[12] Huang, N.E., & Attoh-Okine, N., *The Hilbert-Huang transform in engineering*. CRC Press, 2005.

[13] Fan, Y., Dixon, S., Edwards, R. S., & Jian, X. "Ultrasonic surface wave propagation and interaction with surface defects on rail track head". *Ndt & E International* 40(6): 471-477, 2007.

[14] Song, Z., Yamada, T., Shitara, H. & Takemura, Y. "Detection of damage and crack in railhead by using eddy current testing". *Journal of Electromagnetic Analysis and Applications* 3:546–550, 2011.

[15] Li, Q., & Shengwei, R. "A real-time visual inspection system for discrete surface defects of rail heads." *IEEE Transactions on Instrumentation and Measurement* 61(8): 2189-2199, 2012.

[16] Sun, Y. Q., Spiryagin, M., Wu, Q., Cole, C., & Ma, W. H. "Feasibility in assessing the dipped rail joint defects through dynamic response of heavy haul locomotive." *Journal of Modern Transportation*, 1-11, 2018.

[17] Lee, J. S., Choi, S., Kim, S. S., Park, C., & Kim, Y. G. "A mixed filtering approach for track condition monitoring using accelerometers on the axle box and bogie". *IEEE Transactions on Instrumentation and Measurement* 61(3):749-758, 2012.

[18] Lee, J. S., Choi, S., Kim, S. S., Kim, Y. G., Kim, S. W., & Park, C. "Waveband Analysis of Track Irregularities in High-Speed Railway from On-Board Acceleration Measurement". *Journal of Solid Mechanics and Materials Engineering* 6(6): 750-759, 2012.

[19] Massel, A. "Power spectrum analysis—modern tool in the study of rail surface corrugations". *NDT & E International* 32(8):429-436, 1999.

[20] Bocciolone, M., Caprioli, A., Cigada, A., & Collina, A. "A measurement system for quick rail inspection and effective track maintenance strategy". *Mechanical Systems and Signal Processing*, 21(3) 1242-1254, 2007.

[21] Wei, Z., Boogaard, A., Núñez, A., Li, Z., and Dollevoet, R., "An integrated approach for characterizing the dynamic behavior of wheel-rail interaction at crossings". *IEEE Transactions on Instrumentation and Measurement*, 2018. DOI: 10.1109/TIM.2018.2816800

[22] Molodova, M., Oregui, M., Núñez, A., Li, Z., & Dollevoet, R. "Health condition monitoring of insulated joints based on axle box acceleration measurements". *Engineering Structures*, 123, 225-235, 2016.

[23] Herráiz, J. I. R., Domingo, M. L. M., Real, T., & Puig, V. "Development of a system to obtain vertical track geometry measuring axle-box accelerations from in-service trains". *Journal of Vibroengineering* 14(2): 813-826, 2012.

[24] Wei, Z., Núñez, A., Li. Z., & Dollevoet, R. "Evaluating degradation at railway crossings using axle box acceleration measurements". *Sensors,* 2017 17(10), 2236, 2017.

[25] Lee, S.H., Kim, S.H., Chang, Y.S., & Jun, H.K. "Fatigue life assessment of railway rail subjected to welding residual and contact stresses". *Journal of Mechanical Science and Technology* 28(11): 4483-4491, 2014.

