Track Geometry Monitoring using Measured Data from Commercial Trains towards Predictive Maintenance

Faculty of Civil and Industrial Engineering
Department of Civil, Construction and Environmental Engineering (DICEA)
Ph.D. in Infrastructure and Transport Systems Engineering

Ph.D. Student:

Nadia Kaviani

Supervisor:

Prof. Stefano Ricci

Co-supervisors:

Prof. Riccardo Licciardello

Prof. Luca Rizzetto

A.Y. 2023-2024

Acknowledgements

The journey to my PhD Thesis was not only an academic career for me, but also an incredible life style where I met lots of international experts and friends, who provided valuable insights for both professional and personal growth.

In the accomplishment of this thesis, foremost, I would like to express my heartfelt gratitude to my supervisor Prof. Stefano Ricci, for his guidance and ongoing support. His encouragement led me to participate in the European project "Assets4rail," fostering my research activities and exposing me to a more profound academic environment. Prof. Ricci's continuous support and motivation empowered me to step out of my comfort zone, engaging in various exchange programs abroad. I couldn't have asked for a better advisor and mentor for my thesis.

Next to him, I feel obliged in taking the opportunity to sincerely thanks to my incredible co-supervisors Prof. Luca Rizzetto, and Prof. Riccardo Licciardello for their patience, motivation, enthusiasm, immense knowledge, and their insightful comments. Their guidance played a crucial role throughout the research and writing phases of this thesis. Prof. Rizzetto has been always there for me, and his innovative ideas and challenges broadened my perspective, while Prof. Licciardello provided constant support and guidance in every step of the process, both in person and online.

Beyond Sapienza University of Rome, my thesis took shape at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway, and the UK Rail and Research Innovation Research Centre at the University of Birmingham (UKRRIN). I express my gratitude to my external mentors, including Prof. Gunnstein Thomas Frøseth, Prof. Albert Lau, Prof. Anders Rønnquist at NTNU, and Dr. Mani Entezami at UKRRIN.

A grateful appreciation goes to the projects and institute which had a role in the success of this work. Hence, thanks to the EU Shift2Rail project Assets4Rail (Grand number: 826250), which funded a part of this research under Horizon 2020 Framework Programme. Special thanks to the Italian Railways (Ferrovie dello Stato Italiane), UK National Railways (Network Rail High Speed), and MoniRail's monitoring system for their contributions to the success of this work through test fields and measurements.

I express my deepest admiration to my lovely parents who never gave up on me and took my hand from distance to keep me on the track. It was not easy but only possible with their infinite love, prayers, caring and sacrifices for me to stay motivated. I am grateful to my sibling, Ilia, for his constant presence and support.

In addition, I take a moment to express my sincere gratitude to my second family in Rome, —Christina, Guido, Giorgia, and Francesco—who went above and beyond to make me feel at home and assisted me in overcoming every challenge I encountered. Their continuous support has defined the true meaning of kindness for me. And my special appreciation goes to Laura and Steno for their kind contribution to my journey.

Lastly, I am thankful to my lifetime friends for being a continuous source of inspiration on this journey, in particular the ones who shared the academic journey with me including Zara, Ami, and Hossein. Words fall short compared to the immense support and kindness I have received from every individual throughout these years.

List of appended papers

- ## Paper 1

Development of a contactless sensor system to support rail track geometry on-board monitoring / *Marco Antognoli, Gintautas Bureika, Stefano Ricci, Luca Rizzetto, Viktor Skrickij, Nadia Kaviani* --- Presented at CETRA 6th International Conference on Road and Rail Infrastructure , Zagreb , Croatia.
https://master.grad.hr/cetra/ocs/index.php/cetra6/cetra2020/paper/view/1105

- ## Paper 2

Deep learning based virtual point tracking for real-time target-less dynamic displacement measurement in railway applications / *Shi, Dachuan; Kaviani, Nadia ,ˇSabanoviˇc, Eldar; Rizzetto, Luca; Skrickij, Viktor; Oliverio, Roberto; Ye, Yunguang; Bureika, Gintautas; Ricci, Stefano; Hecht, Markus*. **-** In: CASE STUDIES IN MECHANICAL SYSTEMS AND SIGNAL PROCESSING Journal (**Elsevier**)
https://www.sciencedirect.com/science/article/pii/S0888327021008268?dgcid=author

- ## Paper 3

Track geometry monitoring by an on-board computer-vision-based sensor system / *Matteo Circelli, Nadia Kaviani*, Riccardo Licciardello, Stefano Ricci, Luca Rizzetto, Sina Shahidzadeh Arabani, Dachuan Shi* **---** Presented at AIIT 3rd International Conference on Transport Infrastructure and Systems (TIS ROMA 2022), 15th-16th September 2022, Rome, Italy – Published in Transportation Research Procedia Journal/ Science direct (**Elsevier)**
https://doi.org/10.1016/j.trpro.2023.02.170

- ## Paper 4

Effects of rail vehicle dynamics modelling choices on machine learning analysis / *Nadia Kaviani, Riccardo Licciardello, Sina Shahidzadeh Arabani* – Presented at 28th IAVSD Symposium on Dynamics of Vehicles on Roads and Tracks (IAVSD 2023) in Ottawa, Canada- Published in the Springer Nature (EquinOCS system).

- ## Paper 5

Identifying lateral track irregularities by on-board monitoring of lateral displacement and use of machine learning algorithms / *Nadia Kaviani*, Anders Rønnquist , Gunnstein Thomas Frøseth , Albert Lau Stefano Ricci, Luca Rizzetto* – Presented at 11th Young Researchers Seminar In Lisbon, Portugal **(**In conference Proceedings)

- ## Paper 6

Detecting lateral track irregularities through onboard measurements with monitoring lateral acceleration and lateral displacement of Wheel/Rail /*Nadia Kaviani*, Anders Rønnquist, Gunnstein Thomas Frøseth , Albert Lau, Stefano Ricci* -(IF- Ingegneria Ferroviaria).

Table of Contents

Abbreviations

AHP	Analytic Hierarchy Process
AI	Artificial Intelligence
CNN	Convolutional neural networks
CV	Computer Vision
DL	Deep learning
GNSS	Global navigation satellite system
IMU	Inertial Measurement Unit
LD	Lateral Displacement
LDWR	Lateral Displacement of the Wheel to Rail
MBS	Multi Body Simulation
ML	Machine Learning
NMT	New Measurement Train
PR	Polynomial Regression
PSF	Point Spread Function
RFR	Random Forest Regression
RMSE	Root Mean Squared Error
ST	Standard Track
SVM	Support Vector Machine
TQI	Track Quality Index
TGM	Track Geometry Monitoring
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
TRV	Track Recording Vehicle
UGMS	Unattended Geometry Measurement System

List of figures

List of tables

Abstract

Recognizing track problems is a crucial responsibility for railway engineers to ensure safe train operations. Analysing track geometry (TG) characteristics is essential for developing an effective track condition monitoring strategy. Recent studies have focused on identifying track geometry irregularities using onboard vehicle dynamics data, with a particular emphasis on vertical irregularity. This is because it is easier to understand the relationship between vehicle reaction and acceleration in the vertical direction, and there is a straightforward method for identifying vertical irregularities.

In contrast, lateral irregularities are more challenging to detect due to factors such as wheel conicity, non-linearity of wheel-rail contact, and the Klingel hunting motion. Addressing the problem of lateral irregularities requires tracking the contact point between the wheel and the rail in the lateral direction. This research focuses on detecting lateral irregularities by identifying the relationship between lateral acceleration and the lateral displacement of the wheel in relation to the rail (LDWR).

The first part of this thesis is related to the activities of the EU project Assets4Rail, funded within Shift2Rail in which a method for detecting LDWR has been established by developing a sensor system as an alternative to inertial platforms, and developing the image processing algorithms for analysing the video output data. In addition, the next part of this research uses Machine Learning (ML) models to solve the non-linearity issue for detecting the lateral irregularities by LDWR. The study employs a supervised ML model trained and tested with numerical simulation results from Simpack, and Gensys software for straight and curved sections of various tracks. Testing different algorithms allows for identifying the best models for this purpose and determining the most efficient results.

Hence, this work summarizes the methodology and results of establishing a method for finding the relationship between the lateral acceleration and lateral irregularities from the lateral displacement of the wheel relative to the rail measured on board and proposing an algorithm to detect the lateral irregularities of the track.

The conclusion of this thesis can lead researchers to predictive maintenance in further studies by developing algorithms for an on-board sensor system monitoring capable of detecting LDWR and detecting lateral irregularities Overall, this research highlights the significance of addressing lateral irregularities in track maintenance to ensure safe and efficient train operations.

1) Introduction

In the ever-evolving landscape of transportation systems, the rail sector emerges as a vital arena for innovation and advancement. As societies seek efficient, safe, and sustainable modes of transit, the rail industry stands at the nexus of transformation. In this regard, the health of railway track is a critical issue for both safety and comfort of the passenger. This Ph.D. thesis marks a continuation of my previous academic exploration, extending the path forged during my master's thesis for track geometry monitoring, and aligning seamlessly with the collaborative aspirations of the Assets4rail project—a visionary initiative funded within the framework of Shift2Rail.

The activities of the Assets4ail project, which coincided with the outset of my master's Thesis finished in the middle of my Ph.D., and the results of the project to my doctoral pursuit, was merely a point of transition, not an end. This idea marked the beginning of my research and continued by new discoveries as well as new collaborations. This extension of the academic journey was accompanied by two enriching exchange programs—first at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway and later at the UK Rail Research and Innovation Network at Birmingham University in Birmingham, United Kingdom—where international collaboration not only broadened my horizons but also injected fresh perspectives into the evolution of this research.

1.1) Project Overview (Assets4Rail)

Assets4Rail was began in December 2018 and lasted for thirty months. It is an EU founded project that aims to contribute to the modal shift towards rail by exploring, adapting, and testing cutting-edge technologies for railway asset monitoring and maintenance, in order to ensure a proactive and cost-effective maintenance and intervention system in the assets.

The main objective of the project was to achieve cost efficient and reliable infrastructure developing a set of cutting-edge asset-specific measuring and monitoring devices.

To achieve that, Assets4Rail followed a twofold approach, including infrastructure (tunnel, bridges, track geometry, and safety systems) and vehicles. A dedicated information model (BIM) was the keystone of the infrastructure part of the project. Therefore, it was structured into two Work Streams (WS): the WS1 related to monitoring and upgrading solutions addressed to bridges and tunnels, the WS2 regarding monitoring solutions for trains and track geometry and data collection from fail-safe systems.

The project shared the Shift2Rail view of having an ageing European railway infrastructure that needs to cope with the expected increased traffics in the future. Likewise, reliable rolling stock was required to crystallize the desired modal shift to rail. Both goals relied on a proactive and cost-effective maintenance and intervention system in the assets.

Assets4Rail aimed to contribute to this modal shift by exploring, adapting, and testing cutting-edge technologies for railway asset monitoring and maintenance.

This model with integrated algorithms was gathered and analysed the information collected by specific sensors which monitors subsurface tunnel defects, fatigue consumption, noise, and vibrations of bridges as well as track geometry. Moreover, train monitoring was included by the installation of track side and underframe imaging automated system to collect data for detecting specific types of defects that have non-negligible impacts on infrastructure.

The additional use of the RFID technology could enable the smooth identification of trains and single elements, associated with the identified rolling stock failures.

The combination of mentioned real-time collected data with existing data along the implementation of deep learning techniques for assessing large data volumes paved the way towards a cost-effective and proactive maintenance process of infrastructure and rolling stock. In addition, two innovative intervention methods, noise rail dampers and the cleaning of long tunnel drainage pipes, were validated on field.

Assets4Rail benefited from a strong multidisciplinary consortium committed to concrete exploitation activities aligned towards the achievement of the challenging project objectives.

This work was focused on the workstream 2 challenges, and the layout of the project is shown in the figure below.

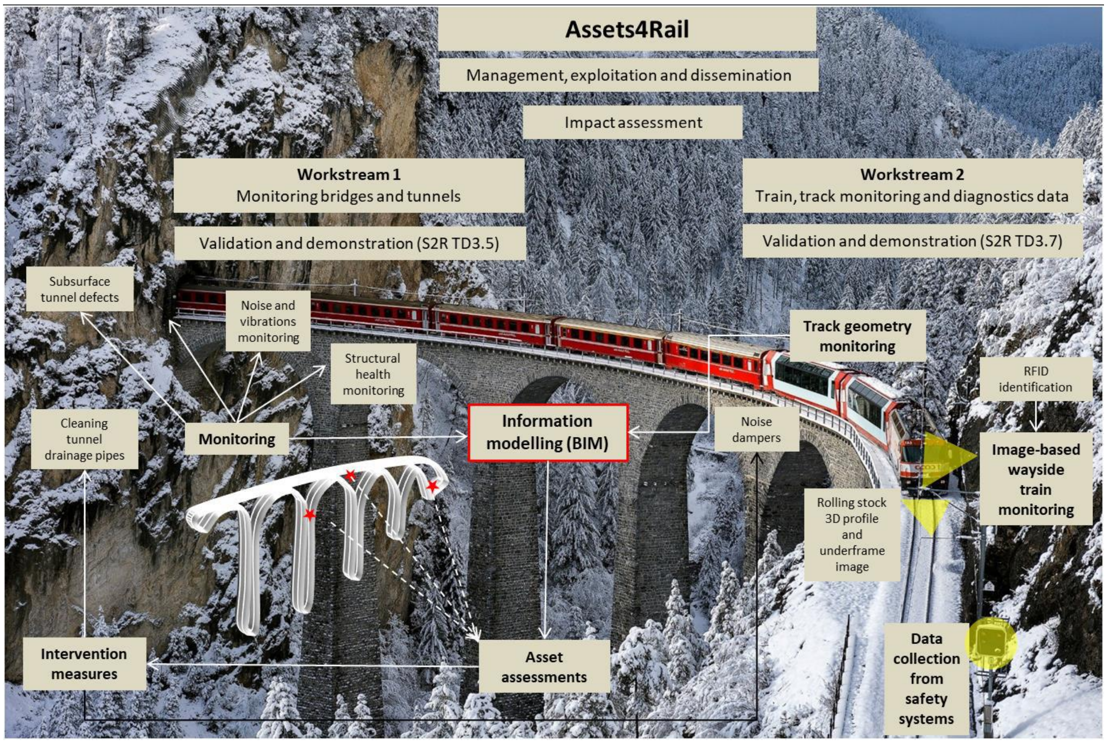


Figure 1 Concept of Assets4Rail

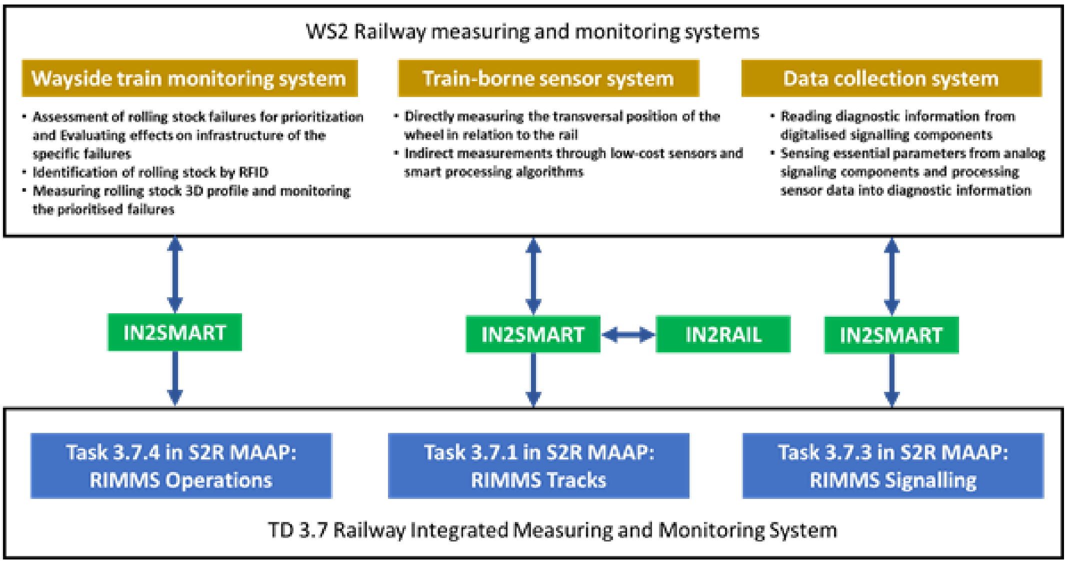


Figure 2 Work Stream 2 Railway Measuring and Monitoring Systems

In particular, the Work Stream 2 (See Figure 1-3) aimed to develop three asset-specific measuring and monitoring systems to achieve the following features:

a) Automatic inspections of specific railway assets by the development of a train-borne sensor system for support of measuring track geometry, a wayside system for monitoring rolling stock and a data collection solution for reading data from signalling components in order to reduce costs of time (money) consuming manual inspections and risks for the personnel involved;

b) Extract relevant maintenance infrastructure-related information from measurement data by the development of smart data processing methods for decision supports of maintenance activities;

c) Appropriate storage and management of generated data/information, considering big data solutions and standardised interface, to allow further data mining and transmission of data to other management systems.

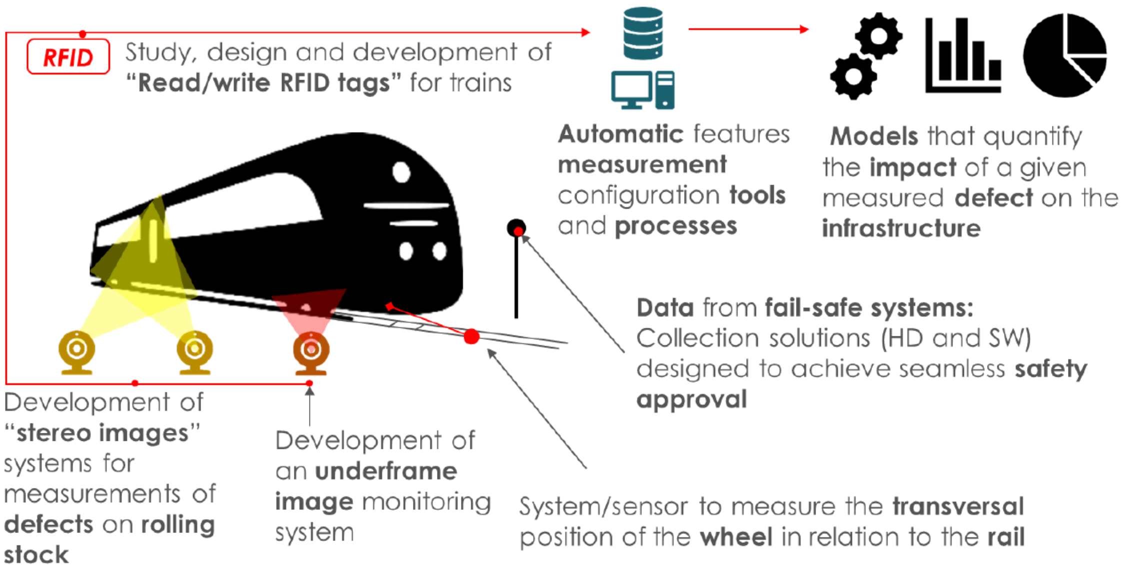


Figure 3 Logical scheme of the Work Stream 2

Within the Work Stream 2, which was the focus of this thesis and in association with Assets4Rail, the Benchmarks and Specifications dealt with the state-of-the-art overview on the relevant technologies and benchmark products for the subsequent developments of an innovative train monitoring system, a sensor system for track geometry monitoring and a data collection for diagnostics from signalling components. Furthermore, it aimed to the definition of requirement specifications for subsequent developments and test specifications for laboratory testing and pilots in the railway environments.

The special part of this project in this thesis work is specifically related to track geometry monitoring system and to the sensors and algorithms that this project had to develop to monitor the transversal position of the wheel in relation to the rail.

Objectives:

The impact of a geometry failure can be severe and include financial losses, interruptions to operations, harm to railroad property and the environment, and possible dangers to human life. Furthermore, poor track geometry leads to track deterioration and has a negative effect on subsequent loading (K. Karttunen, 2012). Maintenance procedures are carefully planned to either keep the geometrical condition at a level suitable for ride comfort and safety or bring it back to that level to prevent such consequences. To carry out an effective and efficient maintenance program, infrastructure managers must create a resource capacity plan that addresses labour, equipment, and maintenance window requirements, among other things. In addition, the creation of a master maintenance schedule is essential, considering the availability of resources, operational requirements, and the facilitation of opportunistic repair during execution. Predicting track geometry conditions and estimating maintenance requirements over medium- and long-term time frames are essential to these procedures.

Due to the interaction of multiple components, including dynamic train loads, substructure conditions, material qualities, and environmental factors, the degradation mechanism of track geometry is inherently complicated(Esveld, 2001)(Hakan Guler, 2014). Predicting the effect of various maintenance tactics on track geometry condition is difficult due to the complexity of the deterioration process and how it interacts with maintenance operations (Reeves et al., 2019, Quiroga & Schnieder, 2012). Consequently, the development of precise predictive models that address important aspects of railway track geometry degradation is crucial. Using these kinds of prediction models could help make inspection and maintenance scheduling decisions easier and better.

Track irregularity condition monitoring is crucial for ensuring the safety and optimal operation of railway network systems. These days, the majority of railway infrastructure managers monitor tracks by using advanced equipment mounted on specialized diagnostic railroad cars to measure geometrical anomalies.(Weston PF, 2007)

While this approach provides precise measurements of the track imperfections, it is an extremely costly and time-consuming procedure. Because of this, a lot of research has been done recently on the creation of novel techniques for track condition monitoring utilizing dynamic measures from in-service vehicles.(Weston PF, 2015).

The main objective of this thesis is to propose a proper on-board geometry monitoring system instead of inertial platforms of diagnostic trains, with an algorithm which could detect, and preferably predict the track geometry faults. Among all the geometry parameters, the focus of this work is on the lateral alignment, and the development of the alignment monitoring system is enabled by the success of the computer vision system for lateral displacement. There have been several numerical simulations, and real test to validate the results.

Research Questions

Specifically, the objectives or research questions that will be addressed through the methodology are as follows:

- To understand the correlation between lateral displacement of the wheel and lateral track geometry (TG) defects.
- To identify the most suitable Machine Learning model for in-service Lateral TG defect monitoring for the purpose of predictive maintenance.

2) Background and motivation

2.1) Literature Review

The safety and comfort of train operations rely heavily on effective track geometry monitoring (Cárdenas-Gallo, I., C. A., et al., 2017; Iman Soleimani, n.d.). The European Railway Standard EN13848-1 defines track geometry (TG) parameters as indicators of track maintenance, including track gauge, cross level, longitudinal level, lateral alignment, and twist (European Committee for Standardization (CEN), 2019), Among these parameters, longitudinal and lateral alignment are the most critical and require careful consideration during maintenance activities and the remaining parameter either decline more slowly than longitudinal level or show strong linear connection with longitudinal level (B. Ripke et al.).

Current track geometry monitoring is conducted using diagnostic trains equipped with heavy and expensive technology, such as inertial platforms (Weston PF, 2007). While this process provides accurate measurements of track faults, it is costly and time-consuming. To increase monitoring frequency, Infrastructure Managers (IMs) are increasingly interested in installing Unattended Geometry Measurement Systems (UGMS) on in-service trains without interfering with regular traffic (Weston PF, 2015).

The rapid advancements in computer vision (CV) over the past decade have led to a noticeable surge in the application of photogrammetry for structural inspections across various sectors. Photogrammetry, traditionally used for deformation measurements in large structures like bridges within civil engineering (Xu & Brownjohn, 2018), is now finding utility in the railway sector as well. For instance, (Zhan et al., 2018) introduced the use of high-speed line scan cameras, calibrated by a 1-D optical target, to measure catenary geometry parameters.(Li et al., 2020) employed CV techniques to monitor track slab deformation, with optical targets attached to the track slab to define regions of interest (RoI). In these applications, optical targets serve as critical reference points for measurements.

However, there are situations where attaching optical targets to structures is not feasible. In such cases, common solutions include edge detection, digital image correlation, pattern matching, and template matching (Baqersad et al., 2017). Nonetheless, these approaches often face challenges, particularly when dealing with complex background scenes.

To overcome these hurdles, (Jiang et al., 2020a) proposed a robust line detection method for measuring the uplift of railway catenary, specifically addressing issues arising from noisy backgrounds. Their measurement setup was static, with the camera system fixed next to the railway.

TG characteristics are important markers for track maintenance, according to the European Railway Standard (European Committee for Standardization (CEN), 2019), Track gauge, cross-level, longitudinal level, lateral alignment, and twist are all included in this list of criteria as mentioned earlier. A statistical analysis carried out as part of a European project (B. Ripke et al., n.d.) reveals that longitudinal level and lateral alignment are the most important variables influencing maintenance choices.

2.2) Track Geometry Monitoring

In this section the main geometric parameters that define the geometrical quality of track are listed. These geometrical parameters are stated by European standards (SS-EN 13848-1: 2004+A1, 2004), that also reports the descriptions about how these parameters have to be measured.

The railway track experiences settling due to permanent deformation in both the ballast and the underlying soil (Iwnicki, 2006). This settling is a consequence of static and dynamic forces generated by traffic, leading to deviations from the intended geometry. The geometry of the track, often referred to as track geometry, plays a crucial role in railway construction, representing the three-dimensional layout of the track. Maintaining high-quality track geometry is essential for ensuring optimal ride comfort for passengers and preventing premature wear of the track, as noted by(Esveld, 2001) and (Jovanovic, 2004). A well-maintained track geometry contributes to overall railway performance.(TRAFIKVERKET, 2017)

The geometry parameters are: Track gauge, Longitudinal level, Cross level, Alignment and Twist.

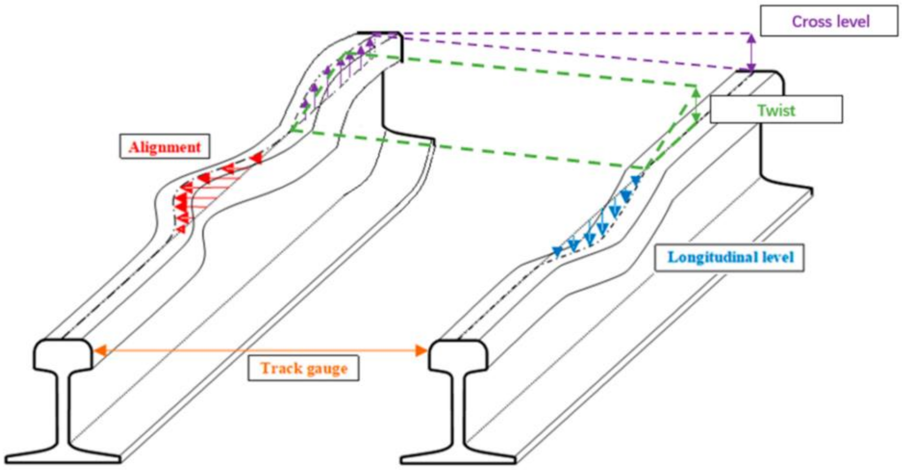


Figure 4 Track Geometry parameters Layout

2.1) Description of the track co-ordinate system

The track geometry quality is described by means of a relative rectangular co-ordinate system centred to the track with clockwise rotation (SS-EN 13848-1: 2004+A1, 2004).
X-axis: axis represented as an extension of the track towards the direction of running
Y-axis: axis parallel to the running surface
Z-axis: axis perpendicular to the running surface and pointing downwards

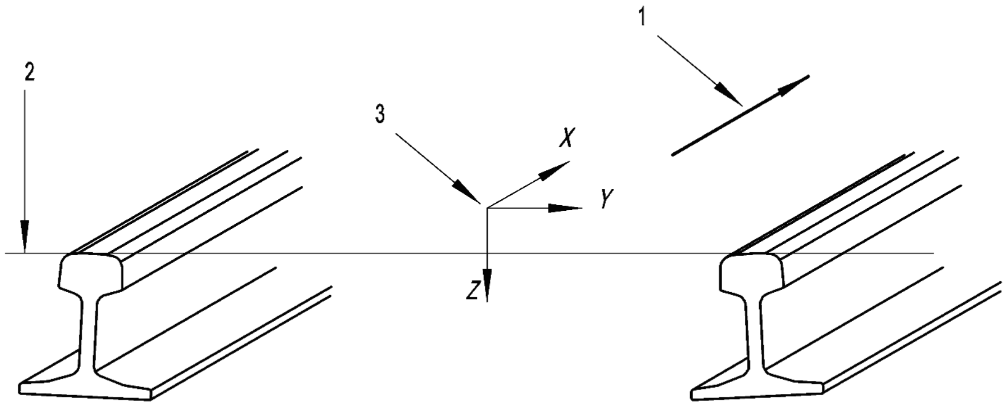


Figure 5 Relationship between the axes of the track co-ordinate system (SS-EN 13848-1: 2004+A1 2008).

Key

1. running direction
2. running surface
3. track co-ordinate system

2.2) Principal Track Geometric Parameters

The following parameters are needed to be covered in any generic track geometry measurement:

1. Track gauge

2. Longitudinal level

3. Cross level

4. Alignment

5. Twist

- Track gauge

Definition:

Track gauge, G, is the smallest distance between lines perpendicular to the running surface intersecting each rail.

Head profile at point P in a range from 0 to Zp below the running surface. Zp is always 14 mm.

In the situation of new unworn rail head the point P will be at the limit Zp below the railhead.

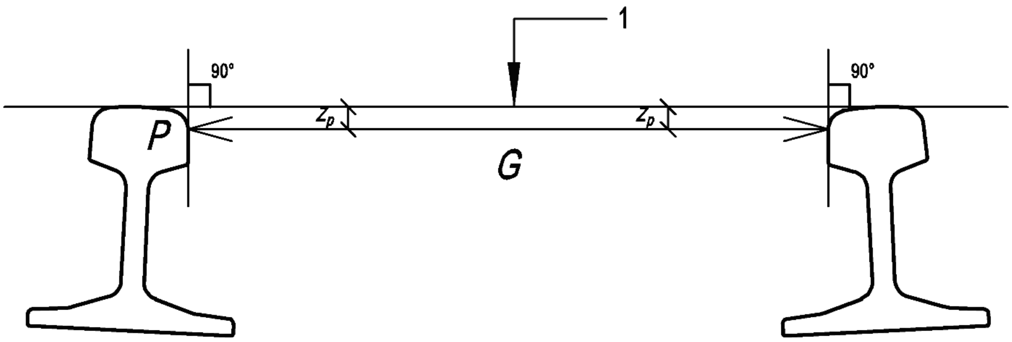


Figure 6 Track gauge for new rail (SS-EN 13848-1: 2004+A1 2008).

In the situation of used worn rail head the point P for the left rail can be different from the right rail.

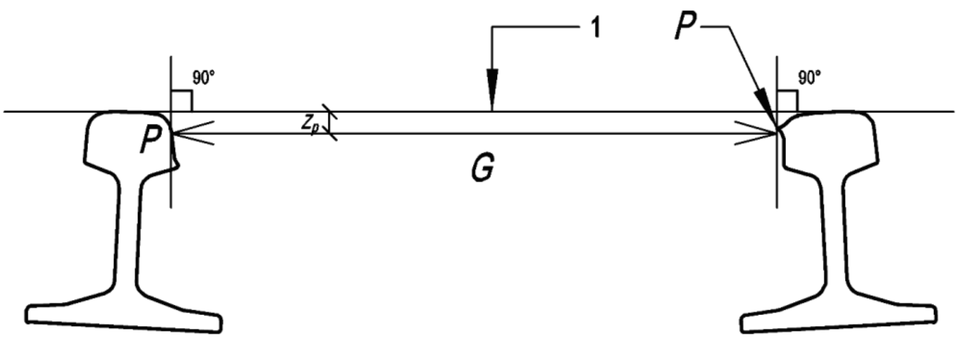


Figure 7 Track gauge for worn rail (SS-EN 13848-1: 2004+A1 2008).

1. running surface

Measurement method:

Track gauge can be measured using a contact system or a non-contact system.

Wavelength range

Not applicable

Resolution

Resolution ≤ 0,5 mm

Measurement uncertainty

Uncertainty ± 1 mm

Range of measurement

The range shall be the nominal gauge -15 mm/+50 mm

Analysis method

Individual defects are represented by the amplitude from the nominal value to the peak value (minimum and maximum peak value).

Output requirements

Measurements taken by recording vehicles, track maintenance machines or trolleys shall be recorded as a consecutive set of readings preferably in digital form and shall also be presented graphically.

For manually operated devices each track gauge measurement shall be recorded as a single value.

Output presentation

As a minimum, track gauge shall be prescribed by the following:

- The identification of individual defects which exceed a prescribed threshold.
- The measured track gauges.
- The difference between the measured track gauge and the nominal track gauge.
- The mean track gauge over a specified distance.
- The variation of track gauge over a specified distance.

❖ Longitudinal level

Definition:

Deviation zp' in z-direction of consecutive running table levels on any rail, expressed as an excursion from the mean vertical position (reference line), covering the wavelength ranges stipulated below and is calculated from successive measurements.

Figure 8 Longitudinal level (SS-EN 13848-1: 2004+A1 2008).

Key

1. running table
2. reference line

Measurement method:

Longitudinal level measurements shall either be made with an inertial system or a versine system (that should preferably be asymmetric) or by a combination of both methods. If the versine method of measurement is used, a re-coloration of the measured signals is necessary in order to eliminate the influence of the transfer function of the versine system.

Wavelength range:

Three ranges expressed in wavelengths (λ) shall be considered:

D1: 3 m < λ ≤ 25 m

D2: 25 m < λ ≤ 70 m

D3: 70 m < λ ≤ 150 m, used for measuring long wavelength defects. Generally, this range should only be considered for line speeds greater than 250 km/h.
In order to detect short wavelength defects, the lower limit of D1 should be reduced to 1 m. parameter. Wavelengths over 70 m can also be taken into consideration by the vertical curvature.
Vehicle responses to track irregularities with short wavelengths differ from those with long wavelengths (Choi, 2013). Long wavelength irregularities increase low-frequency oscillations of the train, while short wavelength irregularities cause vibrations and noise in both the train and the surrounding environment (Iwnicki, 2006). Given that short wavelength defects can pose a danger when their amplitude is high, they require special attention (SIST EN 13848-5:2008, 2008).

Resolution:

Table 1 Longitudinal level: resolution(SS-EN 13848-1: 2004+A1 2008).

Wavelength range	Dimensions in millimetres		
	D1	*D2*	*D3*
Resolution	≤ 0,5	≤ 0,5	≤ 0,5

Measurement uncertainty:

Table 2 Longitudinal level: measurement uncertainty(SS-EN 13848-1: 2004+A1 2008).

Wavelength range	Dimensions in millimetres		
	D1	*D2*	*D3*
Uncertainty	± 1	±3	± 5

Range of measurement:

Table 3 Longitudinal level: range of measurement(SS-EN 13848-1: 2004+A1 2008).

Wavelength range	Dimensions in millimetres		
	D1	*D2*	*D3*
Range of measurement	± 50	± 100	± 300

The ranges of amplitude stated for D2 and D3 will not be applicable for high-speed lines. However, to obtain accurate measurements in all wavelength ranges, the stated ranges of amplitudes should be measured.

Analysis method:

Individual defects are represented by the amplitude from the mean value (low pass filtered value) to the peak value.

Output requirements:

Measurements taken by recording vehicles, track maintenance machines or trolleys shall be recorded as a consecutive set of readings preferably in digital form and shall also be presented graphically. For manually operated devices each longitudinal level measurement shall be recorded as a single value.

Output presentation:

As a minimum, longitudinal level shall be prescribed by the following:

- Isolated defects that exceed a prescribed threshold.
- A standard deviation over a defined length, typically 200 m, in the wavelength range D1.

Other wavelength ranges, particularly D2 and D3, may also be selected for output presentation.

- Cross level

Definition:

The difference in height of the adjacent running tables computed from the angle between the running surface and a horizontal reference plane. It is expressed as the height of the vertical leg of the right-angled triangle having a hypotenuse that relates to the nominal track gauge plus the width of the rail head rounded to the nearest 10 mm.

- For nominal gauge of 1435 mm the hypotenuse is 1500 mm in length.
- For nominal gauge of 1524 mm the hypotenuse is 1600 mm in length.
- For nominal gauge of 1668 mm the hypotenuse is 1740 mm in length.

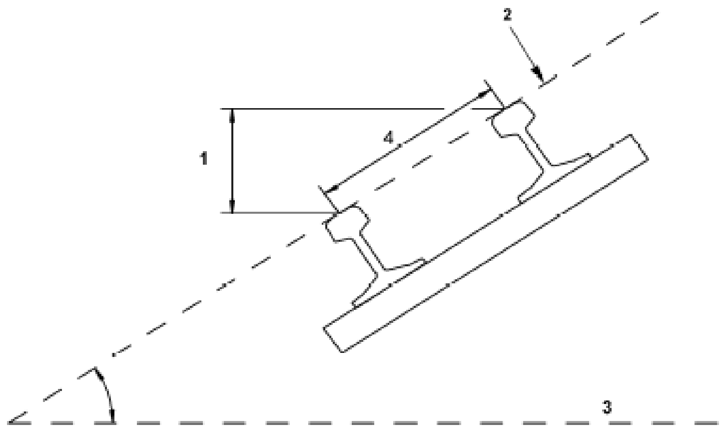


Figure 9 Cross level (SS-EN 13848-1: 2004+A1 2008).

Key

1. cross level
2. running surface
3. horizontal reference plane
4. hypotenuse

Measurement method:

Cross level is determined by measuring either the angle between the running surface and the horizontal reference plane or the difference in height between the two running tables.

Wavelength range:

Not applicable.

Resolution:

Resolution ≤ 0,5 mm.

Measurement uncertainty:

- cross level value: ± 5 mm,
- relative value (difference of successive cross level values) to be used for the twist calculation: ± 1 mm.

Range of measurement:

The range of measurements shall be ± 225 mm.

Analysis method:

Individual defects are represented by the amplitude from the mean value (low pass filtered value) to the peak value.
In addition, the measured values (defined as amplitude between zero and peak values) may be compared with the design values.

Output requirements:

Measurements taken by recording vehicles, track maintenance machines or trolleys shall be recorded as a consecutive set of readings preferably in digital form and shall also be presented graphically.
For manually operated devices each cross-level measurement shall be recorded as a single value.

Output presentation:

As a minimum cross level shall be prescribed by its absolute value.

- Alignment

Definition:

Deviation *yp* in y-direction of consecutive positions of point P (See Figure 10) on any rail, expressed as an excursion from the mean horizontal position (reference line) covering the wavelength ranges stipulated below and calculated from successive measurements.

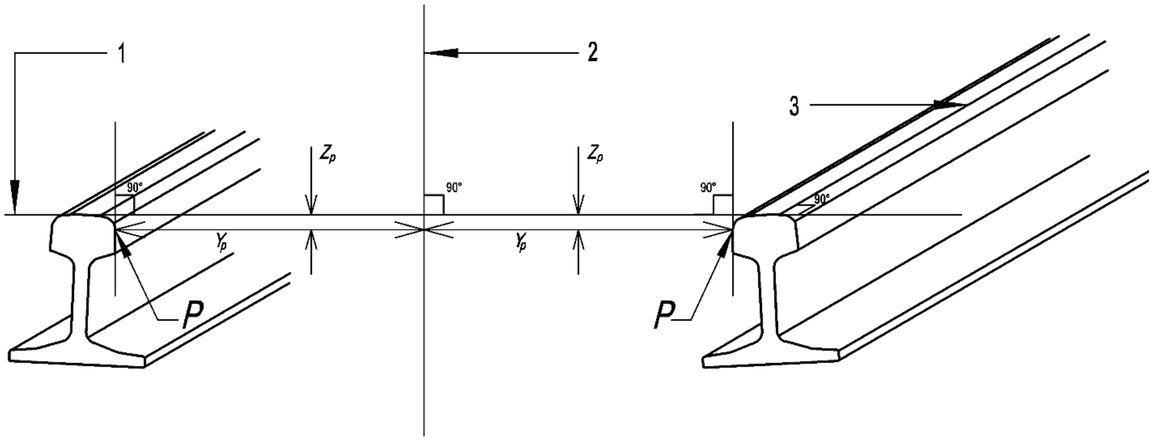


Figure 10 Alignment (SS-EN 13848-1: 2004+A1 2008).

Key

1. running surface
2. reference line
3. centre line of running table

Measurement method:

Alignment measurements shall either be made with an inertial system or a versine system (that should preferably be asymmetric) or by a combination of both methods.

If the versine method of measurement is used, a re-coloration of the measured signals is necessary in order to eliminate the influence of the transfer function of the versine system.

Wavelength range:

Three ranges expressed in wavelengths (λ) shall be considered:

- D1: 3 m < λ ≤ 25 m
- D2: 25 m < λ ≤ 70m
- D3: 70 m < λ ≤ 200 m, used for measuring long wavelength defects. Generally, this range should only be considered for line speeds greater than 250 km/h.

To detect short wavelengths defects, the lower limit of D1 should be reduced to 1 m.

Wavelengths over 70 m can also be taken in consideration by the horizontal curvature parameter.

Resolution:

The requirements are specified in the following table:

Table 4 Alignment: resolution(SS-EN 13848-1: 2004+A1 2008).

Wavelength range			Dimensions in millimetres
	D1	*D2*	*D3*
Resolution	≤ 0,5	≤ 0,5	≤ 0,5

Measurement uncertainty:

Table 5 Alignment: measurement uncertainty(SS-EN 13848-1: 2004+A1 2008).

Wavelength range			Dimensions in millimetres
	D1	*D2*	*D3*
Uncertainty	± 1,5	± 4	± 10

Range of measurement:

Table 6 Alignment: range of measurement(SS-EN 13848-1: 2004+A1 2008).

Wavelength range			Dimensions in millimetres
	D1	*D2*	*D3*
Range of measurement	± 50	± 100	± 500

The ranges of amplitude stated for D2 and D3 will not be applicable for high-speed lines. However, in order to obtain accurate measurements in all wavelength ranges, the stated ranges of amplitudes should be measured.

Analysis method:

Individual defects are represented by the amplitude from the mean value (low pass filtered value) to the peak value.

Output requirements:

Measurements taken by recording vehicles, track maintenance machines or trolleys shall be recorded as a consecutive set of readings preferably in digital form and shall also be presented graphically. For manually operated devices each alignment measurement shall be recorded as a single value.

Output presentation:

As a minimum, longitudinal level shall be prescribed by the following:

- Isolated defects that exceed a prescribed threshold.
- A standard deviation over a defined length, typically 200 m, in the wavelength range D1.
- 18

❖ Twist

Definition:

The algebraic difference between two cross levels taken at a defined distance apart, usually expressed as a gradient between the two points of measurement.

Twist may be expressed as a ratio (%o or mm/m).

Measurement method:

Twist measurements should either be taken simultaneously at a fixed distance e.g., at a distance equivalent to the wheel-base, or be computed from consecutive measurements of cross-level.

Wavelength range:

Not applicable.

Resolution:

Resolution ≤ 0,5 mm.

Measurement uncertainty:

The values are specified in the following table:

Table 7 Twist: measurement uncertainty(SS-EN 13848-1: 2004+A1 2008).

Measuring method	Uncertainty mm/m	
	ℓ ≤ 5,5 m	5,5 m < ℓ ≤ 20 m
direct measurement	± 1 / ℓ	± 2 / ℓ
computed from cross level	± 1,5 / ℓ	± 3 / ℓ

Range of measurement:

The range shall be ± 15 mm/m (%o).

Analysis methods:

Individual defects are represented either by the amplitude from the zero-line to the peak value (V1) or by the amplitude from the mean value (low pass filtered value) to the peak value (V2). For standard deviation the measurement of twist is represented by the amplitude from the low pass filtered value to the current measured value. (See Figure 11).

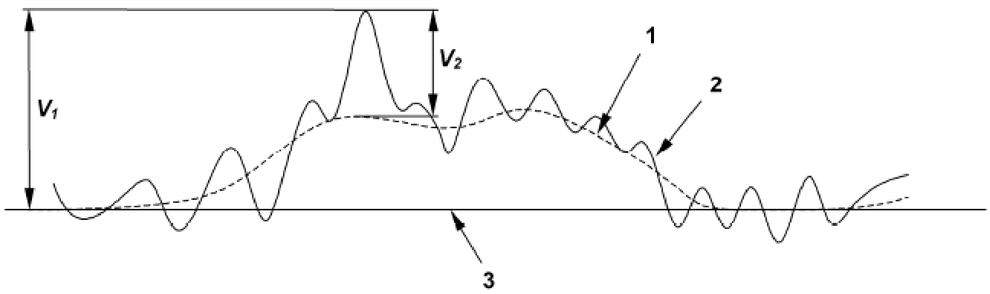


Figure 11 Twist - analysis method (SS-EN 13848-1: 2004+A1 2008).

Key

1. low pass filtered value
2. twist
3. zero line

Output requirements:

Measurements taken by recording vehicles, track maintenance machines or trolleys shall be recorded as a consecutive set of readings preferably in digital form and shall also be presented graphically.

For manually operated devices each twist measurement shall be recorded as a single value.

Output presentation:

As a minimum, twist shall be prescribed by the following:

- Isolated defects that exceed a prescribed threshold;

- A standard deviation over a defined length, typically 200 m.

As we navigate through the realm of rail transportation, two pivotal concepts; conicity and inclination might be necessary to be included in this section. They have some influences on the challenge of this work and need to be considered.

- Conicity

A regular definition and procedure for measuring wheel conicity—a crucial component of rail vehicle and/or bogie stability on tangent and curved tracks—are provided by the current standard, EN 15302:2008+A1:2010, 2008. Additionally, it offers recommendations for choosing and maintaining wheel and rail running profiles for safe dynamic behaviour.

The concept of "conicity," which is applied while creating wheels for railroad tracks, must be understood to comprehend the wheel's profile. Conicity is a property that characterizes a wheel's propensity to roll in a cone shape. To put it simply, train wheels are conical rather than the typical wheels found on other vehicles, which could be seen in Figure 12, and Figure 13. (Wheel Conicity).

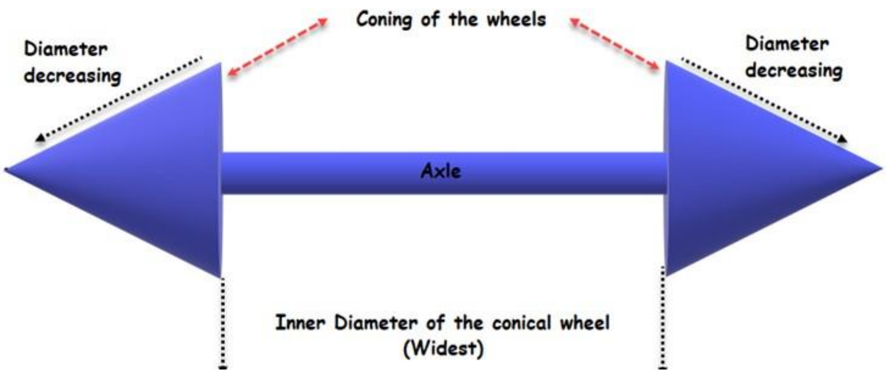


Figure 12 Conical Shape of the wheel

Conicity is used to study how trains and tracks interact dynamically. Based on the behaviour of lateral forces, the flanges are shaped like a cone, with a coning of 1 in 20, this number is the inclination which can be different based on countries, there could be seen rails with different inclinations of 1:30, and 1:40. It consequently influences how well the car steers. The conical wheel positions itself on the track in accordance with the train's steering. The wheels on the track can move freely during bends because to this conical form.

According to the known definitions, the equivalent conicity is equal to tangent tg γ_e of the cone angle γe of the wheelset with conical wheels, transverse movement of which has the same kinematic wavelength as this wheelset (Gerlici et al., 2018).

The equivalent conicity provides an opportunity to compare conditions of the contact of wheels and rails in different states based on the design and maintenance. The linearization approach is used. This approach involves the replacement of nonlinear dependence of the change of the radius of rolling circumference for linear one. The angular coefficient of linear dependence is the equivalent conicity.

It is necessary to adhere to common rules in determining the equivalent conicity in order to able to compare the results obtained for various railways. For this purpose, the International Union of Railways implemented the principles of calculating the equivalent conicity defined by UIC 519 Leaflet (Gerlici et al., 2018)

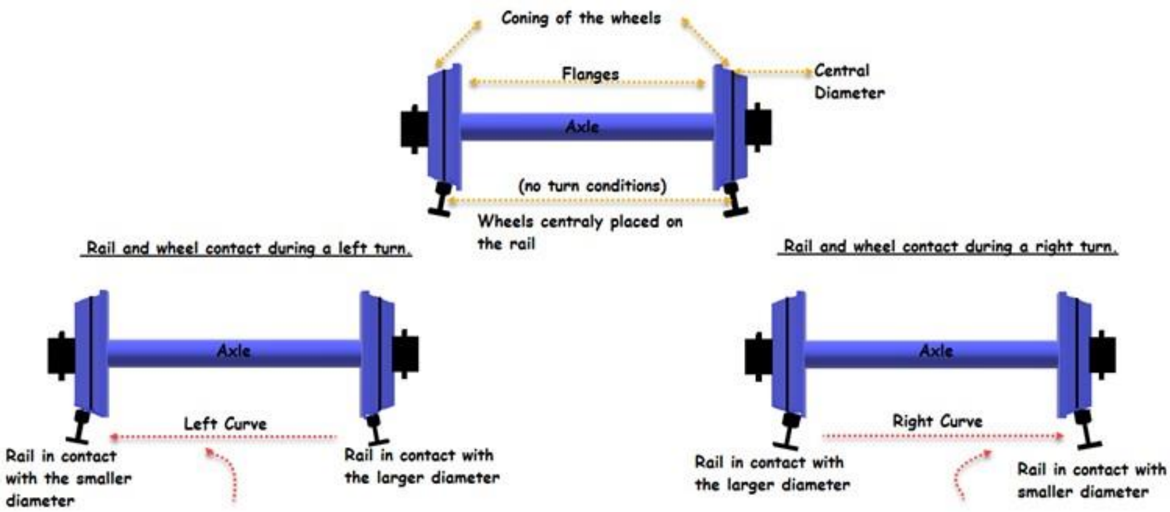


Figure 13 Wheel's movement due to conical shape during turns(Wheel Conicity)

Conicity represents a specific measurement applicable to a single wheelset on the track. Its determination relies on both the lateral movement of the wheelset and the gauge measurement taken at the designated track position. A method derived from a recommendation by (CEN/TC 256/WG10, n.d.) allows for the calculation of conicity across an entire rail segment. The lateral displacement data of vehicle wheelsets on the Gardermobanen line are used for this purpose and could be seen in the following Figure 14.

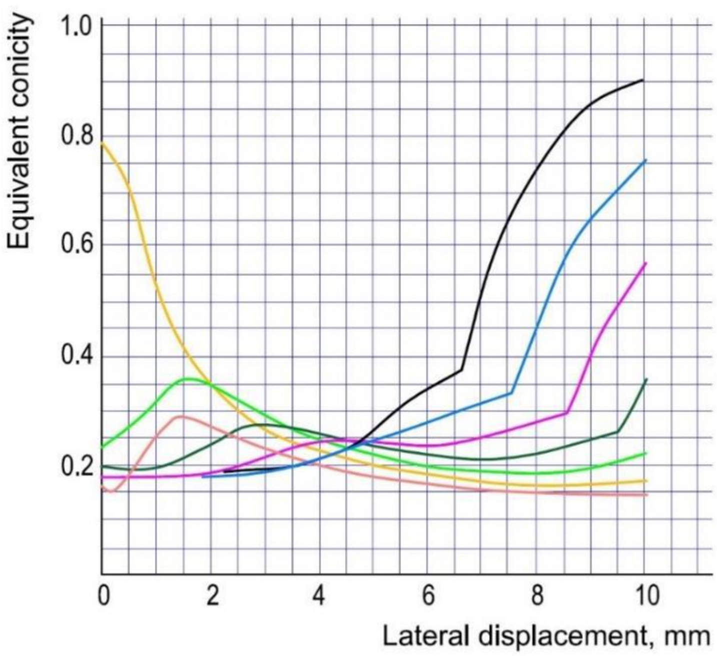


Figure 14 Dependence of wheel equivalent conicity on the lateral displacement (Jönsson & Interfleet, n.d.)

Further details of conicity related to this work will be explained in the following chapter.

❖ Inclination

One well-known and crucial track design element is rail inclination. Given that it influences equivalent conicity, it might have a discernible impact on the running dynamic behaviour of railway vehicles. The inclination of the design rail is unrelated to the lateral alignment. It would not seem to be necessary for assessing alignment at first glance. On the other hand, design inclination influences the vehicle's responsiveness for a specific track alignment. The variable utilized to measure alignment with on-board systems is the vehicle reaction.

Legal documents about rail safety and interoperability, such the Technical Specifications for Interoperability (TSI) of the European Union, employ both inclination and conicity. Here, the TSI pertaining to passenger rolling stock and locomotives is interesting (Commission Regulation of Rail system in the European Union, n.d.). Rail inclination is described as a "basic parameter" in clause 7.1.2.2, which contains guidelines for managing changes in rolling stock. This clause was amended in (Commission Implementing Regulation (EU), 2019). Changes to this clause will require careful study. It is also stated that all rail inclinations (in the EU1:20, 1:30, and 1:40) are considered compatible with rolling stock that has been approved in accordance with the running dynamics assessment standard (EN 14363:2016).

As a definition of rail inclination, shown in Figure 15, the track consists of two rails positioned on sleepers at a specific gauge. In order to generally match the angle γ of the wheelset profile, the rails are laid at an angle β to the sleeper. As the typical response to the contact with the wheel travels through the rail's foot, this helps stabilize the rail against rollover (Ranjha, 2013).

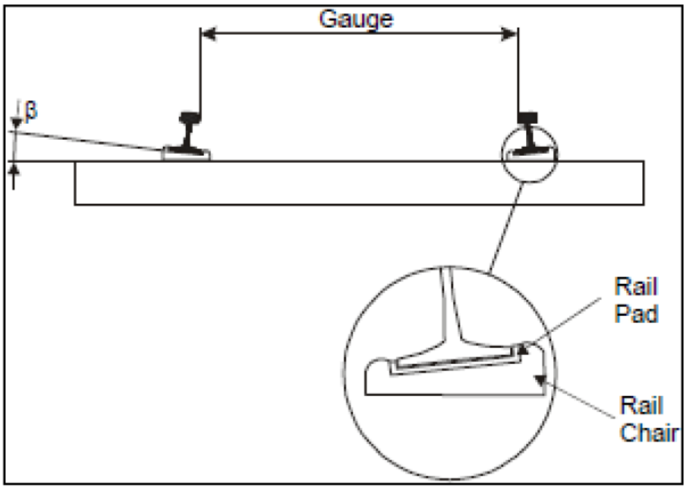


Figure 15 Track Guage and Rail Angle (International Heavy Haul Association, 2001)

As previously mentioned, if conical wheels were to be placed directly on rails arranged perpendicularly, the inherent shape of the wheels would exert a force attempting to push the rails apart. Consequently, the rail fastening system would need to exert greater effort on one side of the rail than the other to prevent the rails from rolling over.

To mitigate the rail spreading effect, rails are frequently laid with a slight inward tilt (inclination) towards the centre of the track. This inclination helps distribute the load evenly through the fasteners, preventing the rails from being pushed apart during operation.

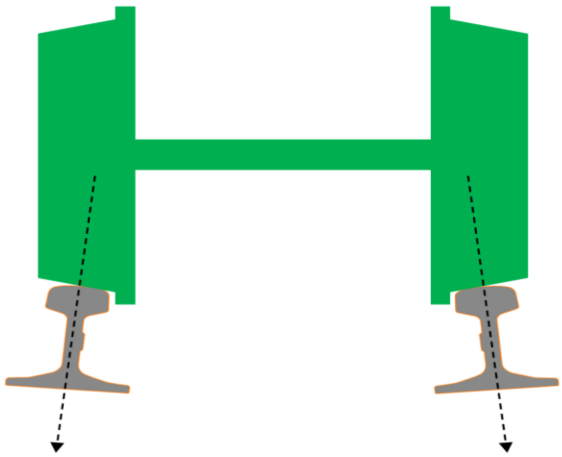


Figure 16 Rail Inclination

There are various common rail inclinations, and the angle of inclination is not the same everywhere in the world. For instance, 1:20 is used in the UK while 1:40 is used in the USA. This is typically taken into consideration when designing the rail sections. Ideally, you don't want the gauge—the separation between the two rails—to alter as the rail ages. To prevent this, rail profiles are made with a taper on the gauge face of the rail and a head shape that ensures the gauge does not alter as the rail wears (Daniel Pyke, 2019).

A BS90R section rail has a parallel head as it was designed to be used without inclination and has a parallel head, which could be seen in the following figures (Daniel Pyke, 2019).

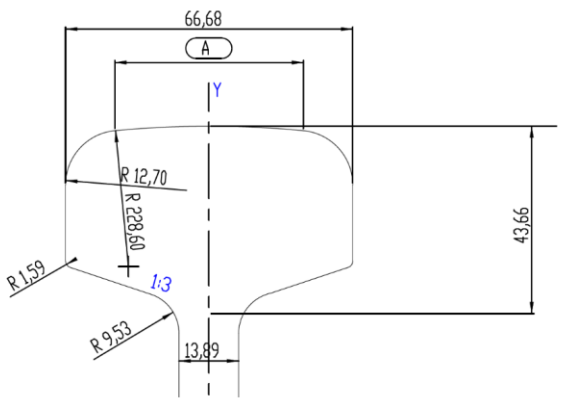


Figure 17 BS90R section rail with a parallel head without Inclination

Following figure shows a BS113A / 56E1 rail which was designed to be laid at an inclination of 1:20 and has a corresponding slope to the rail head.

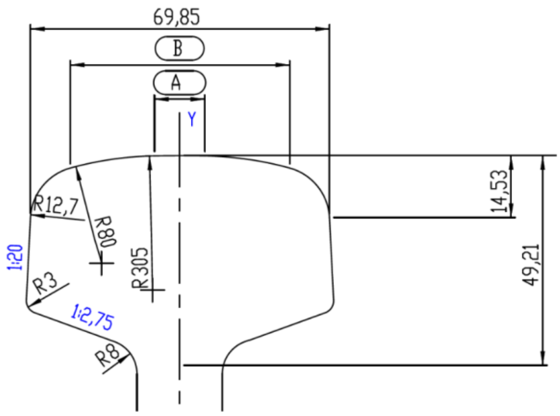


Figure 18 BS113A / 56E1 rail with inclination of 1:20

An AREMA 113RE rail was designed to be laid with a 1:40 inclination.

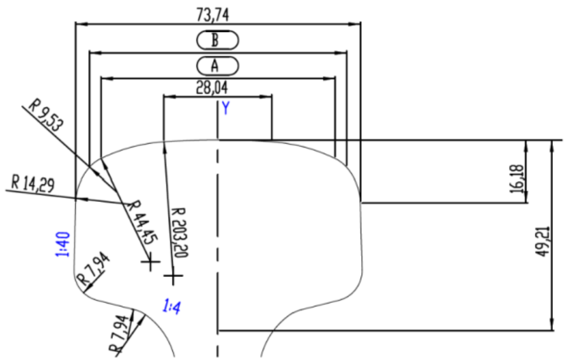


Figure 19 AREMA 113RE rail with 1:40 inclination

❖ Current method of Track geometry measurement process

The track geometry system relies on non-contact optical and inertial technology (Figure 20). Optical sensors are employed to measure both the rail profile and rail positioning, whereas the inertial unit provides data on linear and angular accelerations. By combining information from both optical and inertial sources, it becomes possible to calculate essential rail geometric parameters, including gauge, cant, cross-level, twist, and alignment.

Figure 20 Track geometry system (Mermec System)

The optical measurements are based on the employment of the "light sectioning" approach, which can create a section profile of the railhead from which to extract the points designated "at the top" and "at the gauge" of the rail (Figure 21), helpful to assess the geometric parameter.