[26] Romano, S., Beretta, S., Galli, G.S., & Riccardo, R. "Determination of inspection intervals for welded rail joints on a regional network". *Procedia Structural Integrity* 4:87-94, 2017.
[27] Jamshidi, A., Núñez, A., Dollevoet, R., & Li, Z. "Robust and predictive fuzzy key performance indicators for condition-based treatment of squats in railway infrastructures". *Journal of Infrastructure Systems* 23(3): 04017006, 2017.
[28] Huang, Norden Eh. *Hilbert-Huang transform and its applications*. Vol. 16. World Scientific, 2014.
[29] Velarde, P., Valverde, L., Maestre, J.M., Ocampo-Martinez, C., & Bordons, C. "On the comparison of stochastic model predictive control strategies applied to a hydrogen-based microgrid". *Journal of Power Sources* 343: 161-173, 2017.
[30] Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., Li, Z., & De Schutter, B. "A big data analysis approach for rail failure risk assessment". *Risk analysis* 37(8): 1495-1507, 2017.
[31] Barreiro-Gomez, J., Ocampo-Martinez, C., & Quijano, N. "Dynamical tuning for MPC using population games: a water supply network application." *ISA Transactions* 69:175-186, 2017.
[32] Tian, Y., Cheng, R., Zhang, X., Cheng, F., & Jin, Y. "An indicator based multi-objective evolutionary algorithm with reference point adaptation for better versatility". *IEEE Transactions on Evolutionary Computation* (2017).
[33] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. "A fast and elitist multiobjective genetic algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation* 6(2): 182-197, 2002.
[34] Zitzler, E., Laumanns, M., & Thiele, L. "SPEA2: Improving the strength Pareto evolutionary algorithm". *TIK-report* 103, 2001.
[35] Yang, S., Li, M., Liu, X., & Zheng, J. "A grid-based evolutionary algorithm for many-objective optimization". *IEEE Transactions on Evolutionary Computation*17(5): 721-736, 2013.
[36] Peng, S., & Li, Y. "A multi-objective evolutionary algorithm based on Riemann-sphere". In *Proceedings of the IEEE International Conference on Computer Science and Software Engineering,* pp. 565-568, 2008.
[37] Xiang, Y., Zhou, Y., Li, M., & Chen, Z. "A vector angle-based evolutionary algorithm for unconstrained many-objective optimization". *IEEE Transactions on Evolutionary Computation* 21(1):131-152, 2017.
[38] Li, M., Yang, S., & Liu, X. "Pareto or non-Pareto: Bi-criterion evolution in multiobjective optimization". *IEEE Transactions on Evolutionary Computation* 20(5): 645-665, 2016.
[39] Tian, Y., Cheng, R., Zhang, X. and Jin, Y. "PlatEMO: A MATLAB Platform for Evolutionary Multi-Objective Optimization". *IEEE Computational Intelligence Magazine* 12(4):73-87, 2017.

Conclusion and discussion

The conclusion is presented in four parts. First, conclusion related to the KPIs is discussed and in the second part, the rail maintenance decision support system is concluded. The third part presents major recommendations to ProRail and the last part describes future research and what potentially can be fulfilled in the future.

6.1 Conclusion from the effect of the KPIs on the system performance

Chapter 2 defines a set of rail KPIs for a type of surface defect called squats. The global scheme of the chapter included two major parts, namely condition monitoring and health condition estimation. The chapter relied on ABA measurements for the actual rail health condition. The first step was how to use ABA measurement data in a way that can be applicable to the design of KPIs. However, data transformation from acceleration to an actionable defect characteristic was challenging. We defined a simple model to make the transformation happen. Thus, a model is used to convert the energy value of the ABA signals to visual length of the defects. From the point of view of the infrastructure manager, there should be a difference between defects in terms of their severities. We use the severity model investigated and presented in [1]. The model analyzes mechanical nature of squat and its evolution. Thus, we only rely on the visual lengths and classify the defects according to the visual lengths while by using crack depth of defects, more detailed information on the defect growth can be obtained. In the chapter, we suppose that

over the prediction time horizon, no new squat will develop within the rail. This is on contrary with the nature of the rail which is always possible to have new squats. Thus, a more practical prediction model can take the generation of the new squats in the prediction model into account (In Chapter 4, we set a primary investigation to help for estimating the critical pieces of tracks, prone to defect appearance). Hence, research regarding the effect of the new generation of the defects is for future studies. By having the predictive and robust KPIs, the infrastructure manager can think of a maintenance plan to systematically improve the performance. The chapter resulted in a simple maintenance recommendation in terms of rail grinding and rail replacement. The considerable difference between the rail health condition in doing nothing and the health conditions when doing grinding and rail replacing shows how influential the design of the rail KPIs can be.