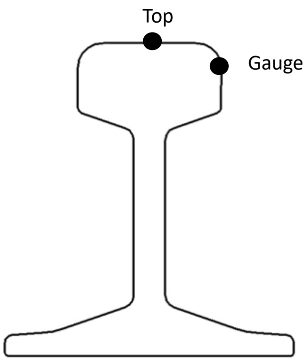


Figure 21 Top and gauge points

Without any mechanical interaction with the rail, the computation is carried out by using the "triangulation" method, which is based on the combined action of a laser and a high-speed camera (Figure 22). The system is made resistant to interference from sunlight or other light sources, allowing it to function in any lighting condition. Each camera is combined with a particular high selectivity optical filter, set on the precise laser emission frequency. ('Kaviani, 2020).

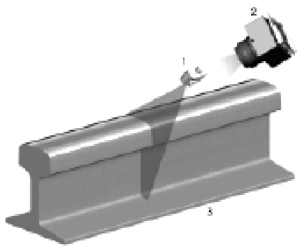


Figure 22 Triangulation Principle (1 laser, 2 camera, 3 rail)

Information obtained from both the optical sensors and the inertial unit is consistently referenced to a shared coordinate system. The optical sensors enable the measurement of distances between the optical boxes and the track, while the inertial sensors establish the position of the optical groups within physical space. Additionally, the system incorporates an inclinometer, ensuring that track geometry measurements are made with reference to an absolute coordinate system, effectively compensating for any movements of the vehicle relative to the rolling plane.

Typically, a track geometry system comprises a minimum of two optical boxes (comprising lasers and cameras) and an inertial unit. This system is installed on the bogie, typically near the axles of the railway vehicle. As mentioned earlier, the track parameters are derived from the combination of data collected by the optical box and the inertial unit. By means of a detailed data processing procedure, the images captured by the camera are transformed into millimetres, allowing for the calculation of points situated "at the top" and "at the gauge" (typically located 14 mm below the rolling plane). These points serve as the foundational data for determining various track parameters.

Data stemming from both optical and inertial sensors are collected in the spatial domain. Ultimately, these track parameters are correlated with information pertaining to the railway line's precise location**.**

The European standard for track measurement systems is EN 13848-2. This European Standard specifies the minimum requirements for measuring principles and systems to produce comparable results. It applies to all measuring equipment fitted on dedicated recording vehicles, or on vehicles specifically modified for the same purpose.

The system needs to be calibrated correctly for accurate measurement data and accurate operation. The inertial unit is usually calibrated in the factory, while the optical boxes are calibrated after they are mounted on the bogie. The optical boxes need to be re-calibrated each time the system is dismounted for vehicle maintenance or when changes are made to the hardware or software throughout the entire system. A manual calliper is also used during the optical calibration process to check for offset errors on the gauge and the cant.

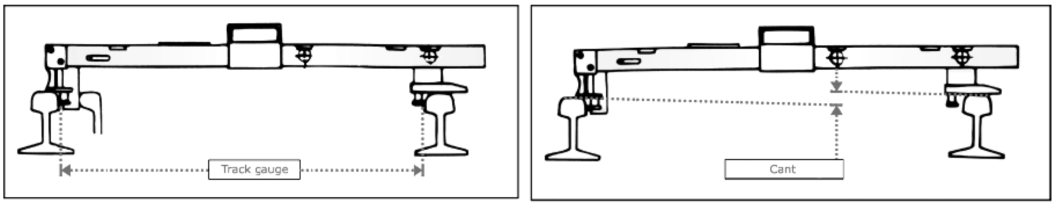


Figure 23 Calibration of a track geometry system

To collect track geometry data for inspection/measurement, infrastructure managers (IM) frequently send out a dedicated track recording vehicle (TRV) or hauled track recording coach (TRC) to travel the rail network (Weston PF, 2015). The measurement method used by a TRV/TRC is an established technology that is standardized in the EN 13848 series. The EN 13848 series addresses several topics related to measurement tools, including measurement methodologies, and track geometry characterization. The parameters of track gauge, longitudinal level, cross level, alignment, and twist are covered. Both an inertial system and a versine system, which is also the measuring concept of TRV/TRC, can be used to measure these parameters. (RIVAS, 2013) examined TRC using the assessment techniques in Switzerland, Sweden, and the UK.

Unattended geometry measuring systems (UGMS) are being installed on in-service trucks in recent years to minimize disruption to regular traffic. To measure the whole track geometrical characteristics and the full rail profile at high speeds, Mermec offers both inertial-based and optical-versine-based UGMS (see Figure 24, Figure 25). UGMS can be mounted on operating vehicles because it is small and lightweight. The 408 km of track of the London Underground was the first metro system in the world to use "unattended" measuring technology (Mermec, 2012).

Figure 24 Mermec inertial based

Figure 25 optical versine based UGMS (Mermec, 2019)

Dedicated TG inspection systems are typically used for the regular evaluation of these parameters. High-quality laser triangulation sensors and/or inertial measurement units (IMU) are frequently used in these systems. However, because these systems' inspection intervals are often set at many months, there is a lack of real-time data regarding track conditions (Soleimanmeigouni et al., 2020).

Track geometry monitoring is something quite distinct. Instead of reconstructing track geometry parameters like inspection/measurement systems do, monitoring attempts to indicate track geometry quality and identify local track geometry abnormalities. Many track condition monitoring systems that do not provide full geometry have been developed, and a few have even been commercialized, despite the desire for UGMS on in-service cars. The relevant monitoring systems employed by IMs as well as in academic and experimental research were evaluated (Entezami et al., 2019) (Bruni et al., 2021;. Weston PF, 2015) .

Suggesting TG monitoring on in-service railway vehicles to enhance available information and facilitate informed maintenance decisions has gained prominence. Over the last two decades, there has been substantial research attention directed towards TG monitoring [(Weston PF, 2015)]. Accelerometers, noted for their affordability and durability, have emerged as the most promising sensors for TG monitoring applications. Robustness is the most important factor for the applications on in-service vehicles, as the monitoring system should not cause additional maintenance, affecting the reliability and availability of the vehicle itself.

The indicators can be defined to show the level of track geometry's quality if the reconstruction of the parameters for the track geometry is not necessary. For example, prEN 13848-1: 2016 suggests computing and analysing mean to peak and/or peak to peak values of axle box accelerations in the specified frequency range that are related to dynamic wheel-rail forces and to isolated faults. Corrugation and/or density of short geometric faults of the rail can be evaluated using standard deviation across a specific distance and frequency range. By comparing the amplitude from the mean value to the peak value or from zero to the peak value with the defined value by the IM, accelerations on bogie and vehicle bodies can also be used to depict individual problems. Perpetuum's wireless sensor network monitors axle box accelerations and establishes the dedicated indications for track condition as an excellent illustration of commercial applications (See Figure 26). The indicator can alert the IM that some track sections require maintenance if it exceeds the calibrated threshold.

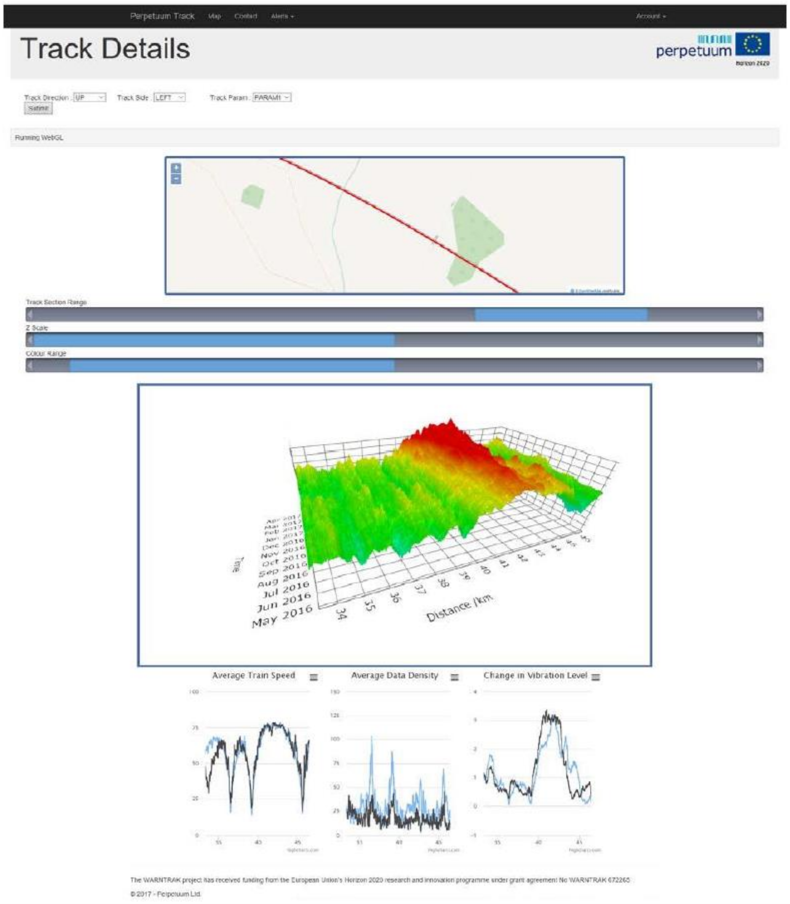


Figure 26 Long-term localised trends in track condition (PER, 2016).

The reconstruction of track geometry characteristics requires the use of particular data handling techniques. For instance, double integrating vertical accelerations can be used to estimate longitudinal level and the vertical rail profile. The most accurate readings are probably obtained by placing accelerometers on the bogie over the axle box and measuring the vertical distance between the accelerometer and the axle box. Although displacement transducers are more precise, they are also more prone to damage and frequently cost more(Weston PF, 2015).

Accelerometers installed on the axle box can be used in conjunction with specific signal processing methods to eliminate the necessity for displacement transducers. RAIDARSS-3, which was placed on shinkansen train sets is an illustration of this strategy (TSUNASHIMA et al., 2012). This illustrates a different method for precisely reconstructing the characteristics of the railway geometry without the additional cost and susceptibility of displacement transducers.

Cross-level and twist can be obtained by sensor fusion of a yaw rate gyro, roll rate gyro, laterally sensing accelerometer, and vehicle speed (Lewis & Richards, 1988).

However, Lateral alignment is hard to measure accurately without optical sensors. Lateral acceleration at the bogie plus lateral displacement from the bogie to the wheelback has been tried, but the displacement component was found to be small and could be done without. The lateral alignment is then the lateral alignment according to the plan view trajectory taken by the wheelsets, which is not the same as the lateral alignment of the track. Determining track lateral alignment and gauge without optical sensors is not practical (Weston PF, 2015).

(Weston PF, 2007) proposed to apply an inverse dynamics model to determine the bogie lateral movement on the rail via bogie yaw rate gyro data or lateral acceleration data in order to estimate the track lateral alignment based on approximate wavelength to be between 35 and 70 m. This method is not simple, though, as it cannot measure alignment at short wavelengths (from 3 to 35 m).

Therefore, monitoring of the longitudinal levelling of track is a straight forward method because there is a strong relation between the vehicle reaction, like vertical acceleration of axle boxes, and the vertical track defects. Hence, monitoring of track geometry with accelerometers is mainly done only for this track geometry parameter like in the ICE2 for several years [DB2003], and previous research has convincingly demonstrated that vertical accelerations can be employed to accurately recover longitudinal level information (R. C. Y. G. et al. Weston PF, 2015). However, due to the inherent dynamics of railroad vehicles, the determination of lateral alignment from lateral accelerations remains a challenge and lacks reliability.

Looking at other parameters of the track geometry such as twist or lateral alignment, which are measured by inspection cars and which are also essential for maintenance and safety issues, it is much more complicated to monitor these parameters with in-service trains using accelerometers.

Monitoring of track geometry with in-service trains is more and more of interest for the infrastructure managers and railway undertakers. As it was mentioned before, in the field of measuring the geometric parameters of the track there are and are commonly applied many sensors suitable for detecting the track geometrical characteristics, applicable to vehicles in normal commercial service. Beside this, an accurate position of the measurements is important to verify the reproducibility and to provide sufficient information for trend analysis to predict the degeneration and to identify the root cause, so, for the location of the measured measurements, inertial platforms (e.g., Velaro in Russia or Mermec) are used, which are generally quite expensive. Therefore, these systems lend themselves to use on special diagnostic trains, rather than to extensive use on the commercial fleet, which is composed of numerous vehicles.

For this reason, this thesis intends to focus on systems that allow the detection of the wheel track position that can avoid the inertial platforms to localize the geometric parameter measures, with the particular focus on the lateral alignment.

To obtain meaningful outcomes in the measurement of wheelset transversal displacement, it is imperative for both equipment designers and manufacturers to possess a comprehensive understanding of the wheelset's interaction with the track and its specific behaviour while the vehicles are in motion on the rails. The unique characteristics of wheelset behaviour, influenced by track conditions and environmental factors, have a notable impact on the accuracy, repeatability, and reproducibility of measurements. Addressing these nuances becomes crucial when establishing parameters for the measuring equipment, determining measurement frequency, and organizing data transmission and collection processes.

❖ Track Maintenance Methodologies and Predictive Models

Maintenance strategies typically fall into two main categories: corrective maintenance and preventive maintenance. Corrective maintenance, also known as reactive maintenance, occurs in response to a defect or failure (Jardine et al., 2006). Traditional preventive maintenance involves scheduling maintenance activities at regular intervals to prevent failures, based on either calendar time or usage.

The primary advantage of preventive maintenance is its ability to be planned in advance and performed at appropriate times (Budai-Balke, 2009). However, traditional preventive maintenance schedules activities regardless of the asset condition, potentially resulting in high maintenance costs and reduced availability (Lu, 2007). To enhance the effectiveness of preventive maintenance, condition-based maintenance strategies are now being implemented.

In a condition-based maintenance strategy, decisions are made based on the actual performance of the asset, determined through condition monitoring, inspection, and testing (CEN 2010, 2010). Predictive maintenance, as a subset of condition-based maintenance, involves scheduling maintenance activities proactively by predicting failure times. According to(CEN, 2010), predictive maintenance is condition-based maintenance undertaken following a prediction based on repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item.

Prognostics and diagnostics can both benefit from condition-based maintenance. While prognostics concentrates on anticipating defects before they occur, diagnostics deals with the detection, isolation, and diagnosis of errors as they occur (Jardine et al., 2006). Prognostic techniques can be broadly categorized into three types: hybrid, physics-based, and data-driven (An, 2015).

Data-driven approaches extract information from training data to identify the characteristics of the current degraded state and predict the future condition of the asset. These approaches can be categorized into artificial intelligence approaches and statistical approaches (Jardine et al., 2006, An, 2015).The prediction performance of statistical approaches depends on factors such as the quality of the model, data quality, and the deterministic, stochastic, or chaotic nature of the degradation process.

Physics-based approaches rely on the assumption that a physical model describes the degradation process, requiring mechanistic knowledge and theory relevant to the monitored asset. Hybrid approaches combine data-driven and physics-based approaches to enhance prediction performance (An, 2015). Prognostic approaches using these methods can predict the remaining useful life (RUL) of the asset, aiding in forecasting maintenance needs.

In an effort to reduce transport disruptions, associated costs, and ensure optimal safety and ride comfort, railway infrastructure managers are striving to shift from a reactive maintenance approach to a proactive one. The evolution of railway track geometry measurement technologies has provided a wealth of event and condition monitoring data, making degradation monitoring more feasible to support a condition-based maintenance strategy. This data not only aids in diagnosis and prognosis of track geometry conditions but also enhances the comprehension of degradation, contributing to improved life cycle management of the track system (Xin et al., 2016). Track geometry degradation models play a crucial role in providing essential insights for decision-making regarding the timing and locations of maintenance activities or inspections. These models must accurately capture the actual degradation process of the track.

The track stands out as a crucial linear asset within railway infrastructure. Unlike point assets, the planning of maintenance activities for linear assets necessitates defining specific points or sections along the asset (Bergquist & Söderholm, 2015). Linear assets' length becomes a critical factor in maintenance planning for this asset category. Given the significance of both location and timing in monitoring the degradation of linear assets, the monitoring scheme requires comprehensive spatiotemporal information (Bergquist & Söderholm, 2015). From a statistical standpoint, condition monitoring data for linear assets typically exhibit spatial autocorrelation. Therefore, to effectively model track geometry degradation, it is imperative to consider the spatial variation in track geometry degradation parameters and potential spatial dependencies.

Final goal of maintenance which is under interest of research, and industry is predictive maintenance, and having an on-board monitoring system could give more frequent geometry measurements which can help to study the temporal evolution of the track faults over years, and this could be a step toward predictive maintenance.

2.2) Statement of the problem

Modelling track geometry degradation over a line section presents several challenges. The primary challenge involves capturing the temporal evolution of track geometry condition while accounting for the impact of maintenance. Existing research suggests that track geometry degradation within a maintenance cycle can be effectively described using linear or exponential models with time or tonnage as the explanatory variable (Esveld, 2001), (Famurewa et al., 2015) (A. R. , T. P. F. Andrade, 2011; Caetano & Teixeira, 2016; Hakan Guler, 2014; Khouzani et al., 2016).

The degradation path consistently exhibits a break point after a maintenance intervention. While maintenance actions, particularly tamping as a primary geometry maintenance activity, enhance the track geometry condition, they cannot fully restore it to a pristine state. Tamping introduces two notable changes in the degradation path: a break point marked by an improvement in geometry condition and an alteration in the degradation rate (Arasteh Khouy et al., 2013; Audley, 2013; Martey & Attoh-Okine, 2018).

To predict the evolution of track geometry over multiple maintenance cycles, it is imperative to assess the impact of tamping on track geometry degradation. The second challenge involves modelling the spatial variation in degradation parameters and addressing potential spatial dependencies. Owing to variations in track structure, substructure, load distribution, and environmental conditions, different track sections experience distinct degradation rates (A. R. Andrade & Teixeira, 2015; Esveld, 2001; Hakan Guler, 2014; Lee et al., 2018). Additionally, there may be spatial correlations between degradation parameters in adjacent track sections (A. R. Andrade & Teixeira, 2013). This correlation could stem from similarities in operational, environmental, and structural conditions among adjacent track sections.

Another challenge involves modelling the occurrence of isolated track geometry defects. The Track Quality Index (TQI), which relies on the standard deviation of geometry parameters, may not offer complete information about severe isolated defects in a track section (Alemazkoor et al., 2018). Isolated defects are brief irregularities in track geometry that can significantly elevate dynamic forces between the wheel and rail, potentially accelerating the growth or occurrence of internal rail defects. The presence of severe isolated defects can lead to passenger discomfort, damage to track components, increased derailment risks, and unplanned maintenance activities. This, in turn, may result in operational interruptions, reduced track availability, and higher maintenance costs (A. R. Andrade & Teixeira, 2018).

Predicting the occurrence of severe isolated defects provides railway infrastructure managers with the means to more efficiently maintain railway tracks while staying within ride comfort and safety limits (Arasteh Khouy et al., 2013; Cárdenas-Gallo, I., C.A., et al., 2017; He et al., n.d.; Sharma et al., 2018) Continuous monitoring of track geometry conditions is essential for predicting when maintenance limits will be reached, enabling effective maintenance planning (UIC, 2008). The inspection frequency plays a crucial role, influencing the detectability of geometry defects, maintenance costs, and the risk of train derailment. Thus, infrastructure managers face the important task of selecting the most suitable inspection interval for track geometry measurement vehicles.

Summarizing the key points discussed earlier, the primary objective is transitioning from diagnostic to prognostic monitoring. This shift necessitates more frequent monitoring, ideally achieved through on-board monitoring, as opposed to relying solely on diagnostic trains.

A significant challenge we face in implementing an on-board monitoring system without an inertial platform is the absence of a well-established relationship between lateral track irregularities and lateral acceleration. As a result, it is currently impossible to detect track irregularities only from accelerometer data, whereas the process is straightforward for the vertical direction. This stresses the complexity and innovation required to bridge this gap and develop our lateral track monitoring capabilities.

The research object is to discover the specific issues of wheelset lateral displacement measurement on the track. The main purpose of this measurement is to allow the track geometry measurement system to correctly measure the geometric parameters, avoiding the possible measurement error due to the transverse displacements of the wheelset with respect to the track. In particular, it is essential to correctly measure the lateral alignment.

When developing devices for the measurement of the transversal displacement of the wheelset, it is necessary to study and evaluate the features of the wheelset movement on the track. When a vehicle is running along a track, the wheelsets move not only straight along the longitudinal axis of the track but also transversely along the lateral axis (Figure 27). This is the so-called Klingel motion of the wheelset, which is a sinusoidal movement.

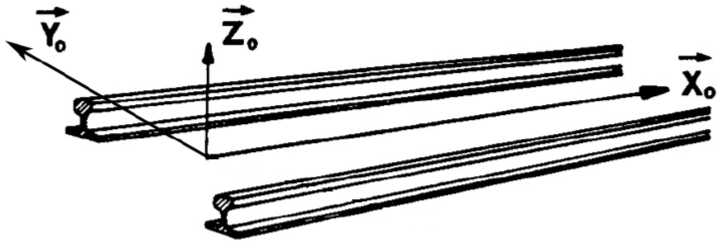


Figure 27 The commonly used axis for considering the track geometry

It should be noted that the wheels of a conventional vehicle are rigidly mounted on the axle so that their angular velocity is always the same when the wheels are rotating.

On the other hand, the rolling surfaces of wheelsets are usually made of a conical profile.

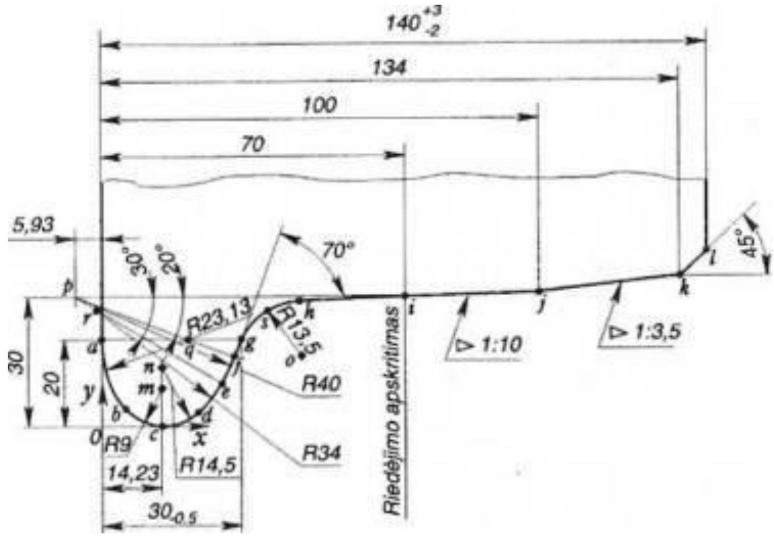


Figure 28 Example of wheel rolling profile

The purpose of wheel geometrical design is to allow wheelsets to centre themselves on the track in operation. If the running surface of the wheelsets were cylindrical, the rolling surface of the wheels would constantly scrub against the rails on one side of the track under any external forces (lateral wind, lateral inclination, etc.). This would cause intense abrasion, overheating and excessive noise.

If the wheel is not positioned exactly ("ideally") in the centre the track, one wheel is running with a greater rolling radius than the other wheel, due to the conicity of the wheel. The difference between rolling radii and linear speeds of moving wheels are presented in Figure 29.

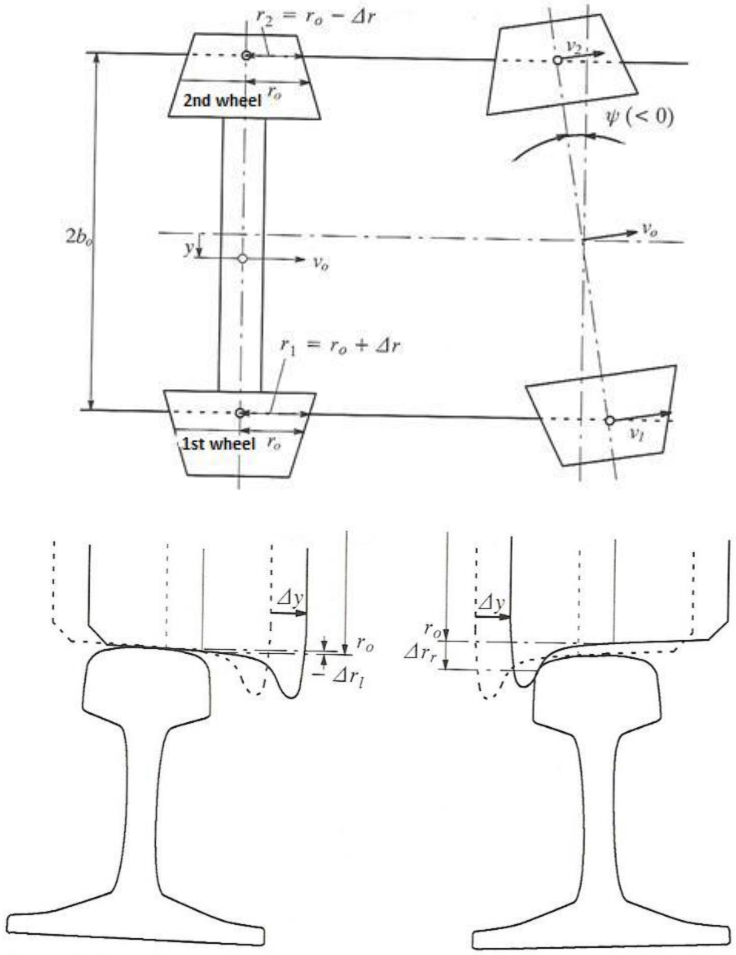
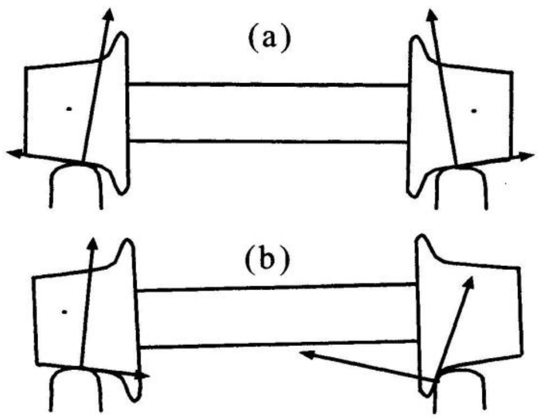


Figure 29 The difference of radii of moving wheels and linear speeds

As the rolling radii of the wheels are not equal, their linear speeds are changing. As shown in Figure 30., the linear speed of the 1st wheel is higher than the linear speed of the 2nd wheel. Due to the difference in the linear speeds of the wheels, they alternately (cyclically) overtake each other, by moving in the transverse direction. The wheel repeatedly moves on the alternating rolling radius.



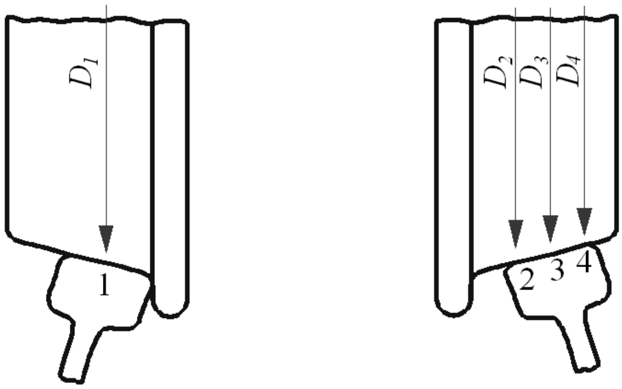


Figure 30 Scheme of a wheel rolling radius alteration due to wheelset transversal moving in the track

To have a closer look on this issue, we could refer to Figure 31. (a) It can be shown that in the horizontal plane, at point P on each rail, which is located 14 mm below the top of the railhead for the standard rail profile UIC 60E1, lateral alignment is defined as the lateral deviation between the actual and reference rail lines (EN 13848-1, 2019). For lateral alignment to be estimated by accelerations, it is expected that the vehicle wheels would follow the impulse of lateral alignment in the lateral direction. The wheels do not, however, follow lateral alignment in the same manner as vertical alignment because the wheel has a freedom of movement in the lateral direction in a clearance between the wheel flange and rail head edge, shown in Figure. 31 (b). Therefore, the problem is to find a solution that could assess detecting the lateral irregularity, which in here could be detecting the lateral displacement of wheel to rail.

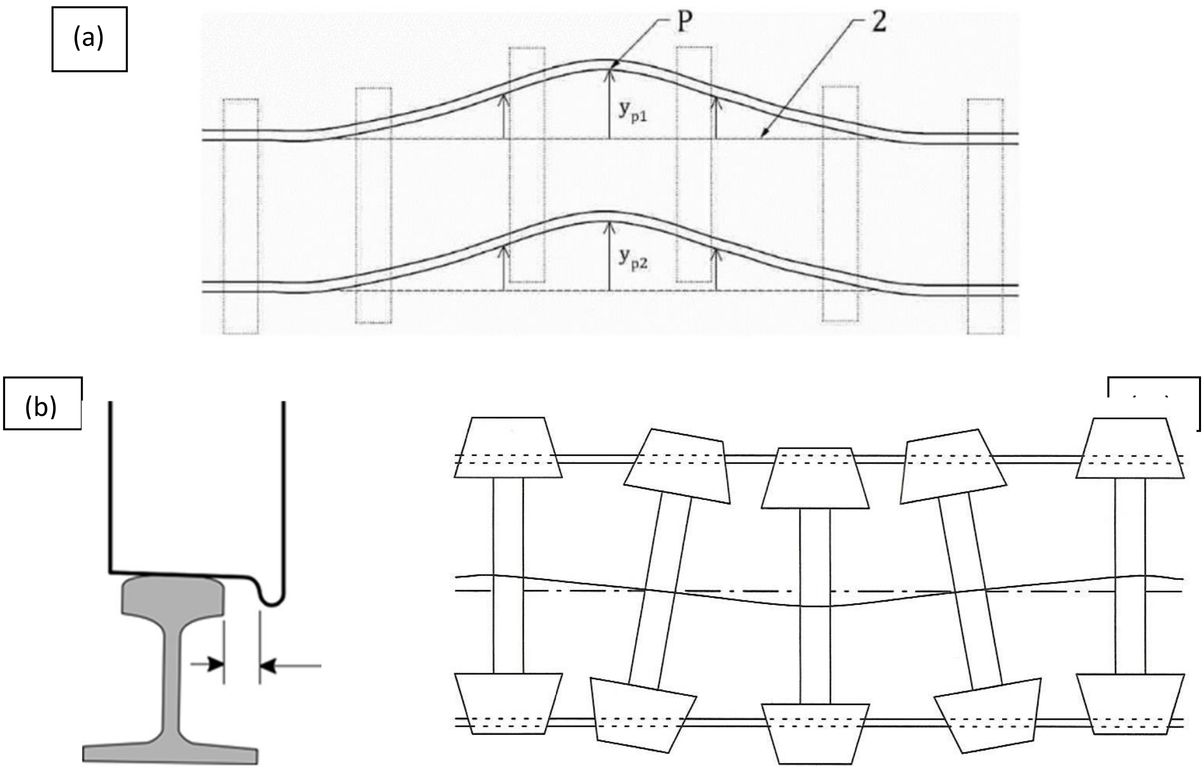


Figure 31 (a) lateral alignment of the left rail y_(P_1) and right rail y_(p_2), where P denotes the rail reference point and "2" denotes the reference rail line (EN 13848-1, 2019); (b) illustration of wheel/rail gauge clearance(L. Rizzetto et al., 2019) (c) sc

The measurements conducted previously were in a static setting, with the camera system positioned alongside the railway. However, the challenge which is tackling here is considerably more complex. The focus is on achieving real-time, target-less dynamic displacement measurements against a backdrop of fluctuating and noisy environments. Specifically, within the railway context, the goal is to monitor the lateral movement of a railway vehicle's wheels in real-time during regular railway operations. This endeavour addresses a previously unresolved issue in the railway domain, specifically related to the monitoring of track geometry (TG).

After solving the on-board monitoring issue, the main objective is to find a Machine Learning model to detect the lateral track irregularity by on-board monitoring and find the pattern for predictive maintenance.

2.3) Related Work (Master's Thesis Contribution)

In response to the initial challenge of measuring lateral irregularities, investigations have been made to develop specialized sensors that can be mounted on commercial trains to monitor track geometry parameters. Various technologies were explored and tested in the previous work during my master's thesis investigation, which was conducted within the framework of the Assets4Rail project. This research aimed to find innovative solutions to identify the lateral alignment through lateral displacement of the wheel in relation to the rail. (Kaviani, 2020).

Some candidate technologies have been examined based both on direct measurements:

- lasers;
- high speed cameras;
- stereo cameras;
- thermographic cameras;

And on indirect measurements:

- accelerations;
- ultrasonic reflection.

Stereo cameras and thermographic cameras seem to be the most promising technologies to be taken into consideration for subsequent developments.

In the two experiments described above, thermographic cameras have already been used, giving good results, even if not for all the speeds and atmospheric conditions that may occur during the normal operation of a commercial train.

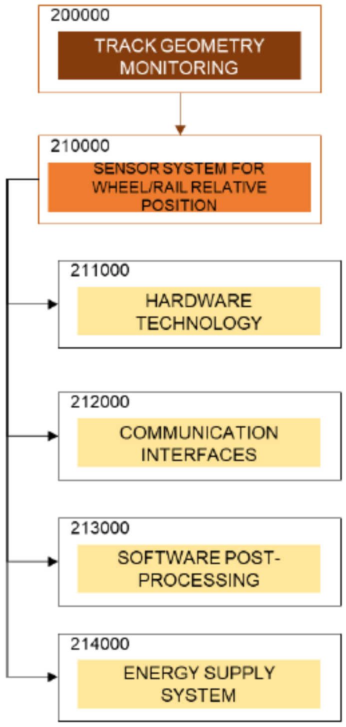


Figure 32 Track Geometry Monitoring Structure

Within the work of Assets4Rail project, several requirements were discovered and divided into four main categories: functional, operational, performance, and safety requirements. ('Kaviani, 2020). The main focus is on building the sensor system for the detection of wheel/rail transversal position within the broader context of track geometry monitoring, unlike some projects that involve splitting down the main system into various subsystems. Therefore, the split of the project into subsystems is not required.

And the sensor system for detection of transversal position could be divided into:

- Hardware technology: that specifically is the camera and supporting system to be mounted on the bogie.
- Communication interfaces: because the data that acquired by the sensor system should be synchronized with whole monitoring system.
- Software post-processing: images which are acquired by devices should be post processed because we will have an image and we must develop an algorithm to detect the position of point.
- Energy supply system: finally, we have to supply energy to the system, so there is a need of providing an energy supply system that could utilize the energy that is already on board (on wagon), or the final goal was to develop a system that provides energy by itself. For instance, the harvesting system that could utilize the vibration of the bogie, etc.

The results obtained in the Assests4Rail project were evaluated using TOPSIS (The Technique for Order of Preference by Similarity to Ideal Solution). It defines alternatives by the distances between the indicators' best and worst values. If the result equals one, it is the best solution, if the result equals zero, it is the worst solution.

Using the estimated values of the criteria and their weights, applying the AHP (Analytic Hierarchy Process) and Direct evaluation methods, and evaluating the selected technologies was performed. The TOPSIS method was applied and the results are presented in Table 8. The technologies' ranking is similar using the two different calculations methods; only the ranks 7 and 8 were identified differently. It has no considerable significance since only the best one is needed. (Kaviani et al,2020)

Table 8 Measurement Technology Ranking

	Technologies							
	Reference technology Plasser Optical Gage Measuring System	Laser profile measurements	Distance sensor laser, time of flight	Time of flight camera	Digital Image Correlation (DIC) of high-speed filming	Laser based systems, displacement	Thermal vision	Stereo vision
TOPSIS DE	0,816	0,761	0,830	0,730	0,789	0,661	**0,847**	**0,879**
Ranking	4	6	3	7	5	8	**2**	**1**
TOPSIS AHS	0,886	0,808	0,888	0,655	0,863	0,801	**0,910**	**0,920**
AHP	4	6	3	8	5	7	**2**	**1**

It was found out that technology based on stereo vision is the best for this specific task. The technology-based on thermal vision came in second place. Therefore, it was decided to conduct the experimentation by investigating two sensors representative of these two technologies through laboratory and on-track testing.

❖ Test platform for real-time testing of sensor solutions

During the experimental investigation, two technologies were under review:

a) Based on stereo camera (the Zed stereo camera developed by STEREOLABS);
b) Based on the thermal imaging camera (the T420 thermal imaging camera developed by FLIR).

Figure 33 Sensors under investigation: Stereo camera ZED by STEREOLABS (left image) and Thermal camera T420 by FLIR (right image)

Algorithms were created to successfully extract the required geometrical information from photos. These algorithms are flexible and work with the two technologies that have been chosen. Furthermore, a system was developed for measuring measurement accuracy, and it was found that it is 0.7 mm for the stereo camera and 1.4 mm for the thermal imaging camera. This methodology can be used to evaluate the measurement precision of various devices that system developers may later take into consideration.

A test platform was developed that is first housed in the facilities laboratory at VGTU in order to permit real-time testing of the sensing systems. This platform includes a test gear and cameras connected to Nvidia Jetson TX2 embedded AI processing chips that are power-efficient. Utilizing experimental data that was gathered in the field by Lithuanian Railways, the data processing technique was validated. It is important to note that the cameras and other measurement equipment have a combined maximum power consumption of less than 20W for both technologies.(Kaviani et al.,2020.)

It has also been checked that data handling and transmission are flexible. Data can be directly stored on the hard drive built into the Nvidia Jetson TX2 device or sent to the data warehouse over Ethernet or Wi-Fi. We've created the algorithms with adaptability in mind throughout the task implementation, allowing for a freely selected output data format as per the particular requirements of the application.

2.4) Challenges and Contributions

In this railway dynamic displacement measurement application, there has been several significant challenges.

Firstly, this task centres around continuous monitoring rather than periodic inspections. Monitoring devices are designed for widespread deployment and seamless automation during operations. Consequently, monitoring devices with high automation capabilities and cost-effectiveness is required.

Secondly, our computer vision (CV) system is mounted on the railway vehicle, focusing on a wheel in motion along the railway track. Attaching an optical target to the rotating wheel is unfeasible. Traditional target-less methods like edge detection, pattern matching, and line detection often suffer in performance, especially in the presence of intricate backgrounds. Railway tracks are accompanied by complex textures such as ballast, sleepers, and varying plant cover, making it challenging to achieve accurate measurements.

Thirdly, a crucial requirement is real-time image processing. Timely processing is essential as the calculated LDWR (lateral dynamic wheel-rail) must be swiftly integrated with acceleration measurements to reconstruct track lateral alignment.

In response to these challenges, an innovative solution involving virtual point tracking was proposed and developed by contribution of project partners in the Technical University of Berlin. To the best of our knowledge, this work represents the first attempt to combine HPE and domain knowledge for displacement measurement.

Furthermore, during the initial real-life testing endeavour, unforeseen delays arose from a comprehensive risk analysis. Unfortunately, this resulted in the absence of the accelerometer on-board during the critical analysis phase, rendering the data and analysis less than fully sufficient. Moreover, the ongoing COVID-19 pandemic disrupted plans to re-do the test.

In response to these challenges, simulations in collaboration with the Norwegian University of Science and Technology (NTNU) was done to complete the measurement. Subsequently, I was able to validate the findings through comprehensive confirmatory testing at the UK Rail and Innovation Research Centre with a similar measurement technology. This phase involved on-board monitoring tests, confirming the reliability and accuracy of our methodology and results.

The contribution of different research activities and institutes of the work could be seen in the following Table 9.

Table 9 Contribution of different research activities on this work

Background research	Ph.D. Activities	National Italian Project "MOST Spoke" 4 T3.1.5, T3.2.4
• Master's thesis within Assets4Rail for Selection of the sensor for on-board monitoring system • Hardware development of the on-board monitoring system • Verification of the system	• Assets4Rail continuation for Field tests with computer-vision system, Image processing algorithm developed by TUB. • Test validation, reporting. • NTNU Collaboration for • Multibody Simulation and • ML algorithm testing on virtual measuremen. • BCRRE Collaboration using on-board monitoring data on a high speed line for Prediction.	• This Ph.D. research has been selected for a sub-task of Most Project for Using the developed ML model and data for predictive Maintenance.

2.5) Structure of the thesis

The thesis follows the structure outlined below. Section 3, titled "Methodology," provides a brief overview of the hardware design, mounting system, and the development of the image processing algorithm. This section also encompasses simulations aimed at completing the data collection process and validating the reliability of the proposed idea. Subsequently, Section 4 delves into data analysis and visualizations aimed at identifying the most suitable algorithm for the intended purpose. The proof of concept is then presented through the implementation of the idea in a real-world experiment. The thesis concludes with a summary in Section 5, highlighting the main results of the study within a short context.

The research methodology scheme could be seen in the Figure 34.

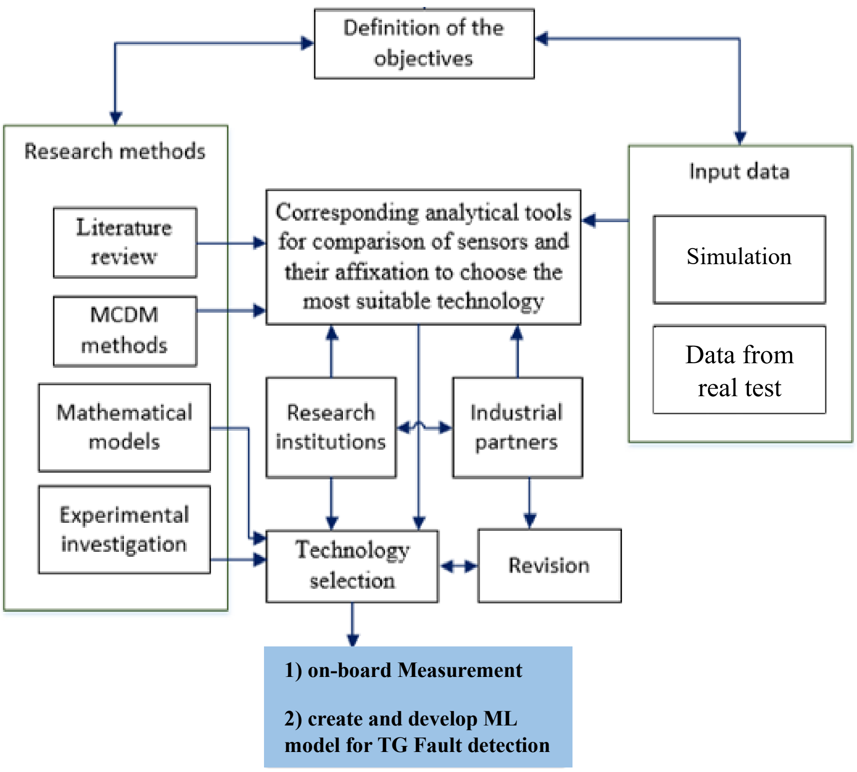


Figure 34 Research Methodology

3) Methodology

The methodology is divided into different steps related to hardware development, Data collection through real test, Complementation of the data analysis based on simulations, and finally developing a machine learning model to detect and predict the lateral irregularities. Afterwards, a data analysis through real on-board monitoring test will be done to confirm the focused issue and solution.

The subsequent sections determine the content covered in the papers referenced at the beginning of this thesis. Sections 3.1 to 3.3 shows the findings of Papers 1 and 2, explaining the collaborative efforts between Sapienza and the Technical University of Berlin in developing the image processing algorithm, along with subsequent validation tests. Then, the initial simulation to assess the methodology (Paper 3) is shown, and subsequently, the remaining facets of the methodology will be discussed in Papers 4, 5, and 6, conducted at NTNU during the exchange period. Another real test to evaluate the whole methodology done at BCRRE would be discussed as well.

3.1) Hardware components of the monitoring system

The proposed monitoring system consists of an off-the-shelf stereo camera, an air cleaning system, a processing unit, a lighting system, and a mounting system with the dampers. The air cleaning system aims to clean the camera lense by blowing the compressed air regularly. This is a standard solution to avoid the dirt in the optical systems (Weston PF, 2015). From the software perspective, we enhance the robustness of the algorithm against the image's visual corruptions. This will be introduced in the Section of experiments and results. For optical sensing, ZED2 stereo camera is used in our system(Stereolabs, 2019) , which is configured to output videos with the resolution of 1920×1080 pixels at the sample rate of 30 frames per second (fps). Any comparable cameras can be used as well. The depth information is merely used for displacement measurement. The algorithms described in this are directly applicable for 2D images obtained by regular CCD cameras. As the processing unit, Nvidia Jetson Tx2 has 256 core NVIDIA Pascal architecture and ARMv8 6 core multi-processor CPU complex, enabling real-time execution of DL models (NVIDIA, n.d.). The mounting system consisting of the vibration dampers, a crossbar and a clamp can be easily installed on different bogie types. The mounting system consists of the vibration dampers, a crossbar, a clamp and camera housing which is equipped with an external lighting system, consisting of a series of LEDs. The entire system is installed on the bottom of the bogie frame, facing the wheel, as shown in Figure 35. Two systems are required to monitor the wheel-rail pair on the left and right sides simultaneously. In the current hardware implementation, the cleaning system is not included. The processing unit is inside the vehicle cabin, connecting to and powering the camera. The hardware of the monitoring system will be further improved for long term monitoring.

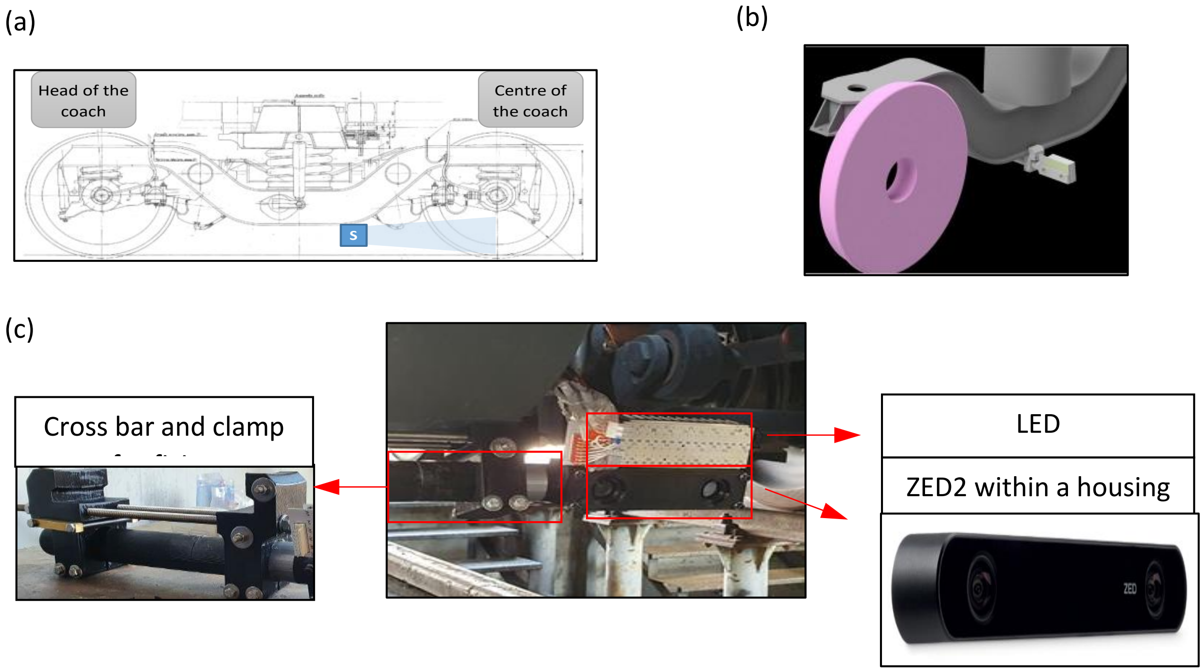


Figure 35 (a) drawing of the camera position on the bogie frame; (b) CAD model of the camera installation; (c) monitoring system

3.2) Approach for virtual point tracking

The attempt of formulating the task of dynamic displacement measurement to track virtual reference points and calculate the distance between two virtual reference points started with technical university of Berlin. Three reference points on the wheel were defined (P_W) and rail (P_{r_1} and P_{r_2}) respectively. P_{r_1} refers to the reference point P for lateral alignment (European Committee for Standardization (CEN), 2019), as introduced in the background. P_{r_2} is the symmetry point of P_{r_1} on the other side of the railhead edge. The distance between P_{r_1} and P_{r_2} is the width of the railhead. The lateral displacement D, of the wheel on the rail (LDWR) is represented by the lateral distance between P_w and P_{r_1}, see Figure 36. The relative lateral motion of the wheel is represented by the changes of D (i.e., ΔD) over time, which is the output of the monitoring system. The point P_{r_2} is defined for tracking mechanism, which is explained in the section of (On-line point tracking by a rule engine).

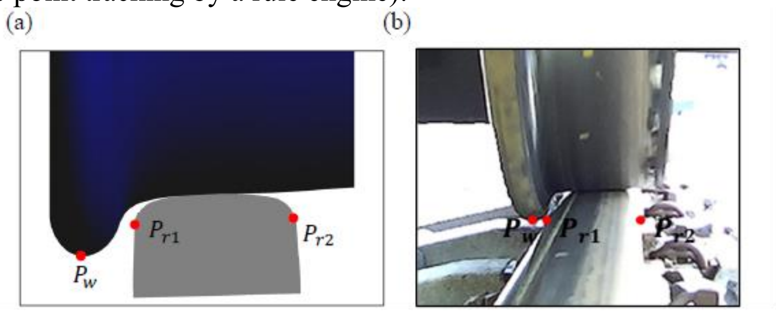


Figure 36 (a) defined key points , P_W, P_{r_1} and P_{r_2} illustrated in the animation; (b) marked in the real photo (right)

Virtual point tracking consists of three steps, as shown in Figure 37 The first step is the calibration, executed for the first-time installation. This calibration process detects RoI, which refers to the wheel-rail contact area. The outputs are the coordinates of the centre point of RoI. Moreover, the distance between the camera and the wheel is obtained as the stereo camera's depth. The distance is an input parameter for displacement calculation. In the case of using a CCD camera, the distance has to be manually measured. The next steps are executed to detect and track virtual points in real time. Next, each step will be introduced in detail (Shi et al., 2022)

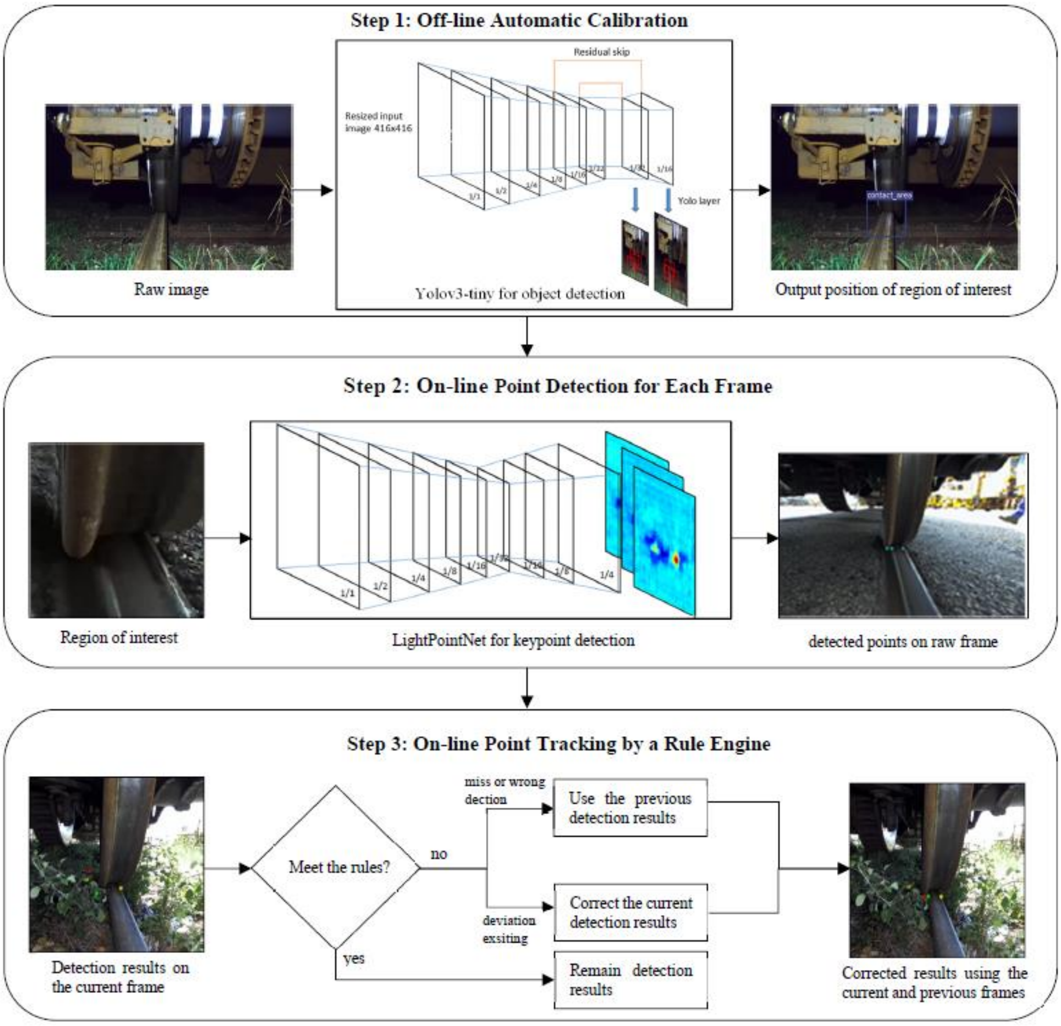


Figure 37 Approach to track virtual points on the wheel and rail

3.2.1) *Step 1: Off-line automatic calibration*

As the resolution of each frame is 1920×1080 pixels, it is necessary to resize the image prior to feeding it to CNN. However, resizing and restoring of the image cause the additional measurement error for point detection. To avoid the step of image resizing, cropping the RoI from the raw image was proposed. A mature object detection technique based YOLOv3 was chosen (J. Redmon, 2018), which is a CNN architecture and has been widely deployed for diverse applications. A modified version of YOLOv3 for RoI detection was adopted, called YOLOv3-tiny. The architecture of YOLOv3-tiny is shown in Figure 37 and Table 10. The first 13 layers are used for feature extraction, known as Darknet.

The input for Darknet is the images with 416×416 pixels downsized from the original 1920×1080 ones. The output 1024 feature maps of Darknet has the dimension of 13 × 13 pixels. The layers 13-16 are to make predictions. YOLOv3-tiny formulates object detection as the regression of the coordinates and sizes of the objects' bounding boxes. Each bounding box has four positioning parameters to be predicted, i.e. x and y coordinate of the centre point as well as its width and height.

Apart from this, an objectness score is calculated according to the overlap area between the predicted bounding box and the ground truth. Each bounding box also predicts the classes contained within this box. In our case, there is only one class, namely the wheel-rail contact area, while others are backgrounds. That means, 6 parameters will be predicted for each bounding box. YOLOv3-tiny pre-defines three boxes with different sizes for prediction, commonly called anchor box. Therefore, 18 parameters, i.e. 3 × (4 + 1 +1), will be regressed. In the 15th layer, there are 18 corresponding feature maps. Each feature map of 13 × 13 is used to regress one parameter. The 16th layer compares the prediction with the ground truth to calculate the loss. The loss consists of classification loss, localization loss and confidence loss. The detailed loss functions can be found in the original paper [31]. YOLOv3-tiny predicts at two different scales. The first scale uses the aforementioned 13 × 13 feature maps, while the second scale uses 26 × 26 feature maps, implemented in the layers 17-23. The 17th layer is a shortcut to the 13th layer. The 13 × 13 feature maps within the 13th layer are upsampled to 26 × 26 and fused with the 26 × 26 feature maps within the 8th layer. The 21-23 layers have an identical function as the 14-16 layers in the first scale. The outputs of YOLOv3-tiny are the candidates of RoI with the dimension of $N \times 18$, where N denotes the number of the candidates. The final prediction is selected by objectness score thresholding and non-maximal suppression(J. Redmon, 2018)

Table 10 Adapted YOLVO3-tiny architecture

Layer Index	Type	Filters	Size/Stride	Output
0	Convolutional	16	3 × 3/1	416 × 416
1	Maxpool		2 × 2/2	208 × 208
2	Convolutional	32	3 × 3/1	208 × 208
3	Maxpool		2 × 2/2	104 × 104
4	Convolutional	64	3 × 3/1	104 × 104
5	Maxpool		2 × 2/2	52 × 52
6	Convolutional	128	3 × 3/1	52 × 52
7	Maxpool		2 × 2/2	26 × 26
8	Convolutional	256	3 × 3/1	26 × 26
9	Maxpool		2 × 2/2	13 × 13
10	Convolutional	512	3 × 3/1	13 × 13
11	Maxpool		2 × 2/1	13 × 13
12	Convolutional	1024	3 × 3/1	13 × 13
13	Convolutional	256	1 × 1/1	13 × 13
14	Convolutional	512	3 × 3/1	13 × 13
15	Convolutional	18	1× 1/1	13 × 13
16	YOLO			
17	Route 13			
18	Convolutional	128	1× 1/1	13 × 13
19	Upsampling		2 × 2/1	26 × 26
20	Route 19, 8			
21	Convolutional	256	3 × 3/1	26 × 26
22	Convolutional	18	1× 1/1	26 × 26
23	YOLO			

3.2.2) *Step 2: On-line point detection for each frame*

Point detection is an essential step in our approach. Inspired by networks for HPE, LightPointNet was proposed, a lightweight CNN architecture for real-time point detection on each video frame. LightPointNet consists of an encoder for hierarchical feature extraction and a decoder for heatmap prediction.

Inspired by MobileNetV3 (Howard et al., 2019) and PoseResNet (B. Xiao & H. Wu, 2018), the key insights behind LightPointNet are the lightweight backbone and the straightforward encoder-decoder structure. These measures reduce the computational complexity of the model with little degradation of the model performance.

Table 11 LightPointNet architecture

Block	Type	NL	SE	Exp size	Filters	Size/Stride	Output
0	Conv	HS	false	-	16	3 × 3/2	128 × 128 × 16
1	Bneck	RE	false	16	16	3 × 3/2	64 × 64 × 16
2	Bneck	RE	true	72	24	3 × 3/2	32 × 32 × 24
3	Bneck	RE	false	88	24	3 × 3/1	32 × 32 × 24
4	Bneck	HS	false	96	40	5 × 5/2	16 × 16 × 40
5	Bneck	HS	true	240	40	5 × 5/1	16 × 16 × 40
6	Bneck	HS	true	240	40	5 × 5/1	16 × 16 × 40
7	Bneck	HS	true	120	48	5 × 5/1	16 × 16 × 48
8	Bneck	HS	true	144	48	5 × 5/1	16 × 16 × 48
9	Bneck	HS	true	192	64	5 × 5/2	8 × 8 × 64
10	Bneck	HS	true	384	64	5 × 5/1	8 × 8 × 64
11	Bneck	HS	true	384	64	5 × 5/1	8 × 8 × 64
12	ConvTranspose	RE	false	-	256	4 × 4/2	16 × 16 × 256
13	ConvTranspose	RE	false	-	256	4 × 4/2	32 × 32 × 256
14	ConvTranspose	RE	false	-	256	4 × 4/2	64 × 64 × 256
15	Conv	-	false	-	3	1 × 1/1	64 × 64 × 3

The architecture of LightPointNet is shown in Figure 37 and Table 11. The first 12 blocks build the encoder, while the last four blocks build the decoder. The whole network is built by stacking three building blocks.