In Chapter 3, the focus is on how to define a risk KPI for rail. A new methodology is developed to analyse rail failure risk probability. Video images are employed to detect the squats for a long track Zwolle-Groningen using a big data approach. Once all the defect images together with their corresponding kilometre positions were obtained, an analysis on their visual lengths were done. Beside the visual length, the authors use the data collected by ultrasonic measurements. The aim is to include the crack depth in the defect evolution model. By having a double source of the data (video images and ultrasonic), a posterior function was provided for the estimate of the failure probability. Thus, the composition of the probability function and the crack growth function is considered in order to give the rail failure estimation. Nonetheless, estimating the failure risk requires consequence analysis as well. We only focus in this chapter on the failure probability and the consequence analysis requires more detailed operation cost analysis which can be the subject of future research. Moreover, the chapter uses yearly train traffic (MGT) to include the effect of defect growth prediction model. Overall, the failure probability can be applicable to replacement maintenance to give the probability on how likely can be for a rail piece to break till next measurement campaign in a given MGT. The proposed methodology can be more practical with using more real-life data for the crack depth analysis as we use limited number of squats for the analysis.

6.2 Conclusion from the condition-based rail maintenance methodology

In Chapter 4, an integrated framework for maintenance decision of the railway infrastructures is proposed. The proposed approach is applied to the condition-based treatment of squats, with big data information coming from a track in the Dutch railway network. After experiencing use of the ABA signals and videos images in Chapters 2 and 3, the proposed algorithm in this chapter makes use of both ABA signals and rail video images, simultaneously, which consists of a huge amount of data. The use of both rail data sources reduces the detection error of the surface defects. Moreover, we have used the track characteristics of the Dutch railway network, enabling the infrastructure manager to link the track influential factors to the actual rail health condition. The interdependency analysis between the track characteristics and the actual rail health condition is conducted over a long track. However, sufficient information was not available for a more detail analysis for the interdependency analysis in more details, e.g. every 10 meter. The proposed maintenance decision system uses a clustering model to perform grinding over the critical

pieces of the rail. The results include the most severe squats covered by the maintenance clusters. Although not all the squats are treated and there might still be some kilometre spots with squats, maintenance cost until the next rail measurement campaign should not be too high as the most critical track pieces are treated. The proposed methodology considers the maintenance time slot as a constraint, to include possible practical limitations as well. A potential improvement to the methodology can be the number of grinding machine passing over the critical pieces to be ground. We only use one passage and the rail is supposed to be sufficiently treated while in the practice, a rail might need multiple passages to become fully healthy.

Chapter 5 is the study on the regional railway networks in which technical support for defect diagnosis and rail maintenance are more challenging compared to regular railway networks due to the difficulties caused by noise in the measurement, less GPS reliability in mountainous regions etc. Furthermore, beside Chapter 4 that is on how an expert system can be used for a rail maintenance decision support system, in this chapter, a Pareto-based algorithm is used to give optimal maintenance decisions. Therefore, in Chapter 5, like Chapter 4, a KPI for defects was directly extracted from the ABA signals. A Hilbert spectrum approach is used for detecting damaged welds. The Hilbert spectrum approach has not been addressed yet in the literature to detect rail defects. The approach helps to acquire the KPIs including the uncertainty of the ABA signals in different growth scenarios. So, a set of robust KPIs is defined. Moreover, a multi-objective based approach is used for optimizing the maintenance decisions. We could use a single objective function but as the trade-offs cannot be clearly tested between performance and cost functions, a multi-objective approach enables the infrastructure manager to see the trade-offs and decide better.

6.3 Future research

Throughout all the chapters, there are some potential improvements that can be taken into account in the future for the proposed framework. The major potentialities are highlighted as follows:

- We rely on ABA measurement systems, rail video images and ultrasonic measurements. However, demand for cheap and easy-to-use condition measurement systems is becoming of major concern of infrastructure managers, in particular, for the regional networks. It means that infrastructure manager requires a cost-effective monitoring system which costs less compared to other options and keeps the performance at an acceptable level. Wireless condition monitoring systems are among the list of the new measurement technologies. Moreover, sensor devices have become cheaper due to the recent advances in networking technologies and mobile ad hoc networking. Therefore, wireless sensor networks (WSNs) can be used for monitoring railway infrastructure. There are different types of the cheap tools for WSNs, most importantly using smart phones, drones and laser camera [2], [3], [4], [5].