The first block “Conv” refers to a convolutional block, consisting of a convolutional layer, a batch normalization layer and the hard-swish function (HS) as the activation function for nonlinearity (NL). In this block, 16 convolutional filters parameterized by the weights $w \in R^{3\times3}$ are performed on the input image I $\in R^{256\times256\times3}$ to generate the feature map F $\in R^{128\times128\times16}$. Then, mini-batch normalization (S. Ioffe & C. Szegedy, 2015) and hard-swish (Howard et al., 2019) are performed on the feature map F to reduce internal covariate shift and add nonlinearity. The swish function aims to solve the dead neuron problem of ReLu, which is the most common activation function for CNN. The hard version of the swish function reduces the computational complexity of the original one, defined as:

$$Hswish = x \cdot (ReLu6(x + 3)) / 6$$

$$ReLu6 = min(max(0, x), 6)$$

The convolutional block is followed by 11 blocks of inverted residual and linear bottleneck (Bneck) (Howard et al., 2019). Bneck is a modified version of the original residual operation (X. Zhang et al., 2016), which enables the skip connection between the input and output feature maps by following a wide-narrow-wide bottleneck structure in terms of the channel number. Bneck uses an inverted bottleneck with a narrow-wide-narrow structure. It is implemented by stacking three convolutional layers. The first one is 1×1 pointwise convolution to expand the input channel dimension c by a factor e, followed by an activation function. The expanded size for each

Bneck block is given in the column "Exp size" of Table 11. The second one is 3×3 depthwise convolution with an activation function, keeping the channel dimension unchanged. Replacing regular convolution with depth wise convolution is an effective lightweight measure. This will be further described in Section (Computational complexity and real-time capability). The third one is 1×1 pointwise convolution to recover the output channel dimension to c , allowing the identity skip connection between the block inputs and outputs. The third convolution layer does not involve an activation function and thus remains linear. Bneck can be combined with Squeeze-and-Excite (SE) (Hu et al., 2017), which improves channel interdependencies of feature maps. The column "SE" indicates the presence of the SE module. Concretely, SE consists of a global average pooling (GAP), a fully connected layer (FCL) with the ReLu activation function, and a FCL with the hard sigmoid activation function. GAP squeezes the input feature map F $\in R^{D \times D \times C}$ to output its spatial average U $\in R^{1 \times 1 \times C}$. FCLs force the interchange of information among the c channels and output the channel-wise weighting factor S $\in R^{1 \times 1 \times C}$, given by:

$$S = HSig(W_2 \cdot ReLu\,(W_1 \cdot U))$$

where HSig denotes the hard sigmoid function and defined as ReLu6(x + 3), , $W_1 \in R^{c \times c/r}$ denotes the learnable weight parameter of the first FCL, $W_2 \in R^{C/r \times c}$ and r denotes the reduction ratio and set to 16 as suggested by the original paper (Hu et al., 2017).

The Bneck blocks extract hierarchical features and downsize the feature maps to 8×8. Afterwards, 3 blocks of "ConvTranspose" are stacked to upsample the feature maps to 64×64. ConvTranspose consists of a transposed convolutional layer, a batch normalization layer, and an activation function.

The final Conv block aims to output the final heatmaps H $\in R^{64 \times 64 \times 3}$ for the defined three virtual points P_w, P_{r1} and P_{r2}, respectively. The deviation between the predicted heatmaps and the ground-truth ones is the optimization objective for CNN training. We define mean square errors (MSE) as the loss function to quantify this deviation.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_l)^2$$

where n denotes the total number of pixels within a heatmap and is $64 \times 64 \times 3$ in our case. x_i is the true pixel intensity, while $\hat{x}_l$ is the predicted one.

We generate the Gaussian heatmaps from the defined points as the ground truth, where the 2D Gaussian function with constant variance, $\sigma = 2$ pixels, is centred at the position of the points. Figure 38 shows an example of a ground-truth heatmap generated by the 2D Gaussian function and feature map predicted by LightPointNet. In the inference process, the maximum point within the predicted heatmap is determined as the detected point.

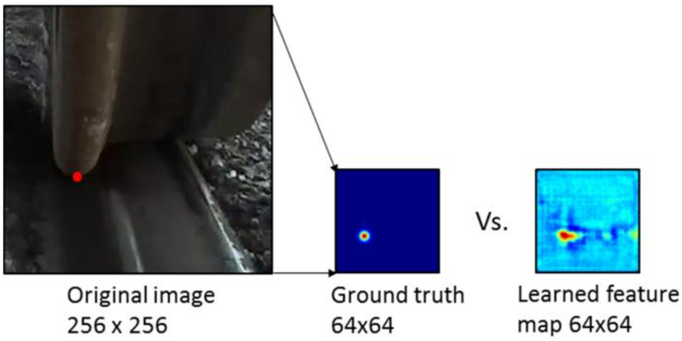


Figure 38 Example of ground-truth heatmap and feature map

3.2.3) Step 3: On-line point tracking by a rule engine

LightPointNet may encounter challenges leading to false detections during regular railway operations, particularly in corner cases. A notable example is illustrated in the third block of Figure 37, where the presence of grass obstructs the visibility of points P_w, and P_{r1}, making accurate detection unattainable within a single frame. To address these false detections, a rule engine was introduced as a point tracker, offering two distinct advantages in our railway application. Firstly, unlike challenges faced in object tracking or human pose tracking domains, our railway scenario typically involves repetitive scenes, such as wheels traversing rails. This repetitiveness allows for the establishment of rules based on intrinsic railway domain knowledge, capitalizing on the spatial correlations among virtual points governed by specific geometric constraints. Secondly, our application encounters specific data availability challenges, a common hurdle in domain-specific contexts. The rule engine, independent of CNN, proves advantageous in mitigating these challenges and refining point detections for enhanced accuracy.

In a recent study focused on real-time human pose tracking (Bazarevsky et al., 2020), the training process utilized a substantial dataset comprising 85,000 annotated images to train an 11-pose tracking network. However, in industrial applications, the collection and annotation of data can be labour-intensive and expensive. Field tests typically yield a more limited dataset for training convolutional neural networks (CNN). Notably, networks featured in previous works (D. Luvizon et al., 2018),(Bazarevsky et al., 2020) faced challenges in achieving satisfactory results on our specific dataset. Additionally, the intricate network architecture presented in (K. Sun et al., 2019) is not conducive to real-time execution on edge devices. To overcome these limitations, we adopt a hybrid approach, combining a DL-based point detector with a domain-knowledge-based tracker. This fusion enables end-to-end real-time point tracking, requiring significantly less training data and offering the flexibility for continuous improvement through additional model training. Furthermore, the rule engine can autonomously identify corner cases once the computer vision (CV) system is deployed for a long-term trial, facilitating the collection of corresponding video frames for ongoing model refinement.

The flow chart illustrating the operation of the rule engine is presented in Step 3 of Figure 37. We have established specific rules, along with corresponding indicators and thresholds detailed in Table 12. Each rule is assessed independently. Rule 1 and Rule 2 impose constraints on the y-coordinates of virtual points, denoting the projection of relative vertical and longitudinal motion between the camera and the wheel in the horizontal plane. These rules ensure that three virtual points are defined at the same horizontal level, i.e., y_{ref}.

The relative movement between rail reference points P_{r1} and P_{r2} is deemed non-existent. Significant disparity in y-coordinates between the wheel and rail reference points typically arises from wheel bounce induced by pronounced rail irregularities. However, such occurrences are infrequent and can be rectified through wheel acceleration measurement. Consequently, we assert that the y-coordinates of the three points should exhibit only a minor variation. To quantify this, the Root Mean Squared Error (RMSE) serves as the indicator for both Rule 1 and Rule 2. If σ_y in Rule 1 surpasses the threshold σ_{yTH_0}, the detection results are considered unreliable.

The detection results from the preceding frame are carried over to the current frame. When σ_y falls within the range of σ_{yTH_0} and σ_{yTH_1}, a correction mechanism is triggered for the detection results. The corrected values are determined by averaging the coordinates from the previous and current frames. The threshold values of σ_{yTH_0} and σ_{yTH_1} are established empirically based on our dataset. Similarly, Rules 3 and 4 restrict the difference in x-coordinates between P_{r1} and P_{r2}, representing the railhead's width, which may exhibit minor variations in practice due to wear. Rule 5 constrains the difference in x-coordinates between P_w and P_{r1}, indicating the potential maximum lateral movement of the wheel relative to the rail. This can be estimated by the maximum instantaneous lateral acceleration a_{ymax} of the wheel within the camera's sample period, with a_{ymax} being statistically estimated as a constant value derived from field measurement data for simplification.

Table 12 Defined rules, indicators, and thresholds in the rule engine

Index	Rules	Indicators	Thresholds
1	Y coordinate of the detected points should remain constant in comparison to the reference one (which can be manually defined in the calibration process or using the detection result on the first frame).	RMSE $\sigma_y = \frac{1}{3}\sqrt{\sum_{i=1}^{3}\left(y_i - y_{ref}\right)^2}$	$\sigma_{yTH_1} = 5mm$ $\sigma_{yTH_0} = 10mm$
2	Y coordinate of the detected points should remain constant in the adjacent frames.	RMSE $\sigma_y = \frac{1}{3}\sqrt{\sum_{i=1}^{3}\left(y_{i,t} - y_{i,t-1}\right)^2}$	$\sigma_{yTH_1} = 5mm$ $\sigma_{yTH_0} = 10mm$
3	The width of the rail head calculated by the x coordinates of the two rail reference points (P_{r1} and P_{r2}) should remain constant.	Difference $d_x = \lvert X_{r1} - X_{r2}\rvert$	$d_{xTH_1} = 5mm$ $\sigma_{xTH_0} = 10mm$
4	The two rail reference points should move in the same lateral direction or remain unchanged in the adjacent frames.	Boolean $\left(X_{r1,t} - X_{r1,t-1}\right) \cdot (X_{r2,t} - X_{r2,t-1})\geq 0$	$B_{TH_0} = True$

5	The wheel lateral displacement between two adjacent frames should be smaller than that calculated by the maximal wheel lateral acceleration.	Lateral displacement $\Delta D = \lvert X_{w,t} - X_{w,t-1} \rvert$	$\Delta D_{TH_0} = 0.5 \cdot a_{ymax} \cdot \Delta t^2$

3.2.4) Implementation of the baseline

The current target-less displacement measurement methods depend on traditional image processing techniques such as edge detection, line detection, and template matching (Jiang et al., 2020b; M. Kim, 2012). To provide a baseline for our application, we have implemented a method for comparison. Template matching was utilized for wheel detection and line detection for rail detection. The process is illustrated in Figure 39. In the calibration phase, template points for the wheel flange and two points on the left railhead edge are manually chosen.

The rail edge slope is computed using the two points on the rail, while the points on the wheel contribute to generating a wheel template. In the pre-processing phase, noise reduction is achieved through median filtering, and contrast enhancement is applied via histogram equalization. Edge extraction employs the Canny edge detector with a Gaussian blur filter. Subsequently, an automatic position detection of the wheel flange is executed using a template matching algorithm based on correlation coefficients. For line detection, a combination of filters is employed to highlight rail lines, and the probabilistic Hough transform is then applied for line detection. Considering the pre-calculated rail line slope, a specific range of slopes is defined to select the desired rail line from the detected candidates. Finally, the lines' extension allows for calculating the horizontal distance from the wheel reference point to the rail line.

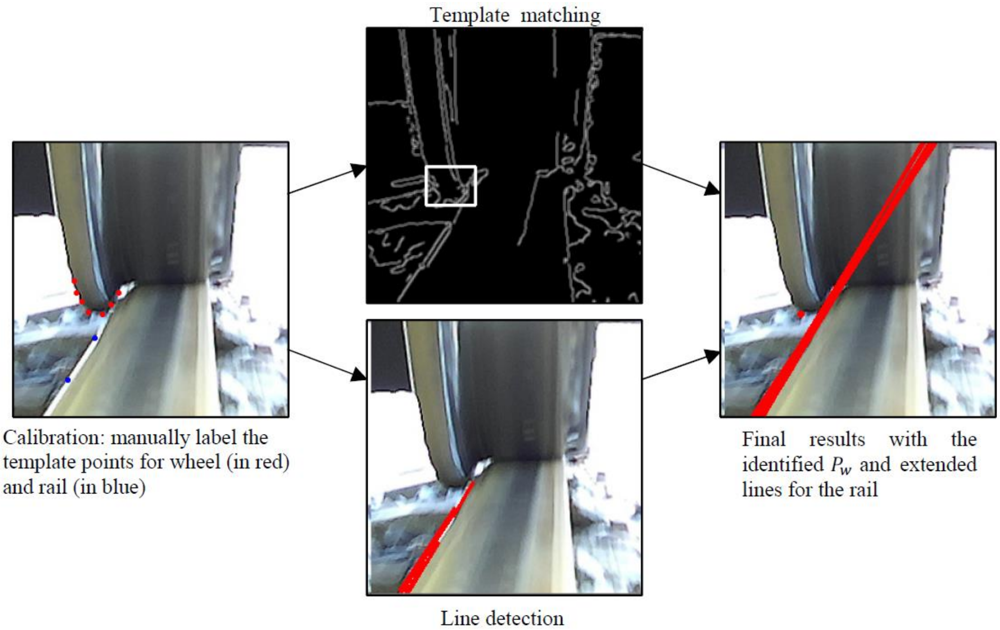


Figure 39 Adapted pipelines of the baseline method

3.2.5) Image corruption for data augmentation

As the computer vision (CV) system operates in a challenging railway environment, a robust housing structure has designed and incorporated an air cleaning system to safeguard the camera lenses from environmental factors. Additionally, a data augmentation procedure was proposed during the deep learning (DL) model training to enhance the model's resilience to potential image corruptions. Drawing insights from prior research on image corruption (A.B. Jung et al., 2020; Hendrycks & Dietterich, 2019), we have modelled relevant corruption types illustrated in Figure 40. For a given image I ∈ $R^{N\times N}$, $I(x,y)$ in the range (0, 255) denotes the original pixel intensity at the position (x,y). Gaussian noise may be introduced during optical sensing. The intensity function of the corrupted image $I_{gn}(x,y)$, injected with Gaussian noise is expressed as follows:

$$I_{gn}(x,y) = I(x,y)\,/\,255 + c \cdot p$$

$$I_{gn}(x,y) = \begin{cases} 0 & If \quad I_{gn}(x,y) < 0 \\ I_{gn}(x,y) \cdot 255 & If \quad 0 \le I_{gn}(x,y) \le 255 \\ 255 & If \quad I_{gn}(x,y) \end{cases}$$

where c is a settable scale representing the severe level and p is the Gaussian distribution.

Shot noise could occur during photon counting in optical systems. The intensity function of the corrupted image $I_{gn}(x,y)$ injected with shot noise is given by

$$I_{sn}(x,y) = f[c \cdot I(x,y)/255]/c$$

$$I_{sn}(x,y) = \begin{cases} 0 & If \quad I_{sn}(x,y) < 0 \\ I_{sn}(x,y) \cdot 255 & If \quad 0 \le I_{sn}(x,y) \le 255 \\ 255 & If \quad I_{sn}(x,y) \end{cases}$$

where f is subject to the Poisson distribution.

The modelled impulsive noise refers to salt-and-pepper noise which could originate from sharp and sudden disturbances in the imaging process. The intensity function of the corrupted image $I_{in}(x,y)$ injected with impulsive noise is given by:

$$I_{in}(x,y) = \begin{cases} 0 & If \;\; c \cdot \alpha/2 \\ I(x,y) & If \;\; 1 - c \cdot \alpha \\ 255 & If \;\; c \cdot \alpha/2 \end{cases}$$

where α is the probability that a pixel is altered.

Defocus blur is that the image is out of focus, which is caused by the fact that the camera integrates the light over areas during sensing. Blur is commonly modelled by convolution of the original image with a uniform point spread function (PSF). The defocus-blurred image I_{db} is given by:

$$I_{db} = I \ * \mathrm{K}$$

$$K(x,y) = \begin{cases} 0 & If \ \sqrt{x^2+y^2} \ < \ r \\ 1/\pi r^2 & If \ \sqrt{x^2+y^2} \ \geq r \end{cases}$$

where K is the parametric PSF for defocus blur and r is the radius parameter of K and linearly correlated to the severe level c.

Motion blur occurs when the vehicle is excited by large track/rail irregularities. The linear-motion-blurred image I_{mb} is given by:

$$I_{mb} = I \ * \mathrm{K}$$

$$K(x,y) = \begin{cases} 1/r & if \ 0 \leq x \leq r \\ 0 & \text{Otherwise} \end{cases}$$

where K is the parametric PSF for linear motion blur and r denotes the extent of the motion blur, relying on the severe level c.

Moreover, various weather conditions were modelled to enhance the robustness of the system. Snowy scenes were generated by introducing randomly placed white motion-blurred particles, simulating snowy conditions, and overall image whitening. Frosty scenes were created by overlaying the original image with templates of frosted glass. Fog was modelled using a plasma fractal generated through the diamond-square algorithm. The sunny/shady effect is simulated by adjusting the brightness of the original image, achieved by modifying the pixel intensity of the first channel in the HLS colour space. Additionally, standard augmentation techniques, such as horizontal flip, rotation, and occlusion, are applied. To replicate images captured from different camera positions and orientations, a 256×256 region of interest (RoI) from the original 1920×1080 image was randomly cropped at various positions. Subsequently, a point perspective transformation is applied to simulate variations in the camera's orientation.

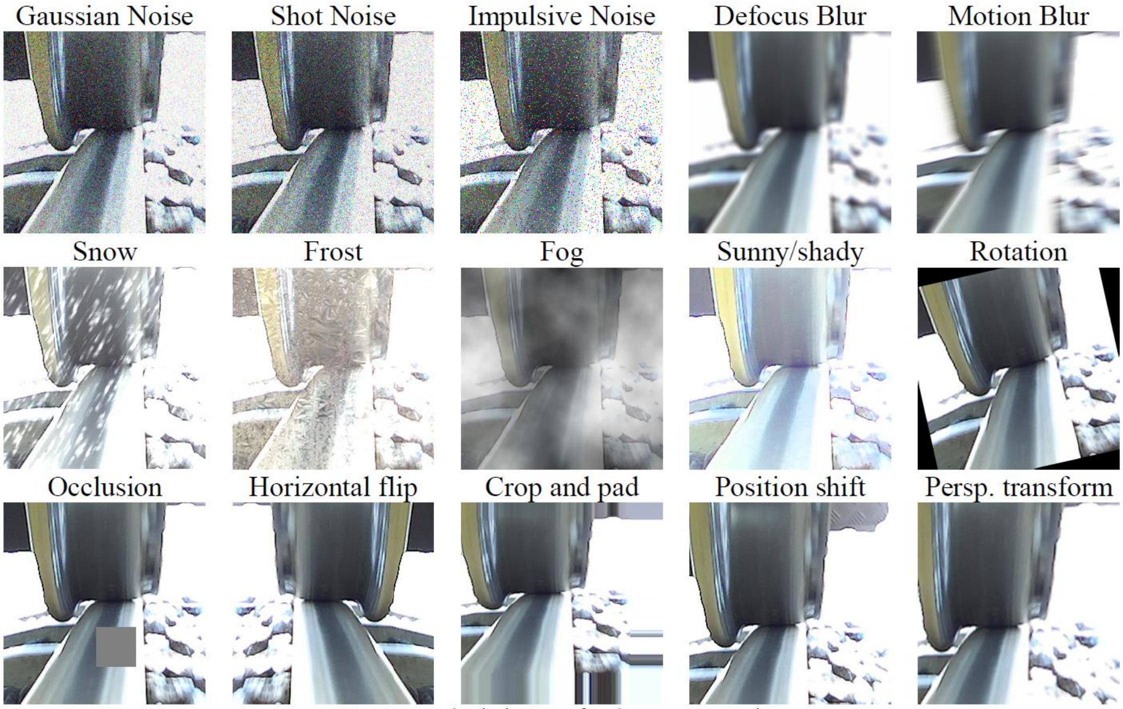


Figure 40 Synthetic images for data augmentation

3.3) Validation Test by Real World Experiments and Results

3.3.1) Field tests and datasets

Two field tests were carried out in Lithuania and Italy under various operational conditions. According to Figure 35 (a), the computer vision (CV) system prototype was mounted on the bogie frame of the track recording coach, Aldebaran 2.0, owned by Rete Ferroviaria Italiana (RFI), the Italian infrastructure manager. It was equipped with a commercial track geometry (TG) measurement system. These tests, which included straight and curved track sections, were conducted in the Catanzaro Lido workshop plant. The two switches with a curve radius of 170 meters and a tangent of 0.12 were corresponding to the curved sections. The Aldebaran 2.0 coach was driven by a locomotive at modest speeds, between 2 and 10 kilometres per hour, during the field test. Four camera configurations for various resolutions and sampling rates were examined, as well as two lateral camera positions with regard to the wheel. The video data from 3 test runs are used for model training, while 3 test runs are used for testing.

The test which was conducted in Lithuania on the mainline near Vilnius is shown in Figure 41. A track recording coach owned and operated by Lithuanian Railways has two CV systems mounted on its bogie frame. Both wheels' videos were captured at the same time. There were two forward runs at roughly 20 km/h and one backward run at a slower pace. Throughout the test runs, the camera setting stays the same. One forward run is used for training, while the other data is used for testing.

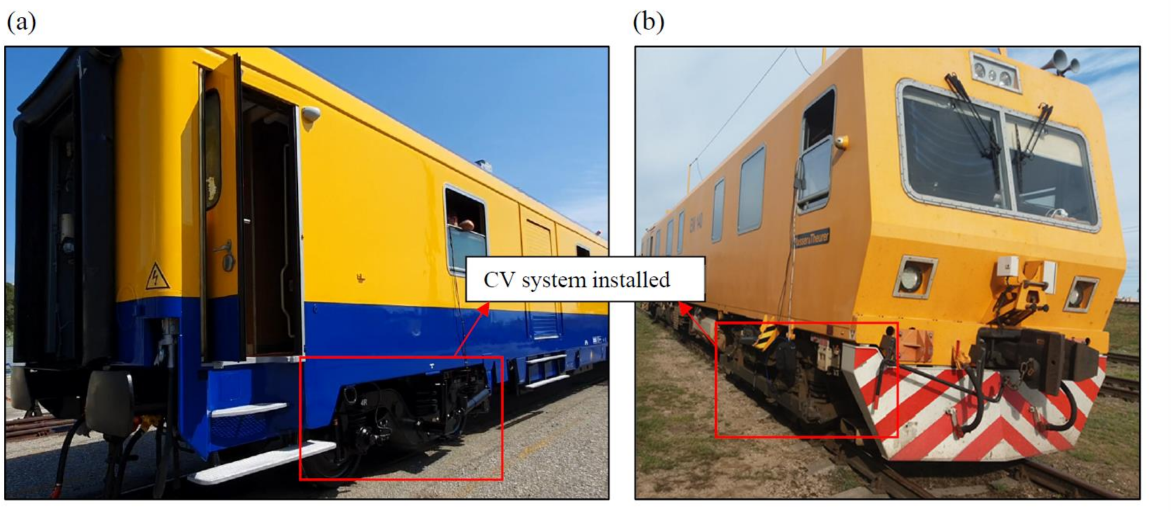


Figure 41 (a) CV system installed on a track recording coach in Italy; (b) CV system installed on a track recording coach in Lithuania

To build the dataset, considering the high sampling rate of 30 frames per second (fps) and the low vehicle speeds, one image was selected every 30 frames from the video data from Lithuania. Likewise, due to the similar low-speed circumstances, one image from every 60 frames of the video data was selected from what was gathered in Italy. Further photos were obtained from related (diiselrong, 2018) and static testing conducted at various places. The ZED2 stereo camera was used in these static testing to take pictures of different bogies that were positioned on the track; Figure 86 in Annex I provides examples of these bogies. Additional frames were retrieved from the YouTube video, which was captured with a GoPro camera during normal train operations, and are shown in Figure 87 in Annex I.

The 1920 x 1080 pixel original images were manually annotated with the required virtual points. For CNN training, the marked points' coordinates serve as the ground truth. In total, there are 767 annotated photographs here. Five 256 x 256 images of ROI was created and clipped at various points on each source image to maximize the amount of annotated data. Hence, there would be 3835 tagged photos in this manner. They are divided into three datasets: two hundred and thirty-one photos for training, seven hundred and sixty-seven images for validation, and a test dataset with a ratio of six to two. To validate the suggested approach, the following comprehensive experiments should have been carried out.

3.3.2) Training and evaluation of YOLOv3-tiny for calibration

In the implementation of YOLOv3-tiny, modifications were confined to the YOLO layers for Region of Interest (RoI) detection, however the first 13 levels, or Darknet, were left unmodified. This strategic approach allowed for the leveraging of pre-trained weights from Darknet for the modified YOLOv3-tiny, significantly reducing the need for annotated images during training. YOLOv3-tiny was first pre-trained using the COCO dataset (T.-Y. Lin et al., 2018), which consists of 123,287 labelled images in total, 3,745 of which are associated with railway scenarios. While the parameters particularly learned by the YOLO layers were lost, it was possible to transmit parameters from the Darknet.

The training dataset incorporated 800 images derived from static tests and the YouTube video. These raw images, initially sized at 1920×1080 pixels, were resized to 416×416 pixels to be compatible with YOLOv3-tiny. The model was trained using adaptive moment estimation (Adam) for 30 epochs, employing gradient-descent-based optimization. Following training, the performance evaluation of the trained YOLOv3-tiny model involved its assessment on 767 annotated images for keypoint detection. The evaluation metric gauged whether the labeled keypoints fell within the predicted bounding box of an image. Remarkably, YOLOv3-tiny demonstrated exceptional performance, achieving a detection accuracy of 100%.

Figure 42 Pipeline for training and evaluation of YOLOv3-tiny: (a) Pre-training on COCO dataset; (b) Fine-tuning on training dataset: (c) Evaluation on test dataset

3.3.3) Training and evaluation of LightPointNet

In the training of LightPointNet, the process involved training the network from scratch using the mentioned training dataset, which consisted of 2301 images resized to 256×256 pixels. The evaluation of the model was conducted on a separate testing dataset containing 767 images. Throughout the training iterations, validation was performed using a separate dataset to prevent overfitting by assessing the interim models created at each epoch.

Three 64x64 feature maps were produced by LightPointNet and were assigned to the specified reference points P_w, P_{r1},, and P_{r2}. Mean Squared Error (MSE), which quantifies the difference between predicted and ground-truth heatmaps, was used as the loss function during the training procedure. Adam with the learning rate of 0.00025 was applied to minimize the MSE loss over 150 epochs.

To ensure reliability, the training procedure was repeated five times, the most optimal model was chosen for further evaluation. This methodology considers the stochastic character of the learning samples as well as the unpredictability resulting from the initial randomization of CNN learnable weights. The evaluation criterion measured the difference in pixels between the ground-truth and predicted x-coordinates of P_w, and P_{r1}.

A comparison was made between LightPointNet's evaluation results and those of the baseline approach. The baseline technique does not require training, in contrast to LightPointNet. As an alternative, it employs a calibration procedure on the training datasets to determine suitable parameters for the Hough transform, edge detectors, and filters used in the baseline methodology.

A comparison between LightPointNet and the baseline approach is shown in Figure 43. Four categories are used to group the detection errors:

- "0-1 pixel" signifies either no error or an error of 1 pixel, which translates to 0.78 mm (depending on image resolution and camera-to-wheel distance).
- Errors within the range of 1-5 pixels are considered tolerable.

- Errors spanning 5-20 pixels are deemed unacceptable.
- Any error exceeding 20 pixels is classified as "miss detection."

LightPointNet shows no examples of miss detection in wheel and rail detection. On the other hand, the baseline approach has a miss detection rate higher than 20%, which can be attributed to its unreliable detection mechanism, especially in situations with noisy and diverse backgrounds. An analysis of the performance reveals that LightPointNet slightly outperforms the baseline in wheel detection compared to rail detection. In contrast, when it comes to rail recognition, the baseline approach performs noticeably worse than wheel detection. This difference results from the fact that rail line detection is less susceptible to failure in noisy and variable backgrounds than wheel detection, which is due to the intrinsic vulnerability of template matching.

On recordings, LightPointNet and the baseline approach were compared. However, a thorough validation by directly comparing predictions with ground truth could not be carried out since the ground-truth points in these sequences were not explicitly identified. In order to tackle this, a rule engine was used for the evaluation, and the number of predictions that surpassed predetermined criteria was taken into account as the measure. A number of video clips with various track arrangements were chosen so that Table 13 could compare LightPointNet's performance with the baseline technique. Especially, compared to the baseline, LightPointNet performed noticeably better in every video. Specifically, in films from Italy—which have lower brightness levels than the ones from Lithuania—the baseline model found it difficult to provide consistent results, nearly failing to deliver any consistent outcomes.

The disparities in lighting conditions between the regions highlight a challenge for the baseline method, necessitating a re-design of pre-processing techniques tailored to different lighting conditions, often requiring expert intervention onsite. However, LightPointNet displayed more robust against conditional fluctuations. By using methods like data augmentation, domain adaptation, and adding a more diverse training dataset, its robustness could be further improved. The fact that these enhancements are software-automated is significant since it highlights LightPointNet's flexibility and potential for growth.

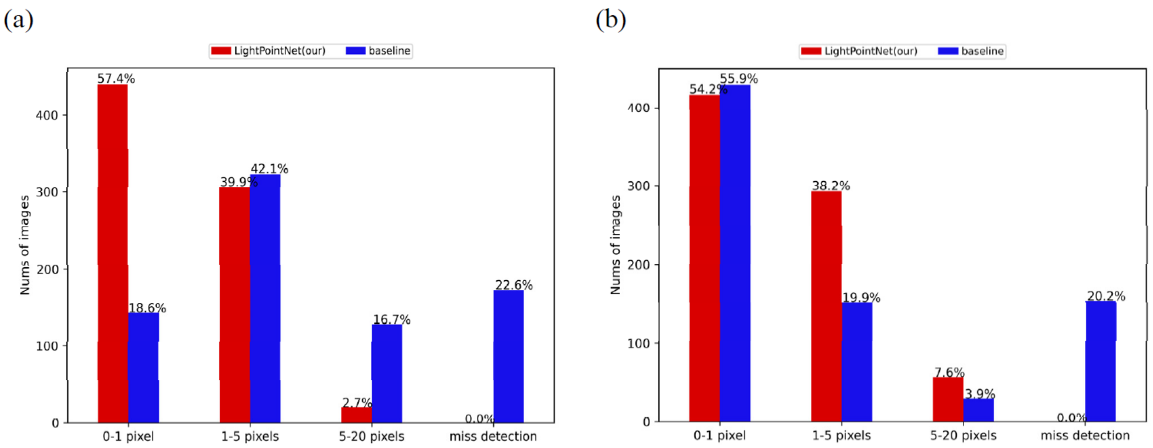


Figure 43 (a) detection errors (in pixels) for the wheel reference point; (b) detection errors (in pixels) for the rail reference point

Table 13 Test results of LightPointNet and the baseline methods on the representative video sequences

Video sequence	Model	Frame number	Implausible detection exceeding TH_0	Percentage of Implausible detection
Lithuania_forward_straight	LightPointNet	4997	622	**12.45%**
	Baseline	4997	3283	65.70%
Lithuania_backward_straight	LightPointNet	7996	102	**1.27%**
	Baseline	7996	670	8.38%
Italy_curve	LightPointNet	2586	263	**10.17%**
	Baseline	2586	2586	100%
Italy_swtich	LightPointNet	492	122	**24.79%**
	Baseline	492	430	87.39%
Italy_straight	LightPointNet	4024	912	**22.66%**
	Baseline	4024	3133	77.86%

3.3.4) Evaluation of rule engine

The evaluation compared the approach with and without the rule engine on the video sequences. In the method excluding the rule engine, LightPointNet processed each video frame, providing predicted coordinates as final results (Table 13 illustrates the percentage of implausible detections using this approach). In contrast, LightPointNet's outputs were processed when the rule engine was used. The purpose of this stage was to find corner instances and either correct or eliminate the corresponding implausible outcomes. Typical cases where LightPointNet faced challenges to produce dependable detection results are shown in Figure 44. These cases frequently included distinct track layouts at switch and crossing zones, as well as particular backgrounds like platforms or factories, which made detection difficult. For example, the rule engine was activated by sudden changes in the y-coordinate of P_w caused by wheel bounce. These unexpected shifts could cause significant motion blur, which would cause LightPointNet to miss detections. These difficult cases will be included in the training dataset so that the model can be further refined and perform better.

For evaluation of the correction mechanism, the trajectory of the LDWR across frames was examined. Figure 45 illustrates the trajectory calculated from the video sequence "Italy_straight," consisting of 4024 frames, comparing results with and without the rule engine. The rule engine-based correction mechanism, highlighted by red impulses, utilizes information from adjacent frames to eliminate unreliable sudden changes in coordinates, as depicted.

The tracking of the wheel's real lateral movement is unaffected by this modification. surprisingly, Fig. 12 shows a significant lateral wheel movement, shown by the blue line between the 2750th and 2950th frames. It's important to note, nevertheless, that the small-scale turbulence in anticipated coordinates could not be satisfactorily smoothed out by the rule engine. The data fusion with the corresponding wheel accelerations may cover this gap.

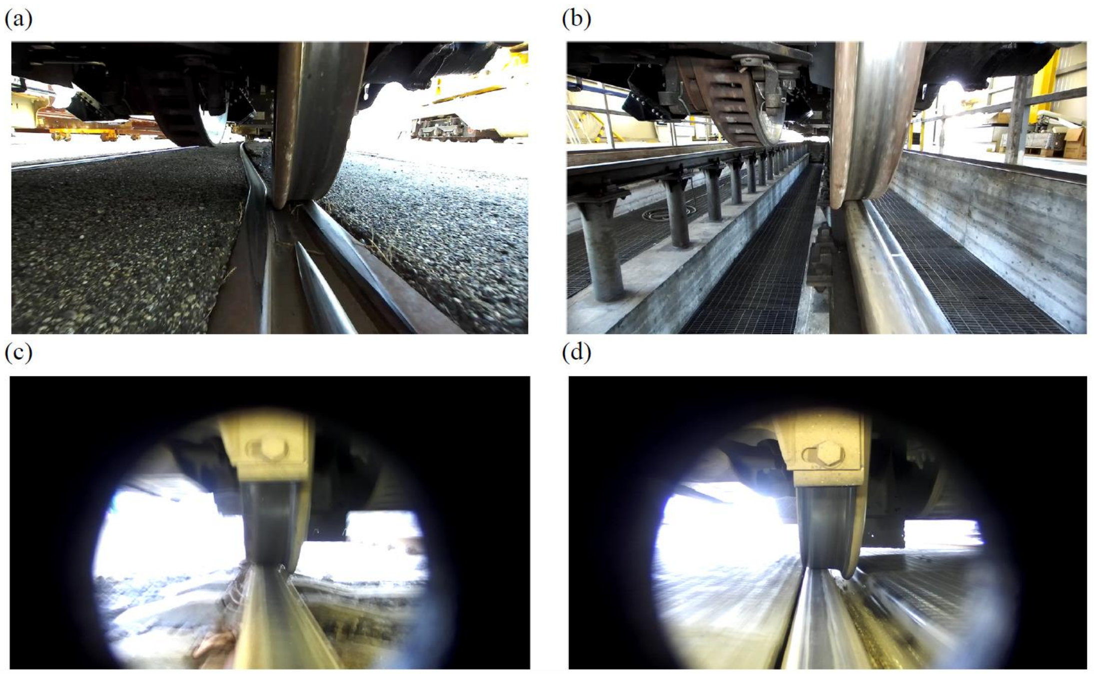


Figure 44 Corner cases detected by the rule engine: (a) switch & crossing zone; (b) workshop zone; (c) wheel bounce; (d) platform

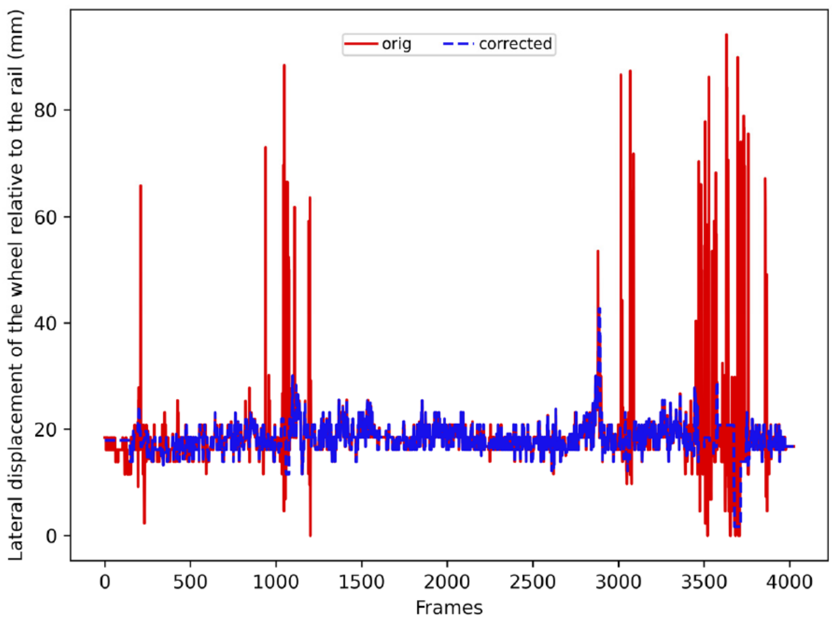


Figure 45 Comparison of the measured lateral displacement of the wheel relative to the rail with (blue) and without (red) the rule

3.3.5) Evaluation of the entire approach

The entire algorithm is run in real time on the Nvidia Jetson TX2 platform. Two video sequences were captured to show the tracking findings. The demo movies only show the 256×256 Region of Interest (RoI) that YOLOv3-tiny automatically detects, as the first stage in our suggested approach, because the spots are difficult to distinguish in the raw 1920×1080 photos. Figure 46 shows the tracked points P_w, and P_{r1} , which are on the rail edge and the wheel flange in the ROI. The lateral wheel displacement can be computed using these points. The point P_{r2}, which is located on the other side of the rail edge, is not shown in the demo movies but gives the rule engine geometric information.

Figure 46 Real-time lateral displacement measurement of the wheel on the rail to support track geometry monitoring (screenshot from the demo videos https://youtu.be/it21cE87LCM *and* https://youtu.be/Nc1bkQdkkSM)

The lateral motion of the wheel has been successfully tracked through virtual points; nevertheless, the detected virtual points show small shifts in the lateral direction, even while the wheel's actual position stays same. This causes abrupt changes in lateral displacement, which total a few millimetres (around 4 mm, based on observations). The main sources of this uncertainty are the displacement calculation and point tracking. The previous study during my master thesis ('Kaviani, 2020) used a stereo camera positioned at different distances and view angles in relation to a standard gauge block to test the displacement calculation method in a laboratory setting. The width of the block was calculated by manually choosing two reference locations on its edges; a measurement uncertainty (standard deviation) of 0.4 mm was found. As a result, LightPointNet's point detection is the main source of uncertainty in this situation.

A portion of this uncertainty results from the model's performance constraints after it was trained on available data. Furthermore, uncertainty is increased by label noise, which results from hand annotation of virtual locations as ground truth. It is difficult to determine the exact point placements on the railhead edge and wheel flange because of the labelling tool's limitations, complicated backdrops, and fluctuating lighting. Given the 1920 x 1080 resolution and the 0.78 mm distance between the camera and wheel, an annotation difference of several pixels within a comparable video frame is typical. With these considerations, a 4 mm measurement inaccuracy resulting from manual annotation makes sense and is difficult to completely remove. Improving ROI image resolution could be one such approach. To reduce these uncertainties, it is being considered to replace the existing camera with one that has a smaller field of vision and a closer focusing distance in future initiatives.

3.3.6) Computational complexity and real-time capability

The computational complexity of a CNN can be theoretically estimated based on the number of parameters and floating-point operations (FLOPs). A standard convolution layer includes 'N' convolutional filters, each characterized by weights $w \in R^{k \times k}$, where 'k' signifies the width. When it takes an input feature map $F_{in} \in R^{D \times D \times C}$ and produces an output feature map $F_{out} \in R^{D \times D \times N}$, the total parameters P_c and FLOPs can be determined using two formulas before, while disregarding the parameter count of bias and the accumulation operation.

$$P_C = k \times k \times C \times N$$

$$FLOP_{s_c} = 2MAC = 2 \times (K \times K \times C \times N \times D \times D)$$

In LightPointNet, regular $K \times K$ convolution is replaced with the combination of 1×1 pointwise convolution and $K \times K$ depthwise convolution, which is named as the depthwise separable convolution (DSC). Its parameter numbers P_{DSC} and $FLOP_{DSC}$ are significantly reduced, given by:

$$P_{DSC} = k \times k \times C$$

$$FLOP_{s_{DSC}} = 2 \times (K \times K \times C \times D \times D + N \times C \times D \times D)$$

The reduction ratio in parameter r_p and in operation r_{FLOP_s} are given by:

$$r_p = \frac{1}{N}$$

$$r_{FLOP_s} = \frac{1}{N} + \frac{1}{k^2}$$

It can be difficult to determine an algorithm for traditional image processing with accurate computational complexity. In our case, we primarily used the Hough transform for rail detection and a template matching technique for wheel detection in the baseline method. The Hough transform's computational complexity is commonly approximated to be $O(D^4)$ (Asano & Katoh, 1996), where D represents the size of the image. However, it can be difficult to give a precise rating because the actual complexity can differ based on the particular implementation.

For a more precise comparison, real-time performance—termed latency—of each algorithm is evaluated. Latency relies on the hardware and software platform used. PyTorch 1.6 is an open-source machine learning framework that is used in this application to create Deep Learning (DL) models. (Shi et al., 2022) For inference tasks, these models are implemented on the Nvidia Jetson TX2 edge computing hardware. OpenCV libraries are used to implement the baseline algorithm. This platform's calculation time is monitored, and average frames processed per second (FPS) is computed throughout a range of test video sequences. With this method, real-time processing speed comparisons between DL models and the baseline can be made.

Besides LightPointNet, two LightPointNet variants and PoseResnet (B. Xiao & H. Wu, 2018) were used for comparison. To reduce the amount of parameters and floating-point operations (FLOPs) while maintaining network performance, LightPointNet employs a variety of lightweight techniques. A linear bottleneck structure, squeeze-and-excitation (SE) modules, depth wise separable convolutions (DSC), and small-sized filters are some of these tactics. Further information about these low-weight metrics is available in an earlier study (Shi et al., 2021). To demonstrate the effectiveness of the lightweight architecture, LightPointNet is compared against PoseResNet-50, which uses ResNet-50 with conventional convolutions as the encoder. Moreover, LightPointNet provides the ability to scale the network design by changing the width and depth of the network. LightPointNet_large and LightPointNet_small were developed for this reason in order to examine how model size affects computational complexity and performance.

Table 17 and Table 18 in Annex II include the architectural details for LightPointNet_large and LightPointNet_small. All models were trained and tested with the datasets that were previously specified. The assessment metric is the average detection errors (in pixels) over the three reference points, P_w, P_{r1}, and P_{r2}. The computational complexity of several models is shown in Table 14. For the baseline, the approximate FLOPs are $O(D^4)$, where D = 256. With a third-party tool, the parameters, and FLOPs of the DL (Deep Learning) models were measured. For the original DL models run in PyTorch, the third row shows the delay in frames per second (fps). The fourth row shows the DL models' latency in Nvidia TensorRT format; this topic will be covered in more detail later.

An initial observation reveals that both the baseline and our LightPointNet operate at the same latency in terms of FPS, running at 20 fps, which falls short of the real-time requirement (i.e., 30 fps). Despite LightPointNet having substantially fewer FLOPs than PoseResNet, LightPointNet shows somewhat less latency than PoseResNet. This implies that platform-dependent latency is significantly influenced by factors other than FLOPs. In terms of parameters, LightPointNet has about 12 times less parameters than PoseResNet, which leads to a much smaller memory use. In particular, a PyTorch model of PoseResNet needs 130 MB, whereas a PyTorch model of LightPointNet needs 11 MB.

Table 14 Computational complexity of different variants of LightPointNet, PoseResNet-50 and the baseline method (M for mllion, G for Giga)

Results	**LightPointNet**	**PoseResnet-50**	**LightPointNet -large**	**LightPointNet -small**	**Baseline**
Parameters	2.86 M	33.99 M	3.35 M	1.22 M	-
FLOPs	2.87 G	14.54 G	3.10 G	1.29 G	4.3 G
FPS	20	16	18	25	20
FPS (TensorRT)	39	26	35	66	-

In terms of model performance, Figure 47 visualizes the detection errors associated with the DL models. LightPointNet displays comparatively inferior performance compared to PoseResNet, especially within the "0-1 pixel" category. This observation highlights a trade-off between performance and model complexity. In comparing various iterations of LightPointNet, LightPointNet_large has an encoder that is wider and longer than the original LightPointNet, but it still has the same decoder structure. As shown in Figure 47, this leads to a slight increase in parameters and FLOPs along with better model performance. On the other hand, LightPointNet_small uses a thinner decoder while keeping the same encoder as the original LightPointNet. The original LightPointNet's parameters and FLOPs are reduced by more than half in this modification. Unfortunately, there is a discernible loss in accuracy because of this reduction.

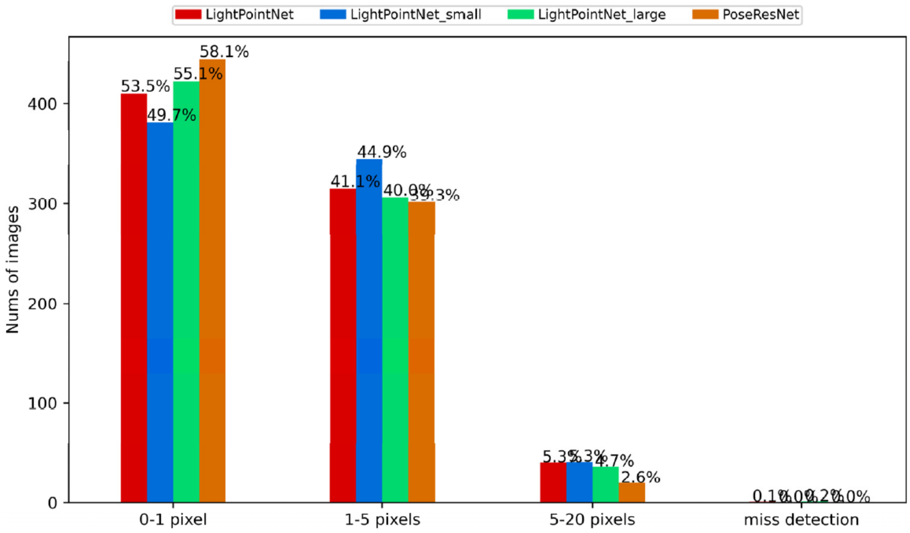


Figure 47 Detection errors (in pixels) of different variants of LightPointNet and PoseResNet-50 averaged over three reference points P_w, P_{r1} and P_{r2}

It seems that on the specified platform, none of the PyTorch models have real-time capabilities. We have transformed the PyTorch models into the TensorRT format, which is intended to accelerate DL model inference on Nvidia's GPUs, to reduce latency. TensorRT converts the learnt weight and bias parameters from their native float32 (32-bit) representation to an 8-bit version by imposing low-precision inference. This modification reduces accuracy loss while greatly speeding up inference. To further increase computational speed, TensorRT also improves a neural network's computation graph. You may read more about this procedure in (H. Abbasian et al., 2020). The last row in Table 14 shows the DL models' latency after being converted to TensorRT format. PoseMobilenet models satisfy real-time requirements with success, apart from PoseRenet. It is important to remember that latency is also a result of the pre-processing process, which involves capturing frames from stereo video recordings using the camera's API. After a thorough analysis, we have decided to favor LightPointNet over LightPointNet_large for application-related reasons.

3.3.7) Data augmentation for robustness enhancement

The resilience of a model in challenging outdoor environments is pivotal. Serious track abnormalities can result in significant vibrations that blur images, while dust or dirt on the camera lens can generate occlusions. Sensor deterioration or interference can also introduce image noise. Furthermore, different weather patterns produce different intensity distributions in photographs. A corrupted testing dataset was created to simulate these image corruptions. A randomly chosen corruption method is used to enhance each of the 767 photos in the original test dataset, as shown in Figure 40. The degree of corruption is controlled by a severity scale c that is randomly assigned between 1 and 3. Both LightPointNet and the baseline on this corrupted dataset was tested to demonstrate how model performance degrades under image corruption.

LightPointNet is initially trained solely on a clean training dataset without corrupted images. Subsequently, for doing the data augmentation, the corruption process on the training dataset was duplicated. Using this supplemented dataset, LightPointNet is further trained, resulting in the creation of an augmented model (LightPointNet_aug). To evaluate how well data augmentation improves the model's resistance to picture corruptions, the updated model with the original one was compared.

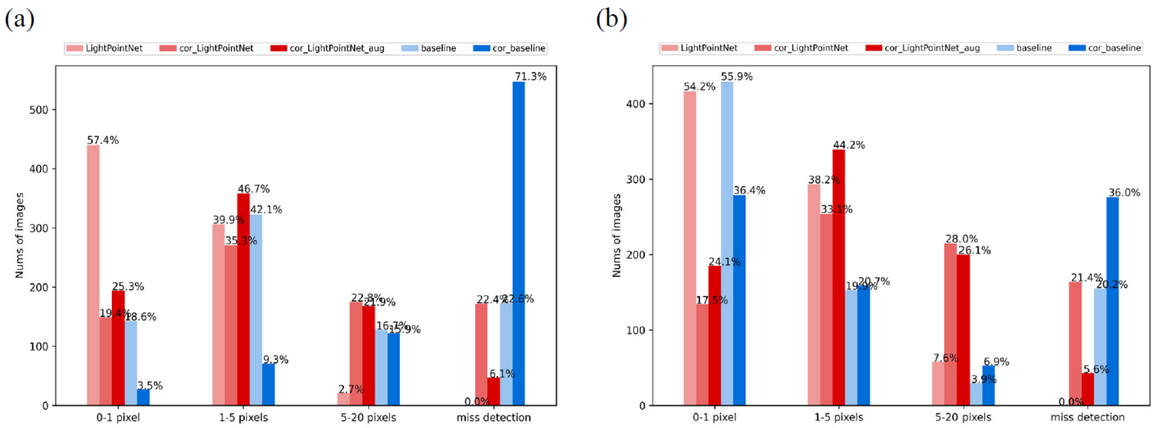


Figure 48 (a) detection errors (in pixels) for the wheel reference point; (b) detection errors (in pixels) for the rail reference point. "LightPointNet" in light red and "baseline" in light blue represent the test results of LightPointNet and the baseline method on the clean test dataset. "cor_LightPointNet" in medium-light red and "cor_baseline" in blue represent the test results of LightPointNet and the baseline method on the corrupted test dataset. "cor_LightPointNet_aug" in red represents the test results of the augmented LightPointNet on the corrupted test dataset.

The experiment's findings regarding detection errors for both wheel and rail reference locations are illustrated in Figure 48. The test results of LightPointNet and the baseline algorithms on the clean test dataset are shown by "LightPointNet" (shown in light red) and "baseline" (represented in light blue), respectively.

LightPointNet and the baseline across wheel and rail reference point identification perform noticeably worse when picture corruption is introduced. The miss detection rates for the "cor_LightPointNet" (medium-light red) show a significant increase, going from 0.0% to 22.4% for wheel detection and from 0.0% to 21.4% for rail identification.

In particular, the corrupted test dataset shows a startling 71.3% miss detection rate for the baseline ("cor_baseline" in pure blue), almost failing in wheel detection. This demonstrates a markedly lower level of robustness in template matching for wheel detection when compared to LightPointNet, which is DL-based. Nonetheless, the baseline's use of a line identification method performs somewhat better than LightPointNet in terms of robustness for rail detection. However, domain adaptation and data augmentation can further strengthen the robustness of DL models without requiring changes to the network architecture.

In this experiment, data augmentation has been successfully validated for enhancing robustness. Notably, "LightPointNet_aug" (pure red) exhibits notably enhanced performance compared to the original LightPointNet. The miss detection rate drops from 22.4% to 6.4% for wheel detection and from 21.4% to 5.6% for rail detection, signifying substantial improvement post-augmentation.

3.3.8.) Validation Test with developed algorithm

After developing the algorithm, there was a need for validation test. The test conducted in Italy covered approximately 330 km along the conventional railway line from Roma Tuscolana station to Pisa Centrale station. Our prototype was strategically installed on the bogie frame of the Aldebaran 2.0 coach, the latest addition to RFI's track recording vehicles. This choice allowed us to directly compare the data collected by our sensor system with the data obtained from the commercial track geometry measurement system (TGMS) provided by MERMEC, enabling a comprehensive benchmarking process.

Despite some line sections permitting a maximum speed of 180 km/h, we restricted the train's speed to 100 km/h. This decision was twofold: primarily due to the Aldebaran 2.0 car still being in its pre-operation phase and owing to the train's low braked weight percentage, comprising only three vehicles (two locomotives and the Aldebaran 2.0 coach as depicted in Figure 49. For image acquisition and processing, the stereo camera was linked via a USB cable to a PC located within the coach. Simultaneously, the lighting system was connected via cable to an Uninterruptible Power Supply (UPS) system situated in the same coach, ensuring continuous and reliable power supply during the operation.

Figure 49 train used for validation test, consisting of two E652 locomotives (leading and trailing) and the Aldebaran 2.0 track recording car.

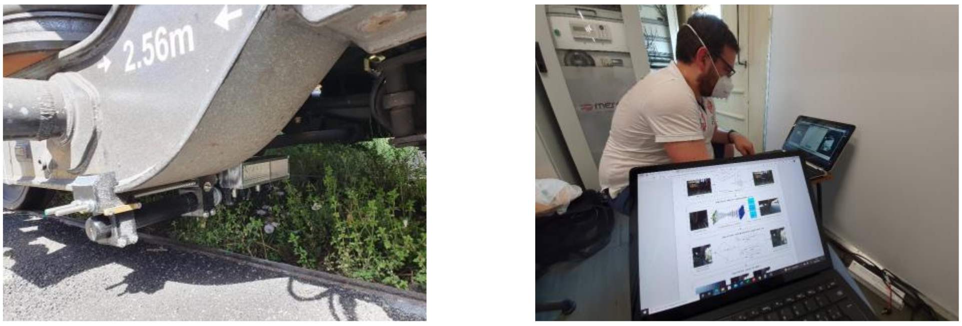


Figure 50 left image: sensor system mounted on the bogie frame; right image: PC for image acquisition and processing, positioned in the coach.

During the test, both the stereo-camera and the MERMEC Inertial Platform system operated simultaneously, capturing distinct sets of data. The stereo-camera outputted the Lateral Displacement with respect to Wheel-Rail interface (LDWR), while the MERMEC system provided Track Gauge, Cross Level, Twist, Longitudinal Level, and Alignment measurements.

For our comparison between the two systems, our focus primarily rested on the alignment of the right rail, specifically within the wavelength range D1: 3 m < λ ≤ 25 m [mm]. This parameter, known as the deviation y_p in the y-direction of consecutive positions of point P (referencing EN 13848-1:2003+A1:2008, section 4.2.1), denotes the excursion from the mean horizontal position (reference line). It's calculated from successive measurements (referencing EN 13848-1:2003+A1:2008(Figure 51).

However, it's important to note that the MERMEC system outputted processed results only, encompassing track geometry parameters and indicators of rail profile wear. Notably, these parameters lacked LDWR data and, consequently, weren't directly comparable to the results provided by our system. Our computer-vision-based sensor system was specifically designed to achieve real-time LDWR measurement. It operates in conjunction with accelerometers and relies on a sensor fusion algorithm to reconstruct lateral alignment accurately.

3.3.9) Comparison of sensor system measurements

The synchronization between the Lateral Displacement with respect to Wheel-Rail interface (LDWR) from our sensor system and the alignment data from the MERMEC system was carried out to explore their relationship. The analysis revealed that while the wheels don't precisely adhere to the lateral alignment as they do to the vertical one—this was expected—the general trend observed was that larger lateral alignment tends to induce more significant wheel movement in the lateral direction.

This observation implies a correlation between the measured alignment (Figure 51) and the LDWR computed by our computer vision-based sensor system, even though it does not constitute a strong comparison validation. It's crucial to remember that lateral wheel acceleration data needs to be included to LDWR in order to properly infer alignment information.

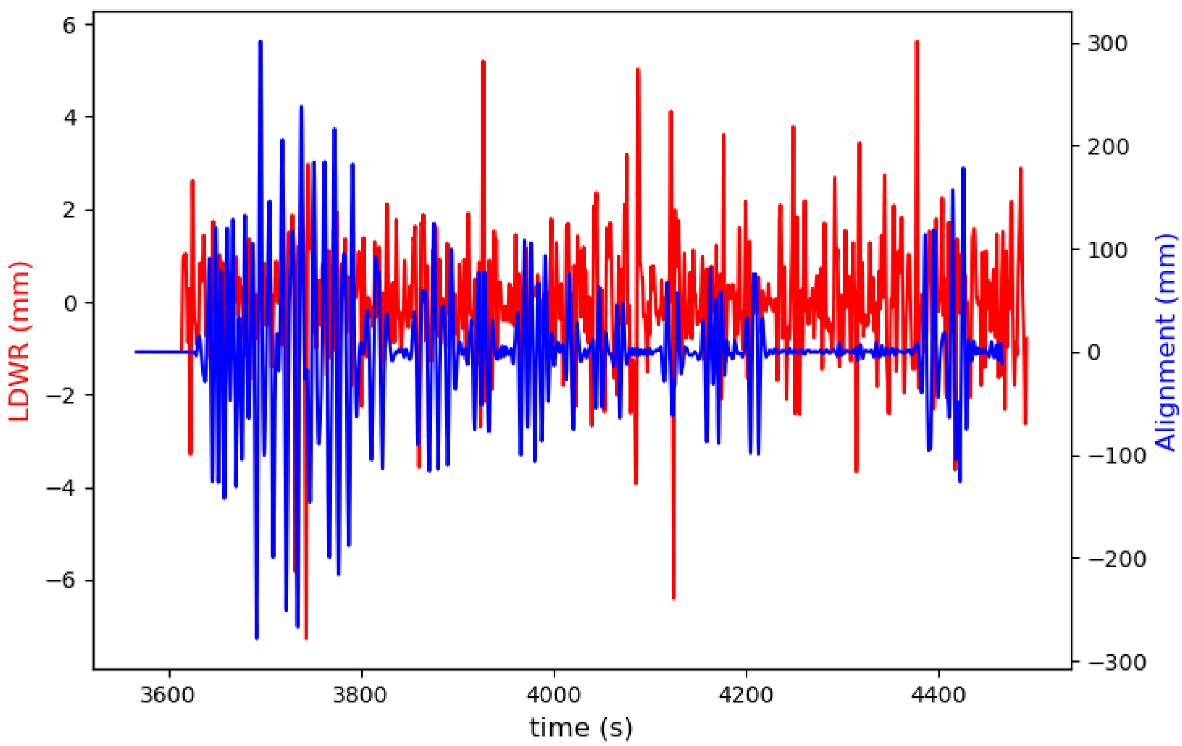


Figure 51 Comparison between LDWR calculated by our CV system and Alignment D3 measured by MERMEC TGMS

3.4) Enhancing Data Integration through Vehicle Dynamic Simulation

Simulations played a crucial role in complementing the validation test, especially due to the unavailability of precise accuracy in the data obtained from the axle-box accelerometer during the validation phase. To ensure a comprehensive understanding and arrive at accurate results, simulations were employed. These simulations were designed to replicate the real test conditions while incorporating necessary parameters. This approach aimed to provide the answers required to address the research questions effectively.

Simpack® and Gensys were the simulation software used in this work. Simpack specializes in multibody simulation, analyzing mechanical interactions in railway systems. Gensys focuses modeling mechanical, electrical and mathematical problems., aiding in simulating and analyzing system behavior and response. These tools allowed for comprehensive modeling and analysis of railway dynamics and control mechanisms in varied conditions. The further explanations can be seen in the following.

3.4.1) Simpack Simulation

It was chosen to replicate the 26 km of the ride that Aldebaran 2.0 made during the test by using the Simpack multibody simulation program, including the virtual axle-box acceleration measurement this time. A high-pass and low-pass filter were used to filter the virtual axle-box accelerations initially. The low-pass frequency, Vmax/3, and the high-pass frequency, Vmin/25, of the filter band were determined by calculating the relationship between the train speed and the defect wavelength band D1, which is located between 3 and 25 meters.

To comprehend the difficulties in utilizing accelerometers for the evaluation of wheel displacement, the virtual lateral accelerations were double-integrated and compared with the virtual lateral displacement of the wheel. When negotiating curves, where the quasi-static component of the lateral acceleration matters, a Kalman filter was employed to enhance the match. The outcomes are displayed in Figure 52, where it is evident that the values of the lateral wheelset displacement obtained by Kalman-filtering and double-integration ("calculated LD" – magenta line) closely correspond to the benchmark ("real" LD – blue line), especially in curves and on straight tracks (the track's curvature is indicated with an orange line).