- While this dissertation is focused on the analysis of squats and welds, the results are applicable to analyze other types of rail defects like corrugations, damaged insulated joints, and other types of RCF defects.

- A rail maintenance cost analysis could be added to the framework in order to reduce life cycle cost (LCC). The LCC considers costs associated with the lifetime of a rail infrastructure, namely, construction costs, operating costs, maintenance costs, energy costs, and taxes, and capital costs. This is not only useful for the maintenance cost analysis, but also could be used for a consequence analysis of different maintenance solutions.
- Another topic for further research can be evaluation of the methodology at different regions to investigate the influence of exogenous factors like environmental and socioeconomic factors to the decision rules and consequently the maintenance decisions. Hence, it is required to perform an assessment considering economic (cost-benefit analysis), social and environmental effects in order to estimate the real benefit provided by maintenance strategies. To do so, multi-criteria analysis of the factors can be used to evaluate the influence by quantifying both impacts, i.e. quantitative and qualitative factors. Multi-criteria decision-making methods, particularly ELECTRE (elimination et choix traduisant la réalité), AHP (analytic hierarchy process) and TOPSIS (technique for order of preference by similarity to ideal solution) have been widely used for this purpose.
- We select a scenario-based approach to cope with the uncertainties related to the ABA measurements throughout the dissertation. However, there are different strategies to include uncertainties in optimization problems. Some of them are of future prospect of this research. Stochastic optimization methods to tackle the real-world problems give better view of the optimization problem compared to deterministic methods. When the variables of the problem are known only within specific data interval, robust optimization can be a solution. Stochastic programming models are substantially the same but with a difference in the fact that probability distributions governing the data are known or can be estimated. Thus, beside the scenario-based approach used in this dissertation, the other methods can be also used, e.g. stochastic linear program and statistical inference approach [6], [7].

6.4 Recommendations for ProRail

This PhD dissertation is sponsored by ProRail (the Dutch railway infrastructure manager). The objective of this project (which involves three PhD researchers) was the development of a new asset and risk management methodology. The results can consider multiple parties in order to guarantee the long-term objectives of the Dutch railway infrastructure manager. Thus, this dissertation is supposed to be insightful on how to design new KPIs and use the exiting KPIs for maintaining network performance in a cost-effective way. According to the three major steps described in the summary, this dissertation can help ProRail in the following recommendations:

- As the proposed KPIs include predictability and robustness, ProRail can have a better prospect on what is needed in future operations.
- This dissertation tests three different infrastructure sources for condition monitoring: (1) the ABA, (2) high quality rail video images and (3) ultrasonic measurements. Therefore, the infrastructure manager comprehends how efficient each source can be on the performance improvement of the infrastructure.

- At the moment of writing this dissertation, the Dutch railway network mostly relies on cyclic maintenance decisions. According to this dissertation, this maintenance approach is not always the best choice (Chapter 2) and can be replaced with a condition-based maintenance approach. Therefore, instead of maintaining periodically (normally, every 6 month), Prorail can make a balance between preventive maintenance, reactive maintenance and condition-based maintenance (maintaining only critical rail pieces) in order to save maintenance operational time and cost.
- The dissertation framework is presented in such a way that it can be applicable for other infrastructures. Thus, not just Dutch rails, but other infrastructures under the management of ProRail can use the framework such as bridge, tunnel, etc. To redo the framework for a new infrastructure, the framework structure would be substantially the same as introduced in this dissertation, with considering all differences that new infrastructure brings along compared to the rail infrastructure.
- Although the dissertation is based on real life data and validated and improved relying on railway expert opinions, both in the KPI's part and the maintenance part, still there should be a gap between what can be implemented in the real case and what is implemented in the proposed framework. The following ideas need to be improved for using directly the proposed framework in the field: (1) defect prediction model which can be improved as mentioned in 6.2, and (2) including life cycle cost analysis and (3) detailed involvement of track influential factors in the rail degradation. Also, the framework must be tested and validated in different Dutch railway tracks to analyze results consistency.
- Design of a business case can be taken into account. A track from Dutch railway network can be selected for the design. During the current research, a discussion was held with a grinding operation manager to apply the proposed framework of the dissertation to a business case; however, due to the lack of time, it was not feasible to do the case. The business case should be carried out with corresponding company partners for grinding operations in The Netherlands including ProRail (asset manager), main contractor (e.g. Speno International) which is in charge of grinding machines and personnel, main subcontractor (e.g. BAM Rail) in charge of planning, safety and cleaning and multiple smaller subcontractors. Thus, coordinating with the mentioned set of the partners for the business case is of important challenges. Additionally, as a business case requires a list of detailed information and enough track-related historical data, data management can be of difficulties to make the business case happen.