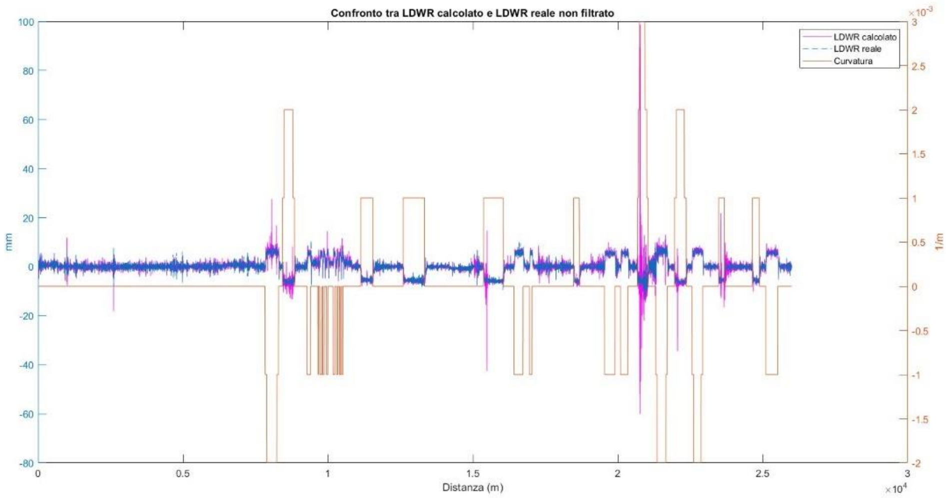


Figure 52 Comparison between the double-integrated virtual accelerometer signal ("calculated") and the "real" benchmark LD.

An interesting comparison may be found in Figure 53 between the values of the "calculated LD" and the total of the left and right "L+R" alignment (i.e., the measured TG values used as simulation input). They visually appear to be somewhat correlated (the orange line represents the L+R alignment, while the blue line represents the LD).

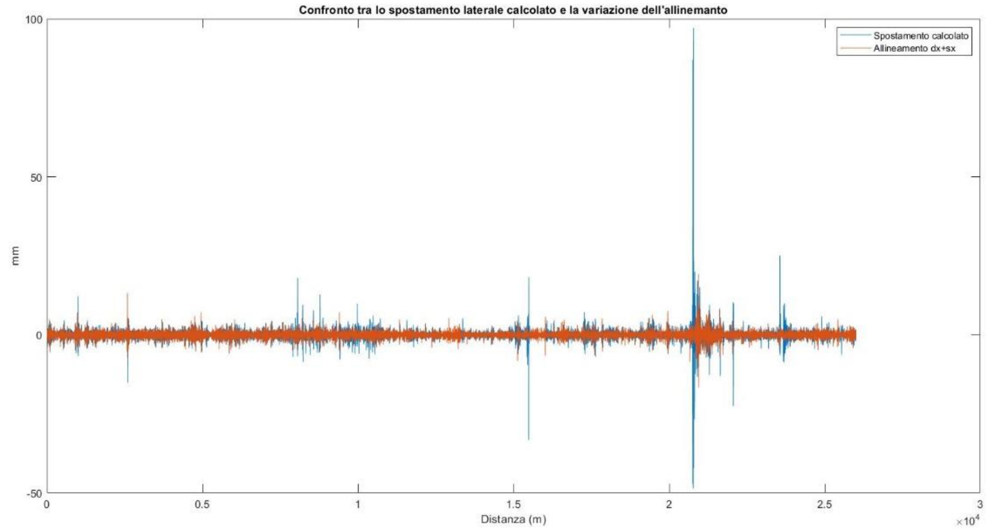


Figure 53 Comparison between the "calculated LD" (double-integrated, Kalman-filtered) and the track alignment (sum of the right and left rail alignment) along all 26 km of the simulation

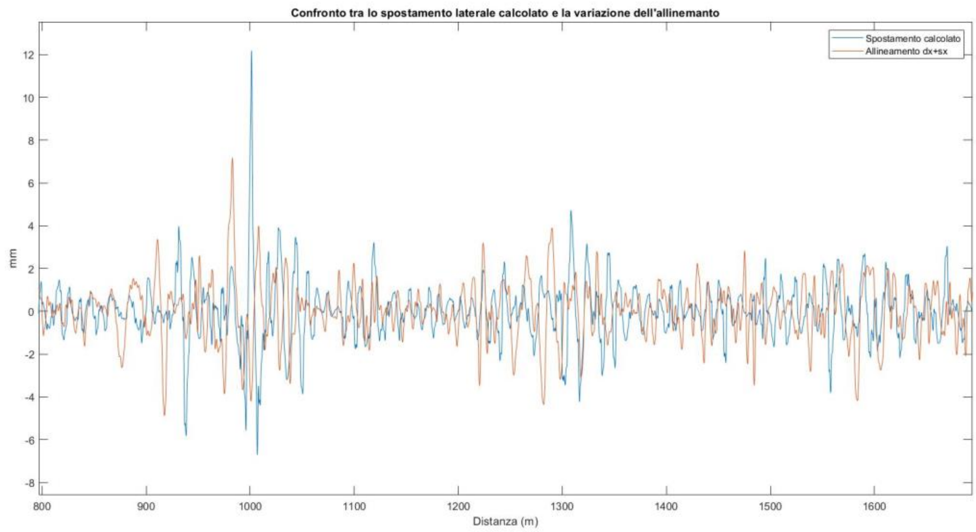


Figure 54 zoom on the previous figure on a section of 1 km

As expected, a phase shift is visible between the peaks of the LD and the L+R alignment. This is dependent upon the vehicle's lateral inertia as well as the suspension's properties and the wheels' corresponding conicity. It is intriguing to investigate if this phase shift influences other TG values and the standard deviation over 200-meter sections, which correlates to important maintenance indications when applied to both left and right alignment. Throughout the whole 26 km of the simulation, the standard deviations over 200 m-sections of the "calculated LD" and the measured L+R alignment were compared (Figure 55).

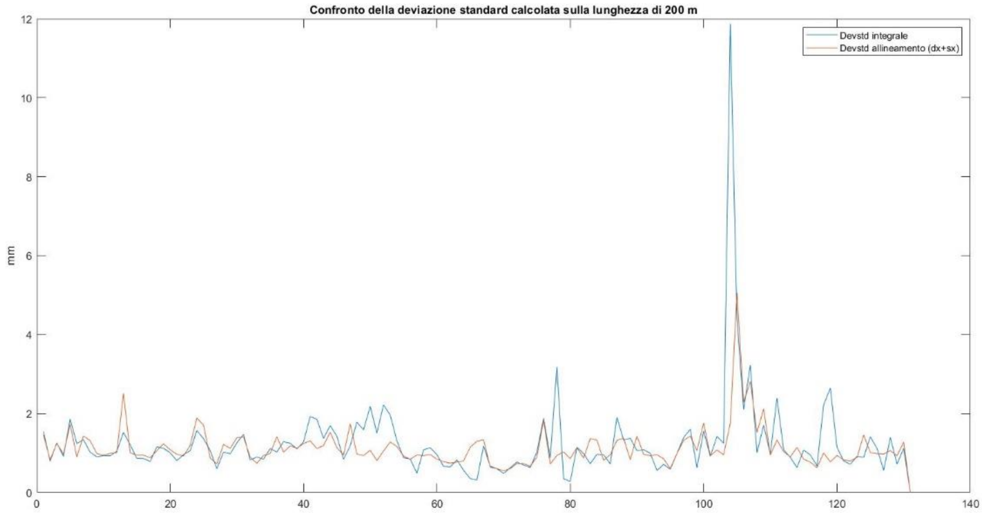


Figure 55 Comparison of the standard deviations on sections of 200 meters of the calculated LDWR and the measured alignment along all 26 km of the simulation (130 points)

This comparison also shows a reasonably good visual correlation between the LD and the L+R alignment. As mentioned above, the curves could undergo further improvement in the match with further work. The 12 mm peak at approximately the abscissa 104 also shows a less satisfactory match. It corresponds to a line section in which there are numerous switches leading to higher lateral displacement values. The R^2 representing the correlation between the two standard deviation curves turns out as 0.5 if the entire section is considered. Exclusion of the data associated with the 12 mm value peak causes it to rise to 0.7. A further rise to 0.8 occurs if only the first straight section is considered.

So far, the multi-body simulation of the running dynamics of a virtual replica of the Track Recording Vehicle used for the tests showed, mostly on straight track, a good correlation between the lateral wheelset displacement values obtained by integrating the axle-box accelerations with the ones obtained directly from Simpack and with the sum of the measured left and right alignment.

In addition, the stability of the train ride on the rail with the inclination of 1:20 was facing some challenges, that is why there was a need to confirm this issue with using another simulation software, too. Therefore, the outputs of Gensys simulation which will be explained further could be the inputs of the Machine Learning model.

3.4.2) Gensys Simulation

This part of the work was done by cooperating the railway group of NTNU, during my exchange program using the Gensys; Swedish multibody simulation software.

A train running on two different standard tracks (ST) are used in this work. The train has one car body with two bogies and 4 wheelsets. Car body, bogies and wheelsets are rigid masses coupled by spring and dampers. The characteristic of the train can be found in (Gensys).

The characteristics of the standard tracks(ST), are as follow.

- ST1 – Czech high speed test track , VUZ VELIM including 2 curves, with a total track length 13.276 km (VUZ).
- ST2 – Polish lower speed test track ZMIGROD including several curves, with a length of test track 7725 m (Zmigrod).

The choice of these tracks was based on the availability of access to their characteristics, and good design of the layouts. The lateral irregularity on the track was modelled using Swedish standard irregularity on the first quality level (QN1) based on the (UIC 518: 4ED, 2009) and (EN 14363, 2005). The simulation test was performed at three different speeds for each track, with ST1 speeds of 130, 160, and 210 km/h and ST2 speeds of 50, 80, and 120 km/h. The simulations were performed with different inclinations, including 1:20, 1:30, and 1:40. The only geometry parameter considered in the study was lateral irregularity, and 18 simulations were conducted, which could be seen in Table 15.

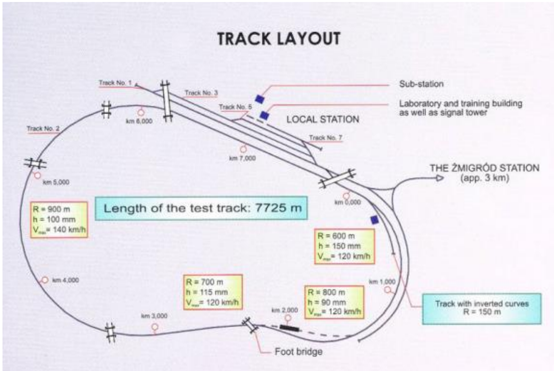


Figure 56 Standard Tracks: Czech VUZ VELIM track (left), Polish ZMIGROD track (right)

Table 15 Gensys Simulation

Scenario	Track type	Speed (km/h)	Inclination	Irregularity type
1	ST2	50	1 in 20	K0
2	ST2	50	1 in 30	K0
3	ST2	50	1 in 40	K0
4	ST2	80	1 in 20	K0
5	ST2	80	1 in 30	K0
6	ST2	80	1 in 40	K0
7	ST2	120	1 in 20	K0
8	ST2	120	1 in 30	K0
9	ST2	120	1 in 40	K0
10	ST1	130	1 in 20	K0
11	ST1	130	1 in 30	K0
12	ST1	130	1 in 40	K0
13	ST1	160	1 in 20	K0
14	ST1	160	1 in 30	K0
15	ST1	160	1 in 40	K0
16	ST1	210	1 in 20	K0
17	ST1	210	1 in 30	K0
18	ST1	210	1 in 40	K0

The Swedish wheel profile defined in EN S1002 T32.5, and the rail profile was 60 EI with conicity of I20 and the focus among the results of the simulation, is just the lateral acceleration of the wheel, and LDWR which could be used for the Machine Learning model inputs, with the goal to find the relationship between lateral track irregularities and the lateral displacement.

The results of this simulation were used to create and develop a Machine Learning model to detect the lateral irregularity.

3.5) Regression Models by Machine learning approach

Recently, the capability of ML models to monitor the health of structures and track geometry has been proven and is in a high interest of researchers and industry. As it was mentioned before, it is a powerful method due to the non-linearity of the problem.

Regression methods are used in this work to investigate the relationship between the variables and to confirm the simulation results. Regression is a technique which helps in examining the relationship between independent features or variables and a dependent feature, outcome, or variable. To ensure that altering one algorithmic parameter does not drastically change the results of the entire model, it is intended to start with simple regression algorithms that can support related case studies without relying on many algorithmic parameters.

For this project, supervised learning was chosen since it is a machine learning task of learning a function that maps an input to an output based on example input-output pairs (Kotsiantis, 2007). Regression entails choosing the best model to suit the provided data set and then using the model to make predictions in the future (Breiman et al., 2017).

In different situations, various regression models are utilized, and their appropriateness depends on the type of data and the relationships between the variables. For instance, when dealing with multiple variables showing a polynomial relationship, the preferred choice is a Polynomial Regression Model. On the other hand, for data exhibiting a linear relationship between two variables, the most suitable option is a Linear Regression Model.(Vladimir Lyashenko).

Different algorithms which are used in this study are random forest regression, polynomial regression, and support vector machine regression, to predict lateral irregularity by using numerical simulated lateral acceleration of the train bogie and lateral displacement of the leading wheelset.

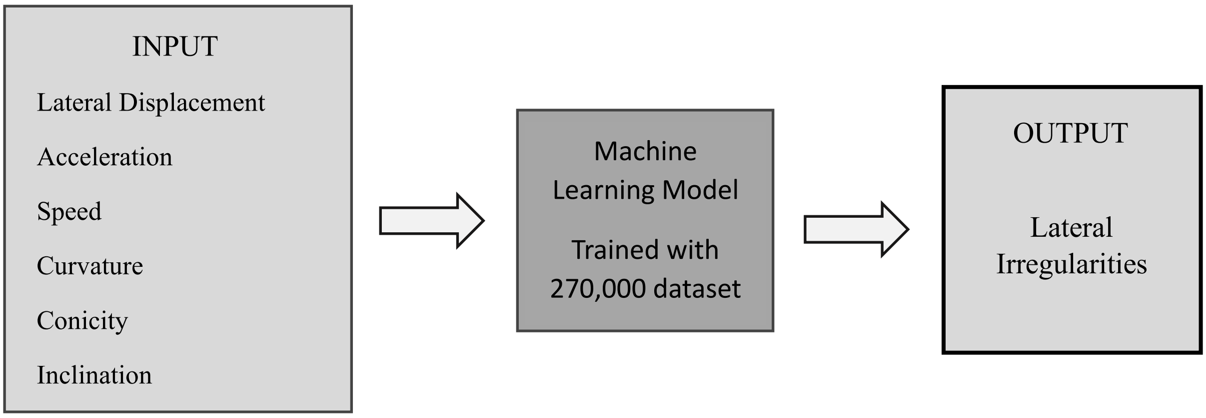


Figure 57 Input-Output figure for Machine Learning Model

3.5.1) Data management

The selected features for training the model are lateral acceleration of the wheel, LDWR, conicity and the lateral irregularity of the track. the lateral irregularity is employed throughout the training phase because it is supervised learning. As a result, the model is learning how to recognize lateral irregularity only by having the lateral acceleration and LDWR in future.

Several attempts were made at the train-test split, and it was found that changes in inclination could significantly affect the results if the model was not trained and tested on the same inclinations. To address this, three different groups were created based on the different inclinations, and the sample model was selected based on the highest accuracy of each group. Each sample model was then trained using the full dataset of that group and tested on other case studies with the same inclination.

The selected sample model was the standard track 1 in Czech, which was trained at a speed of 130 km/h for three distinct inclinations of 1:20, 1:30, and 1:40. To evaluate the results and select the most effective method for the objective, Random Forest Regression, Polynomial Regression, and Support Vector Machine Regressor were applied. During the testing phase, the track was divided into different sections including the whole track, straight track, and only the curve section of the track. This division was made to gain a better understanding of the model's errors and functionality on each section by comparing their accuracies.

After applying the model on different case studies, the errors were calculated regarding to the two metrics which will be discussed below.

3.5.2) Performed metrics

The evaluation of the model's performance involved calculating the Root Mean Squared Error (RMSE) and R-squared (R2) values. These two indicators are commonly used by statisticians to assess the fit of a regression model to a dataset. The RMSE measures the average difference between the predicted and actual values in the dataset and is not scaled to any value. On the other hand, R2, also known as the coefficient of determination, ranges between 0 and 1, where values closer to 1 indicate a better fit of the model to the dataset. The aim was to obtain a low RMSE and a high R2 to demonstrate the effectiveness of the model (Zach, 2021).

They are calculated as:

RMSE = $\sqrt{\sum}(P_i - O_i)^2 / n$

where:

Σ is a symbol that means "sum"

P_i is the predicted value for the i^{th} observation

O_i is the observed value for the i^{th} observation

n is the sample size

R^2 = 1 – (RSS/TSS)

where:

RSS represents the sum of squares of residuals

TSS represents the total sum of squares

3.5.3) Data aggregation to develop the model

Through data analysis, important parameters were identified among the simulation outcomes that could have a significant impact on the model's ability to predict results. Therefore, all of these parameters were included as input features to create a general model that could be trained and tested using a 70/30 split. The model incorporated various track characteristics, such as LDWR, lateral acceleration, speed, inclination, curvature, and conicity, although conicity cannot be measured in real world. As a result, the model was tested twice, once with conicity included and once without it. The following chapter presents the results of these tests and the distribution of errors.

3.6) Sensitivity Analysis

As it was mentioned earlier, in engineering, machine learning (ML) methods have been widely used, especially to solve high-dimensional and non-linear problems (Chen et al., 2019; Tao et al., 2019; L. Zhang et al., 2017). However, handling high-dimensional data poses difficulties for decision-making and data analysis. Feature selection (FS) techniques have been developed with the goal of efficiently processing high-dimensional data in order to address this problem (Gao et al., 2018; Hosseini & Moattar, 2019). To improve model correctness, the FS method looks for the best possible feature set that yields the most information (P. Zhang, 2019). Based on a variety of parameters, Cai et al. (Cai et al., 2018) categorized FS approaches into multiple groups. The three general categories of FS methods are filter, wrapper, and embedding models, which are distinguished by their respective connections to learning approaches.

In this work, Importing the importance of the Permutation feature was chosen for sensitivity analysis. When the data is tabular, a model inspection method called permutation feature importance can be used to any fitted estimator. This is particularly helpful for unclear or non-linear estimators. The reduction in a model score resulting from a single feature value being shuffled at random is known as the permutation feature significance (Pedregosa et al., 2011)

As a result, it can evaluate the relative impact of each quantity on the issue; the outcomes are displayed in the conclusion chapter.

3.6.1) Outline of the permutation importance algorithm

- Inputs: fitted predictive model m, tabular dataset (training or validation) D.
- Compute the reference score s of the model m on data D (for instance the accuracy for a classifier or the R^2 for a regressor).
- For each feature j (column of D):

For each repetition k in 1,…,k:

- Randomly shuffle column j of dataset D to generate a corrupted version of the data named $\widetilde{D}_{k,j}$.
- Compute the score $s_{k,j}$ of model m on corrupted data $\widetilde{D}_{k,j}$.

Compute importance $i_{k,j}$ for feature f_j defined as: (Pedregosa et al., 2011)

$$i_j = s - \frac{1}{k}\sum_{k=1}^{k} s_{k,J}$$

3.7) Analysis of multi-annual measurements with a similar technology towards predictive maintenancce

This chapter delves into the proof of concept, drawing from a series of real-world tests conducted during an exchange program in the UK Rail and Research Centre at the university of Birmingham, United Kingdom. It will explore the rationale behind the tests, detailing the setup, and the data which was collected. The outcomes of these tests from a critical piece of evidence, shedding light on the applicability and robustness of the proposed solutions in a live rail infrastructure setting.

This part of the work of my Ph.D. was used within the national Italian "MOST" project for the task 3.1.5 called Applicable and robust learning models from on-board monitoring for Predictive Maintenance. The project is still on going and focused on finding the temporal evolution of the track faults.

The MoniRail system was the monitoring system that was utilized to measure the data. A spin-out company from the Birmingham Centre for Rail Research and Education (BCRRE), MoniRail was founded lately. This system is robust and provides ongoing monitoring of the train and the track, as well as the ability to recognize and identify train and track-related performances, the system overview could be seen in the Figure 58 (MoniRail, 2019).

Figure 58 MoniRail system overview

3.7.1) MoniRail System

❖ Hardware

The MoniRail Inertial Measurement System is a low cost, robust, accurate system that can collect data from the bogie, body and axlebox levels.

MoniRail hardware consist of three parts:

1) Vehicle Body Logger (VBL), installed on secondary suspension (body)
2) Track geometry monitoring (XIS), installed on primary suspension (bogie - optional)
3) Axlebox system (AXS), installed on axlebox (optional)

a. The components, data communication, power and overall diagram of the hardware is shown in Figure 59.

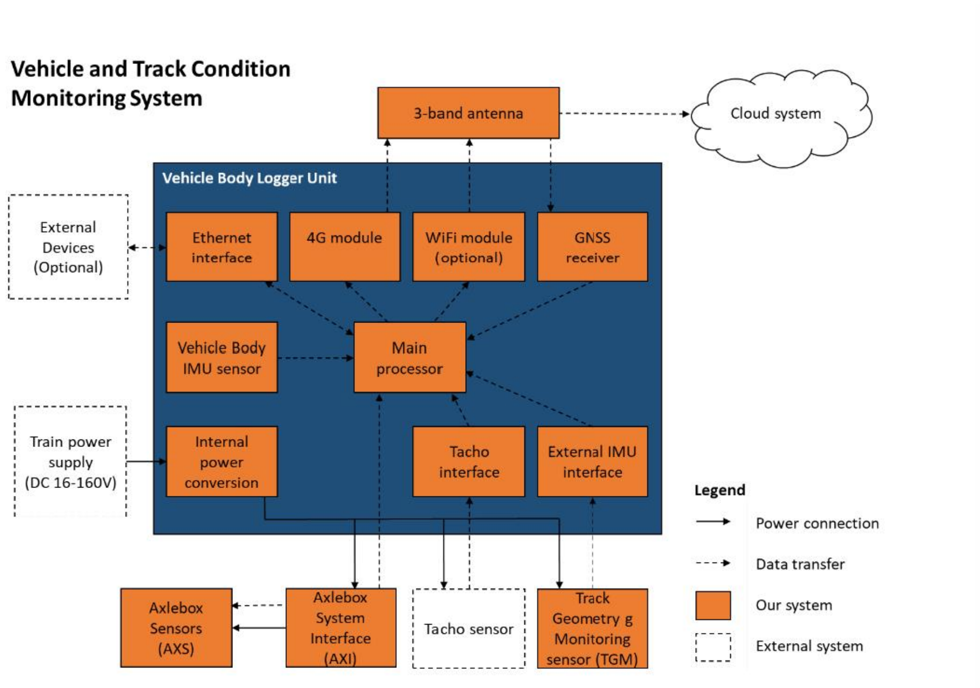


Figure 59 Overview of the hardware components

Examples of the hardware and their capabilities are shown in Figure 60.

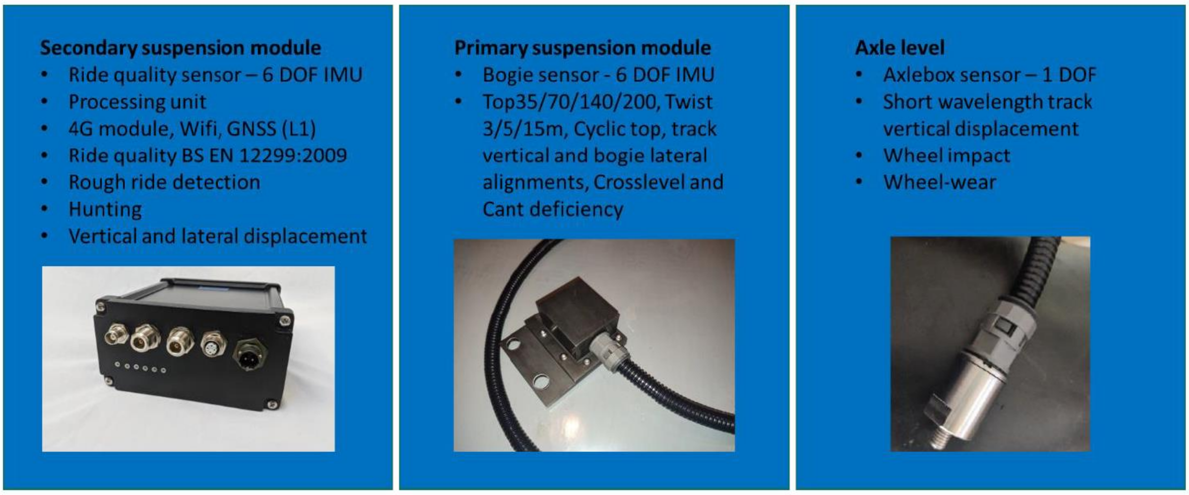


Figure 60 Body, bogie and axlebox modules

❖ Track Geometry Monitoring (TGM)

Figure 61 shows the location of a single TGM sensor unit on the vehicle bogie (main suspension). A VBL unit is necessary to use this sensor. Normally, the Vehicle Body Logger (VBL) unit is mounted on the secondary suspension, or carriage. The units are equipped with industrial-grade inertial sensors that are of excellent quality and cost-effectiveness. The enclosure is 45x45x28 mm and is designed to comply with BS EN 50155 and 30 g rms. It has an IP68 rating and can operate in temperatures ranging from -20˚C to 60˚C.

Figure 61 Used Track Geometry Sensor

The bogie unit can be mounted to existing fixing points, shown in Figure 62 (left), does not need to be centrally mounted as offsets can be accounted for during data processing. The sensor can be affixed to a clear and flat surface on the bogie using adhesive, Figure 62 (right).

Figure 62 Bogie mounted system on an in-service train (left) using existing fixing points - C374 (right) using adhesive - C170

The following track geometry and vehicle parameters are calculated:

- Vertical alignment, including wavelengths of 35m, 70m, and 140m (with the option to add other wavelengths) – (±1mm accuracy)
- Lateral alignment of the bogie, including wavelengths of 35m, 70m, and 140m (with the option to add other wavelengths) (±1mm accuracy)
- Cyclic Top 4m, 9m, 13m, and 18m (with the option to add other wavelengths)
- Twist 3m and 5m (with the option to add other wavelengths) – (accuracy depending on the distance between the wheelsets on the bogie – Twist 3m measurement accuracy on 2.6m wheelset spacing is above 70% compared to NMT, 5m is above 90%)
- Cross level (±1mm accuracy)
- Curvature
- Cant deficiency (line speed information is required)
- Bogie instability

❖ Software

The software system was designed and developed at BCRRE. Increased accuracy for reliable data is possible when Global navigation satellite system (GNSS) time synchronization is combined with geolocation solutions. An enhanced image of the location, time, and circumstances around an incident can be obtained by combining sensor data with open-source and location data.

During the initial monitoring phase, a baseline for track geometry is established by cross-checking in-service Track Geometry Monitoring (TGM) data with data collected by Track Recording Vehicles (TRVs). Ongoing comparisons with up to date TRV data enhance confidence in the in-service track condition information. The Vehicle Body Response Monitoring (VBL) data captures a fundamental 'signature' for each monitored vehicle's dynamic response, creating a database of these 'signatures' over time. This comprehensive database, including time, location, and accurate speed data, is developed through successive runs. The integration of Driver Reported Rough Rides (DRRRs) feeds into the system, allowing mapping to other system-captured data. This approach establishes a cross-referenceable database for all runs, facilitating the detection of changes from the initial signature and correlation with track conditions monitored by vehicles equipped with track sensors. Incorporating vehicle speed data ensures adjustments for speed differences during the 'signature' mapping process. The normalized data undergoes thorough analysis to unveil significant patterns and relationships with key predictors. The process of training a statistical model to estimate these predictors involves partitioning data into training, test, and validation sets. Employing various machine learning algorithms, both supervised and unsupervised, and comparing for accuracy yield reliable outputs.

Hence, the developed software empowers users to select various rides and directly visualize their measurements within the application. This includes an assessment of track conditions to check for thresholds and defects, evaluate ride comfort, and, with the availability of location data, allows users to refer to the map. This feature proves valuable in pinpointing the measurement location, aiding in understanding the causes behind issues like signal loss, especially in scenarios involving tunnels, etc.

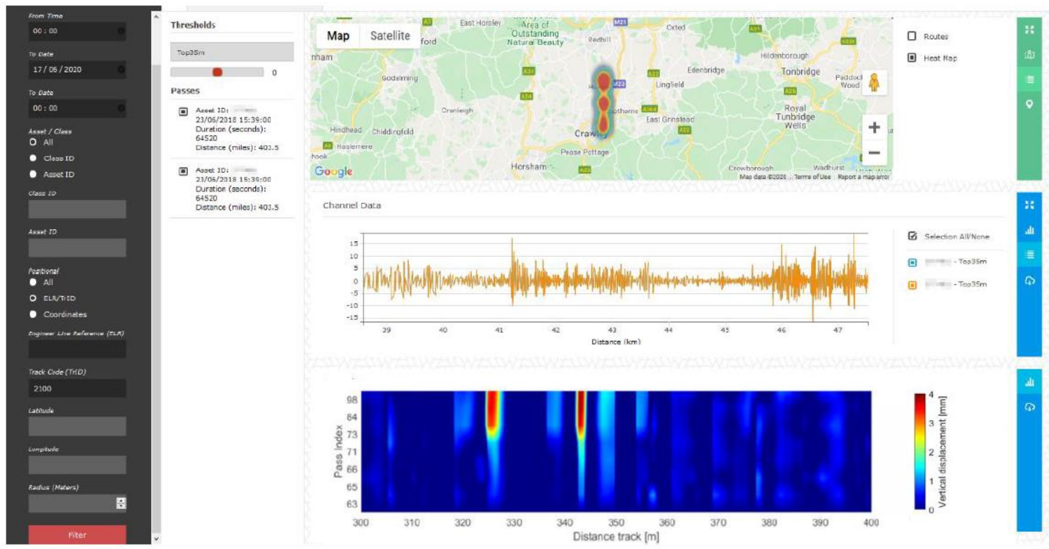


Figure 63 Example Visualisation of Track Condition Data

❖ Comparison to measurement fleet

As it was mentioned in the previous section, both passenger and freight vehicles come equipped with an inexpensive inertial measurement unit (IMU) (Entezami et al., 2022) This device can capture high-frequency accelerometer and gyroscope readings, which are promptly transmitted to a centralized data storage system.