6.5 References

[1] Li, Z., Dollevoet, R., Molodova, M., & Zhao, X. (2011). Squat growth—Some observations and the validation of numerical predictions. *Wear*, *271*(1-2), 148-157.

[2] Flammini, F., Gaglione, A., Ottello, F., Pappalardo, A., Pragliola, C., & Tedesco, A. (2010). Towards wireless sensor networks for railway infrastructure monitoring. *IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)*, 19-21 October, Bologna, Italy.

[3] Flammini, F., Pragliola, C., & Smarra, G. (2016). Railway infrastructure monitoring by drones. *IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)*, 2-4 November, Toulouse, France.

[4] Berlin, E., & Van Laerhoven, K. (2013). Sensor networks for railway monitoring: Detecting trains from their distributed vibration footprints. *IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 80-87, 21-23 May, Massachusetts, USA.

[5] dos Santos, S. G., de Araújo, I. R., Anjos, E. G., Araújo, R. C., & Belo, F. A. (2017). Using Accelerometers to Improve Real Time Railway Monitoring Systems Based on WSN. *International Conference on Computational Science and Its Applications*, 761-769, 3-6 July, Trieste, Italy.

[6] Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2009). Lectures on stochastic programming: modeling and theory. *Society for Industrial and Applied Mathematics and mathematical programming society*, Philadelphia, USA.

[7] Ruszczyński, A., & Shapiro, A. (2003). Stochastic programming models. *Handbooks in operations research and management science*, 10, 1-64, Elsevier Science B.V.

Curriculum Vitæ

[Jan 2014-Jan 2018]	Doctoral candidate in Civil engineering PhD researcher at Faculty of Civil Engineering and Geosciences, Railway Section, Delft University of Technology in The Netherlands working on the project "PYRAMIDS - Multi-Party Risk Management and Key Performance Indicator Design at the Whole System Level". Main task: maintenance planning, infrastructure performance measurement, risk analysis in railway infrastructures projects in collaboration with ProRail (https://www.prorail.nl/). Dissertation: Design of Key Performance Indicators for Asset Management of Railways.
[Sep 2006-Dec 2008]	Master degree in Civil Engineering, University of Tehran (Iran). Dissertation: Developing a model for safety risk analysis of pipeline infrastructure.
[Sept 2002-Sep 2006]	Bachelor degree in Civil Engineering, BU-Ali Sina University, Hamadan, Iran. Final project: Failure analysis of water pipelines.

Publications

Journal papers

1. A. Jamshidi, A. Núñez, R. Dollevoet, and Z. Li, "Robust and predictive fuzzy key performance indicators for condition-based treatment of squats in railway infrastructures". Journal of Infrastructure Systems, Volume 23, Issue 3, September 2017, DOI: 10.1061/(ASCE)IS.1943-555X.0000357. *(Impact Factor 2018: 1.538)*
2. A. Jamshidi, S. Faghih-Roohi, S. Hajizadeh, A. Núñez, R. Babuška, R. Dollevoet, Z. Li and B. De Schutter, "A big data analysis approach for rail failure risk assessment". Risk Analysis, Volume 37, Issue 8, Pages: 1495-1507, August 2017. DOI: 10.1111/risa.12836. *(Impact Factor 2018: 2.564)*
3. A. Jamshidi, Z. Su, S. Hajizadeh, M. Naeimi, A. Núñez, R. Dollevoet, B. De Schutter, and Z. Li, "A decision support approach for condition-based maintenance of rails based on big data analysis". Transportation Research Part C: Emerging Technologies, Volume 95, Pages: 185-206, October 2018. DOI: 10.1016/j.trc.2018.07.007. *(Impact Factor 2018: 5.775)*
4. A. Núñez, A. Jamshidi, H. Wang, "Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration". IEEE Transactions on Industrial Informatics. Volume 15, Issue 3, Pages: 1496-1507. May 2018. DOI: 10.1109/TII.2018.2847736. *(Impact Factor 2018: 7.377)*

Collaboration journal papers

5. Z. Su, A. Jamshidi, A. Núñez, S. Baldi, and B. De Schutter, "Multi-level condition-based maintenance planning for railway infrastructures – A scenario-based chance-constrained approach". Transportation Research Part C: Emerging technologies, Volume 84, Pages: 92-123, November 2017. DOI: 10.1016/j.trc.2017.08.018. *(Impact Factor 2018: 5.775)*
6. Z. Su, A. Jamshidi, A. Núñez, S. Baldi, and B. De Schutter, "Integrated condition-based track maintenance planning and crew scheduling of railway networks". Transportation Research Part C: Emerging Technologies, Volume 105, Pages: 359-384. June 2019. DOI: 10.1016/j.trc.2019.05.045. *(Impact Factor 2018: 5.775)*

Collaboration chapter book

Z. Su, A. Jamshidi, A. Núñez, S. Baldi, B. De Schutter, "Distributed Chance-Constrained Model Predictive Control for Condition-based Maintenance Planning for Railway Infrastructures", Predictive Maintenance in Dynamic Systems, Pages: 533-554, March 2019. Springer, Cham. DOI: ://doi.org/10.1007/978-3-030-05645-2_18.

Conference papers

1. A. Núñez, A. Jamshidi, H. Wang, J. Hendriks, I. Ramirez, J. Moraal, R. Dollevoet, Z. Li "Evolutionary multiobjective optimization for rail maintenance operations in a regional railway network", IEEE World Congress on Computational Intelligence (WCCI 2018), Windsor Convention Centre, Rio de Janeiro, Brazil, July 08-13, 2018.
2. A. Jamshidi, S. Hajizadeh, M. Naeimi, A. Núñez, and Z. Li, "Influencing factors for condition-based maintenance in railway tracks using knowledge-based approach". Proceedings of the First International Conference on Rail Transportation (ICRT2017), Chengdu, China, July 10-12, 2017.
3. A. Jamshidi, S. Faghih-Roohi, A. Núñez, R. Babuska, B. De Schutter, R. Dollevoet and Z. Li, "Probabilistic defect-based risk assessment approach for rail failures in railway infrastructure". Proceedings of the 14th IFAC Symposium on Control in Transportation Systems, CTS 2016, Istanbul, Turkey, 18-20 May, 2016. IFAC-PapersOnLine 49(3), pp. 73-77 DOI: 10.1016/j.ifacol.2016.07.013
4. A. Jamshidi, A. Núñez, M. Molodova, Z. Li, and R. Dollevoet, "Key performance indicators using an interval based fuzzy prediction modeling to treat squats in railway infrastructures". Proceedings of the Third International Conference on Railway Technology: Research, Development and Maintenance, RAILWAYS2016, Cagliari, Sardina, Italy, 5-8 April, 2016, Paper 159. DOI: 10.4203/ccp.110.159
5. Z. Su, A. Núñez, A. Jamshidi, S. Baldi, Z. Li, R. Dollevoet, and B. De Schutter, "Model predictive control for maintenance operations planning of railway infrastructures". Proceedings of the 6th International Conference on Computational Logistics, Delft, The Netherlands, September 23-25, 2015. DOI: 10.1007/978-3-319-24264-4_46
6. A. Jamshidi, A. Núñez, Z. Li, and R. Dollevoet, "Fuzzy maintenance decision support for treating squats in railway infrastructures". Proceedings of the Joint Rail Conference 2015, JRC2015, San Jose CA, USA, 23-26 March, 2015. DOI: 10.1115/JRC2015-5676

7. A. Jamshidi, A. Núñez, Z. Li, and R. Dollevoet, “Maintenance decision indicators for treating squats in railway infrastructures”. Proceedings of the 94th Annual Meeting of the Transportation Research Board, Washington, USA, January 11-15, 2015.