Through a comparison of data obtained from on-board measurements and Network Rail high-speed data, it was reaffirmed that lateral acceleration alone cannot effectively reconstruct lateral irregularities. This underscores the significance of my research in addressing alignment issues. Another notable contribution of this work to my Ph.D. thesis lies in evaluating the proposed concept for predictive maintenance in real-life applications. It highlights the potential of on-board monitoring systems in studying the temporal evolution and discerning patterns of track irregularities.

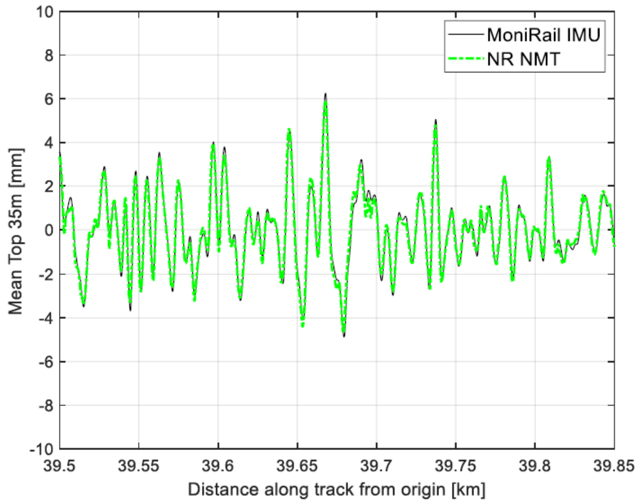


Figure 64 Data comparison, MoniRail IMU vs Network NMT

3.7.2) Degradation analysis

Following figures are some examples of this work which shows the comparison of measurement by IMU system, and the new train measurement of the Network Rail (NMT). It can show that both lines are quite aligned as well as case of lateral alignment detection of the track and the bogie. Therefore, it is a continuation of previous work by the on-board measurement with a similar technology for acquiring large amount of data and it is proved that on-board monitoring systems with overcoming the challenge of lateral alignment detection could replace the inertial platforms of diagnostic trains and make it possible to take a step toward predictive maintenance.

The Figures below shows some analysis of data on a specific section of the UK track as seen in Figure 65 for different parameters including:

- AI 35 – Bogie lateral alignment 35m filter
- Top 140 – Track Vertical alignment 140m filter (similar to D3)
- Curvature
- Speed

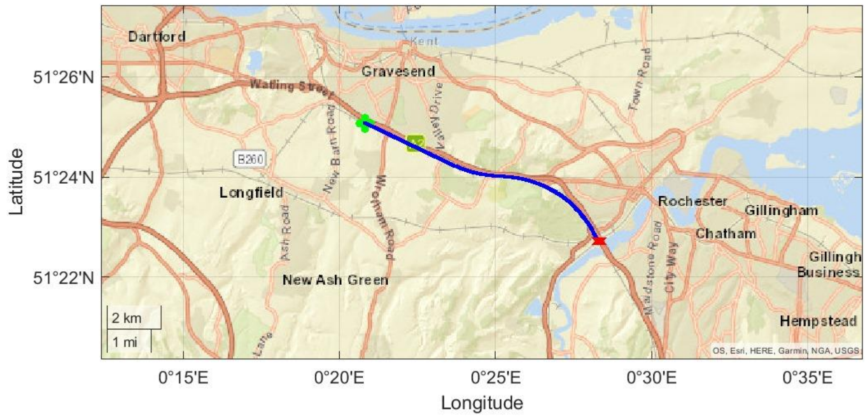


Figure 65 Map of chosen section

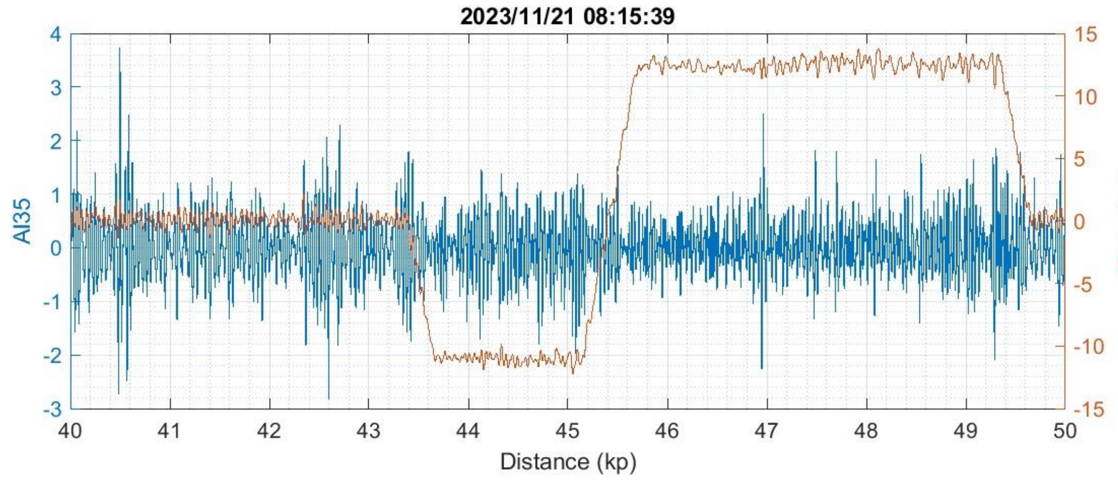


Figure 66 Curvature vs Alignment measured by IMU – D1

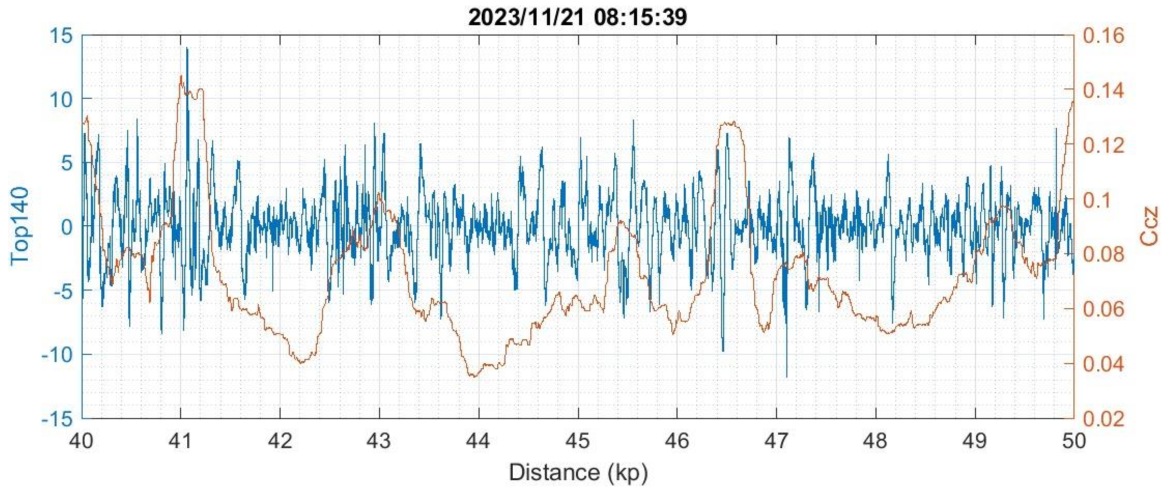


Figure 67 Ride comfort vs Vertical displacement - D2

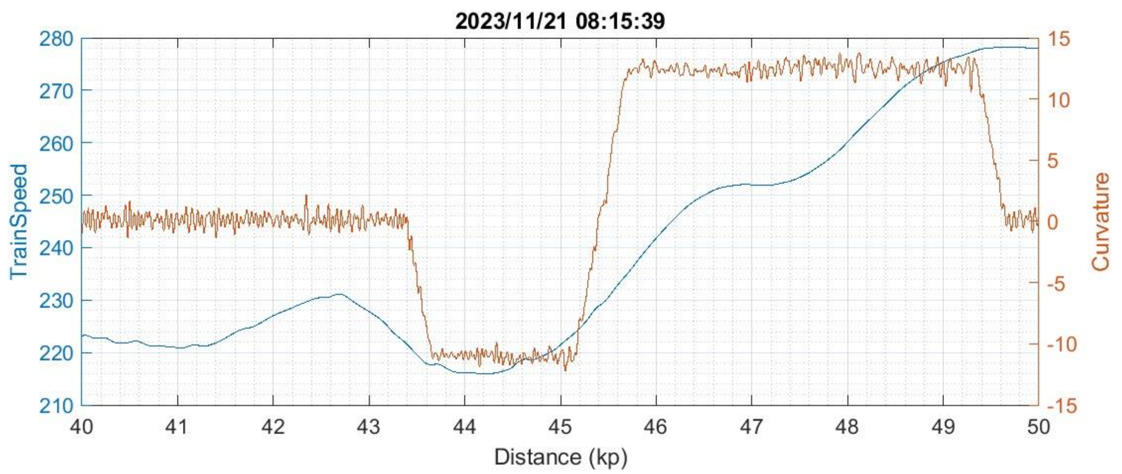


Figure 68 Train Speed vs Curvature

Comparison of IMU and NMT in the section between London and Ashford will be shown in the following figures to show the efficiency of the IMU sensor, and difference of the track condition for different years. The dark green shows the NMT data(reference), and the pink and light green are showing the evolution for 2022, and 2023, respectively.

Figure 69 Map of Chosen section (London - Ashford)

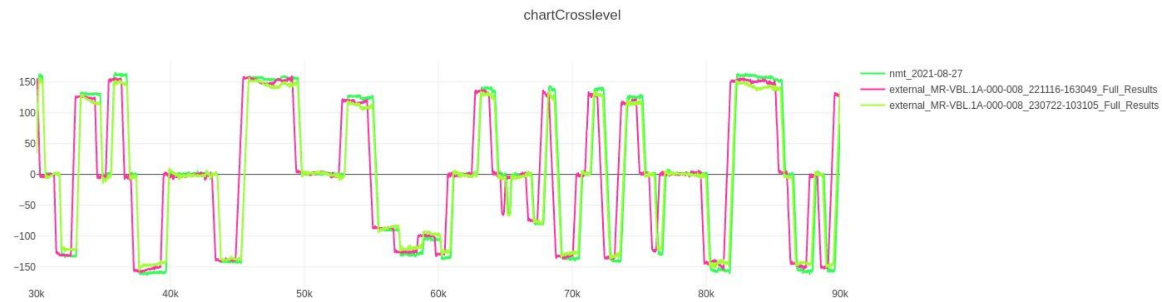


Figure 70 Comparision of Cross Level for NMT (Green) vs IMU (Pink)

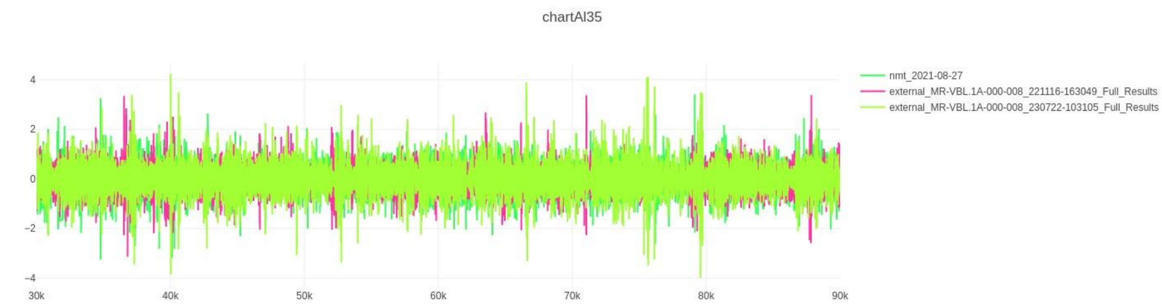


Figure 71 Comparision of Lateral Alignment for NMT (Green) and IMU (Pink)- D1

In Figure 72, one section of the track (55 _ 56 km) is zoomed and shown for better understanding. The dark green shows the NMT data(reference), and the pink and light green are showing the evolution for 2022, and 2023, respectively.

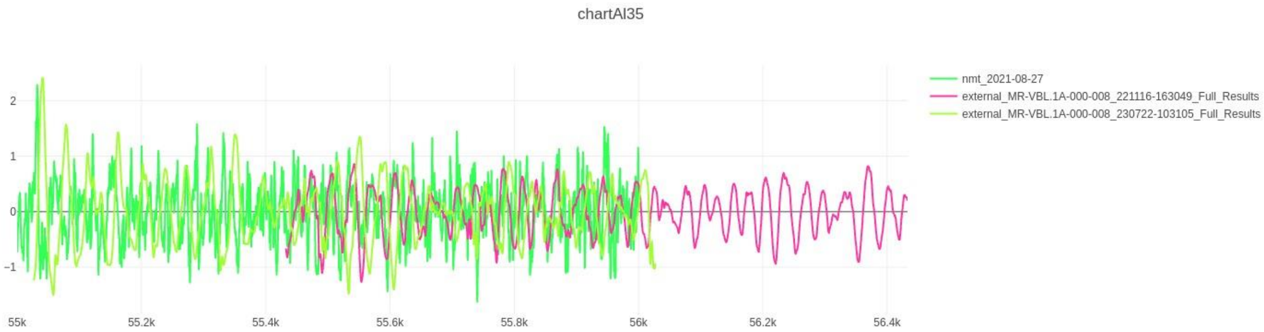


Figure 72 Comparision of lateral alignment for Zoomed section of the track between 55-56km

4) Results and discussion

4.1) Monitoring System Development

The first steps for the final goal of this thesis were focused on the development of a sensor system to facilitate track geometry monitoring as part of the EU Shift2Rail project Assets4Rail. The preliminary outcomes of the research, focusing on track geometry monitoring through systems suitable for installation on commercial trains, led to the identification of the System Breakdown Structure (SBS) and the requirements for an on-board sensor system dedicated to detecting wheel-rail relative positions. Stereo cameras and thermographic cameras emerged as the most suitable and promising technologies for implementation.

Algorithms for extracting necessary geometrical data from images were formulated, ensuring compatibility with both selected technologies. A methodology for estimating theoretical measurement accuracy was established, and the measurement accuracy for the devices under investigation was determined: ±0.7 mm for the stereo camera and ±1.4 mm for the thermal imaging camera. For instance, in track gauge measurement (as per the EN 13848-1:2019 standard), an accuracy of ±1.0 mm is mandated. The developed algorithm allows the calculation of measurement accuracy for various devices, providing valuable insights for potential system developers.

A test platform for real-time testing of sensing solutions was created, initially in the laboratory facilities of VGTU. This platform incorporated cameras connected to a power-efficient embedded AI computing device, Nvidia Jetson TX2, and a test rig. The data processing algorithm was validated using actual experimental data collected on the tracks of Lithuanian Railways (LTG).

Both the stereo vision and thermal imaging cameras employed in this study boast a robust sampling frequency of up to 60 Hz. The hardware utilized in our investigation is capable of processing data at impressive rates, reaching up to 120 frames per second, making it highly suitable for real-time applications. These selected technologies exhibit versatility, allowing for implementation in trains across a diverse range of high velocities, with the flexibility to choose devices featuring even higher sampling frequencies.

Addressing the mounting of sensors on the bogie frame, it was discerned that additional protective housings are imperative to shield the cameras from ambient conditions. These protective housings, when securely attached to the bogies via dampers, provide a resilient defence for the equipment, safeguarding it against overloads of up to 20g. Ensuring adequate airflow to these housings is crucial for optimal functionality. The designed system, with a modest power consumption of approximately 3W, effectively protects against environmental elements like dust, rain, and snow, ensuring reliable performance even in challenging low-temperature conditions.

To sum up, the computer-vision-based sensor system prototype was tested in regular train operating environment that was typical of the TRL 5. The hardware system worked well for data acquisition up to the maximum test speed of 100 km/h. Shock and vibrations have no effect on the obtained photos. The LEDs employed in the lighting system can produce light with a colour spectrum that is extremely like the sun spectrum, the lighting system was assured to provide enough lighting in tunnels, outdoors, and in various environmental circumstances. The primary issue that arose throughout the validation tests was a minor loss of frames from the obtained images.

This resulted from the USB cable used to connect the stereo camera to the processing unit—a laptop that was inside the car in the validation test—being overly long. As a result, the processing unit for the sensor system should be mounted beneath the vehicle in the final prototype. The measurement uncertainty can be reduced by swapping out the existing camera for one with a tighter focusing distance and a narrower field of vision, if needed.

The convolutional neural network (CNN) algorithm for monitoring the lateral displacement of train wheels on the rails, known as LDWR, underwent enhancements and validation through numerical experiments. The computed LDWR was compared with the lateral alignment obtained from a commercial track geometry inspection system. The comparison revealed a correlation between LDWR and lateral alignment, emphasizing the necessity of combining LDWR with wheel acceleration for an accurate derivation of lateral alignment. The standalone use of LDWR or wheel acceleration alone is insufficient for this purpose.

4.2) Simulation

The running dynamics of a virtual replica of the Track Recording Vehicle were simulated using multi-body simulation. The results indicated a strong correlation, especially on straight tracks, between the lateral wheelset displacement values obtained by integrating the axle-box accelerations and those directly obtained from Simpack. This correlation was also observed with the sum of the measured left and right alignment. With this correlation established, the next phase of the research aims to develop an algorithm for calculating rail alignment based on the right and left lateral dynamic wheelset readings from the stereo-camera and the lateral displacement values derived from double-integrating the axle-box accelerations. However, the results of the next chapters are based on the Gensys simulation as explained earlier in the methodology.

Some key Simpack simulation results are shown in Figure 73. The LDWR of the left wheel of the leading wheelset of the selected coach running over the same portion of track are simulated for the following different conditions, all with a speed of about 27 m/s: (Kaviani et al., 2023)

a. 1:20 rail inclination, single-carbody model;
b. 1:20 rail inclination, 3-carbody model, middle coach;
c. 1:30 rail inclination, 3-carbody model, middle coach;
d. 1:40 rail inclination, 3-carbody model, middle coach.

The lateral motion in case a) is clearly unstable on straight track, with a characteristic frequency of approx. 0.6 Hz. The lateral oscillation always spans all the available flangeway clearance, the amount of which is evident from the central part of the plots that refer to the negotiation of a low-radius curve. In case b) the lateral oscillation related to the Low-Frequency Body Hunting (LFBH) mode is present but reduced, and its frequency, as observed over longer stretches of straight track, is also reduced to about 0.4 Hz on average. In cases c) and d), such lateral oscillation is no longer visible. Without the otherwise dominant LFBH, the effects of track geometry irregularities appear more clearly.

The corresponding eigenvalue analyses are in line with the above: in case a) there is one unstable mode (-0.04 damping factor) with an undamped natural frequency of 0.41 Hz and a modal shape corresponding to LFBH, in case b) the same mode remains, this time with a positive damping factor (+0.14), and a frequency of 0.35 Hz.

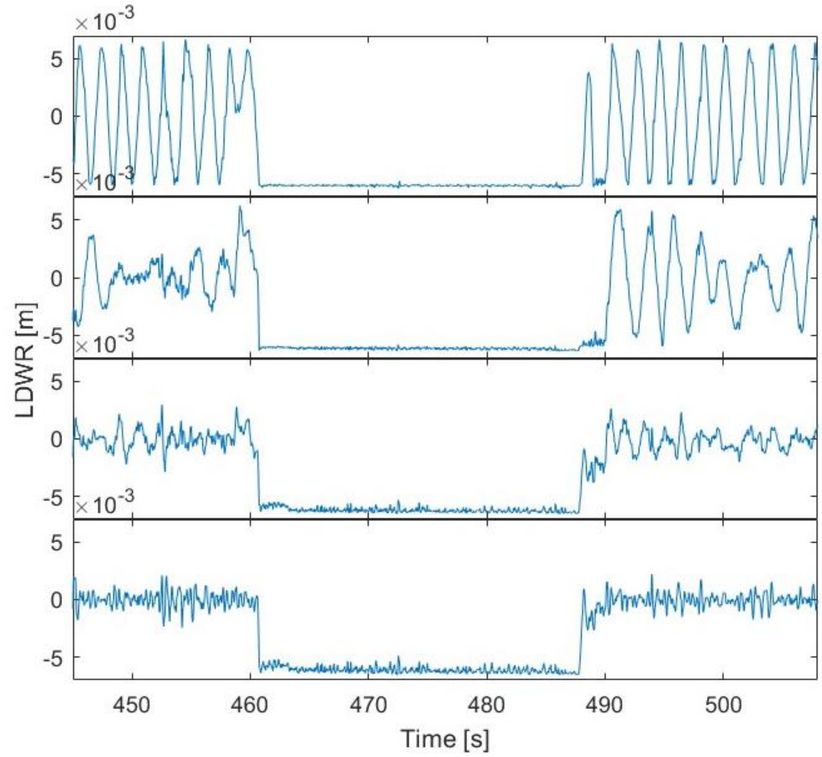


Figure 73 Comparison of the lateral displacements of the left wheel of the leading wheelset for: a) 1:20 rail inclination, single-carbody model; b) 1:20 rail inclination, 3-carbody model, middle coach; c) 1:30 rail inclination, 3-carbody model, middle coach; d) 1:40 rail inclination, 3-carbody model, middle coach.

Further simulations with varying coupling parameters highlight that, while those for which literature values are available (drawgear stiffness, buffer stiffness/damping) do not have a large effect on the LFBH modes, the buffer plate contact friction coefficients do. The final chosen values represent conditions which were both considered plausible and providing good LFBH mode damping.

4.3) Rail inclination influence on the lateral irregularity estimation

While analysing the data output from the Gensys simulation for model creation and development, a notable finding emerged as an initial result. This interesting finding revolves around the impact of rail inclination on the model.

Figure 74 illustrates the initial validation attempt, which resulted in failure. This outcome underscores a crucial finding: the model's testing accuracy is contingent upon training with the same inclination. For instance, the model was trained with an inclination of 1:20 at a speed of 130 km/h and tested on a similar track and speed but with an inclination of 1:30. The figure vividly portrays the failure in detecting lateral irregularities, manifested by a substantial disparity between real and predicted values, indicative of notably low accuracy.

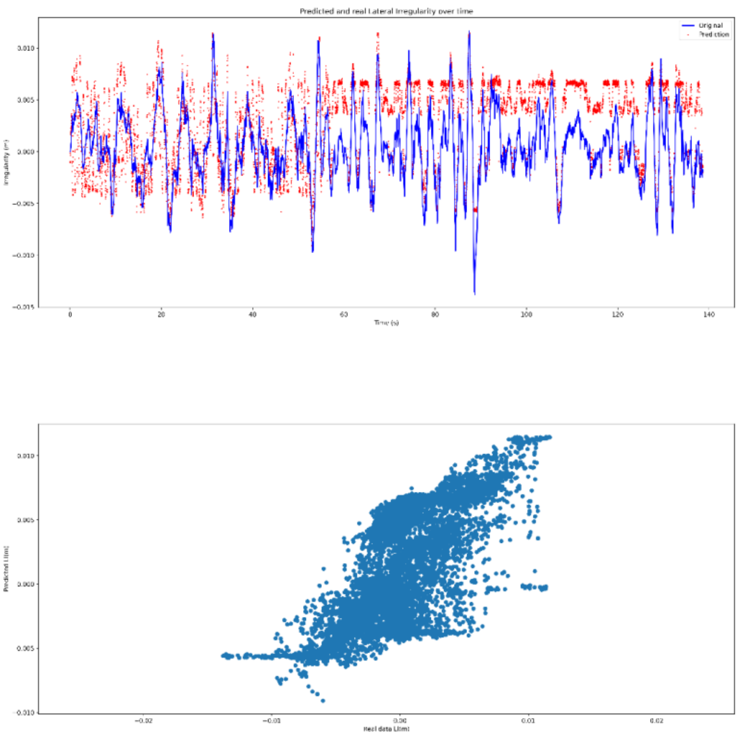


Figure 74 Failed trial

Hence, it was observed that the model exhibited optimal performance when trained and tested under consistent rail inclinations. The accuracy of the model proved to be dependent on maintaining a uniform inclination between the training and testing phases. Therefore, when the inclination remains constant during both phases, the model can effectively identify lateral irregularities even in the presence of varying track characteristics and speeds.

Consequently, a categorization procedure was started, dividing the data into three different groups according to different inclinations. Every group had distinct phases of training and testing, with consistent inclinations within the respective group. Using Random Forest

Regression, the following figures show one group's best and worst performances. This algorithm was selected because it performed better at inclinations of 1:20, 1:30, and 1:40 than Polynomial Regression and Support Vector Regression. The subsequent sections of this thesis will predominantly present results based on Random Forest Regression owing to its heightened accuracy and efficiency.

It should be noted that similar special limitations have not been found for other variables than inclination, like different characteristics for different tracks. Therefore, it is possible to test the case with the model which is trained on another track with different characteristics and speed but the same inclination.

4.4) Data Analysis and Visualizations

The sample model, trained with an inclination of 1 in 20 and a speed of 130 km/h, was used to test the different case studies. Among the case studies, the highest accuracy was achieved for the ST1 track at a speed of 160 km/h, while the ST1 track at 210 km/h had the worst performance in this group. The figures below show the accuracy results for the 160 km/h and 210 km/h case studies for the sections of whole track, straight and curve, respectively.

Table 16 Case Studies Characteristics

Case Studies	Speed (km/h)	Inclination	Wheel Profile	Rail Profile	Curves	Track Length
ST1	130,160,210	1:20, 1:30, 1:40	EN S1002	60 EI	2	13.276 km
ST2	50,80,120	1:20, 1:30, 1:40	EN S1002	60 EI	Several	7725m

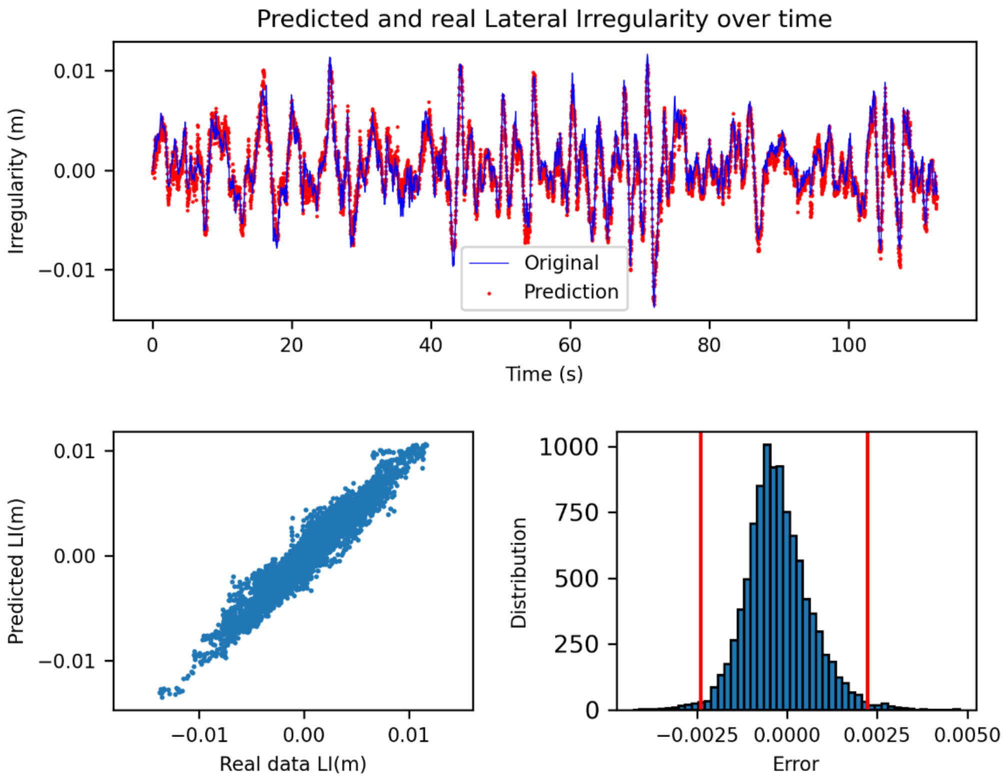


Figure 75 Predicted and real Lateral Irregularity over time, and error distribution for the whole track tested on the Velim Track with the Inclination of 1:20, 160 km/h

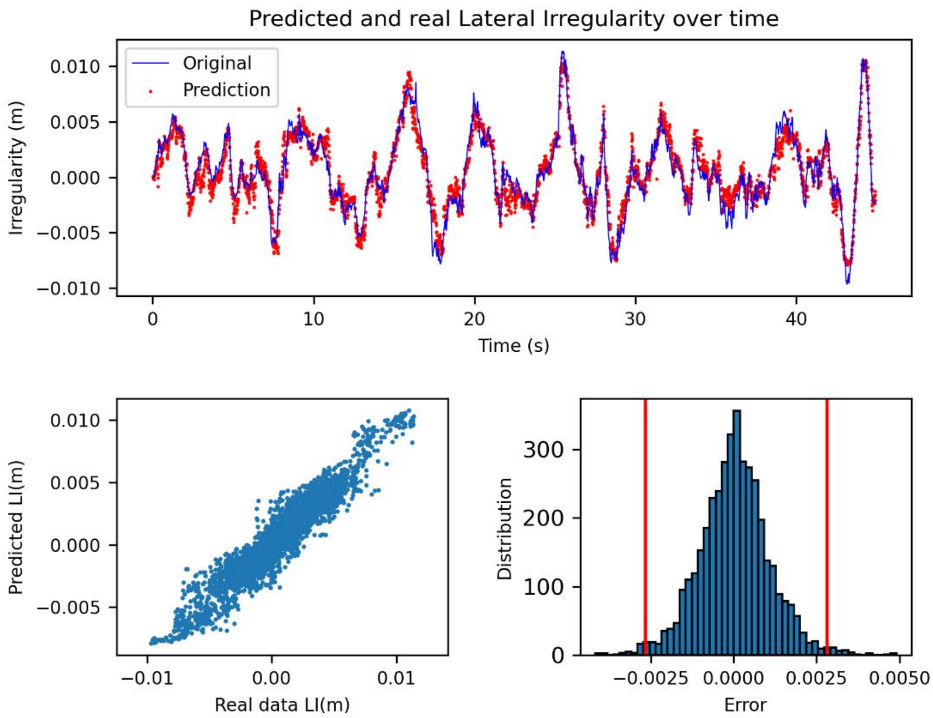


Figure 76 Predicted and real Lateral Irregularity over time and error distribution for the Straight track tested on the Velim Track with the Inclination of 1:20, 160 km/h

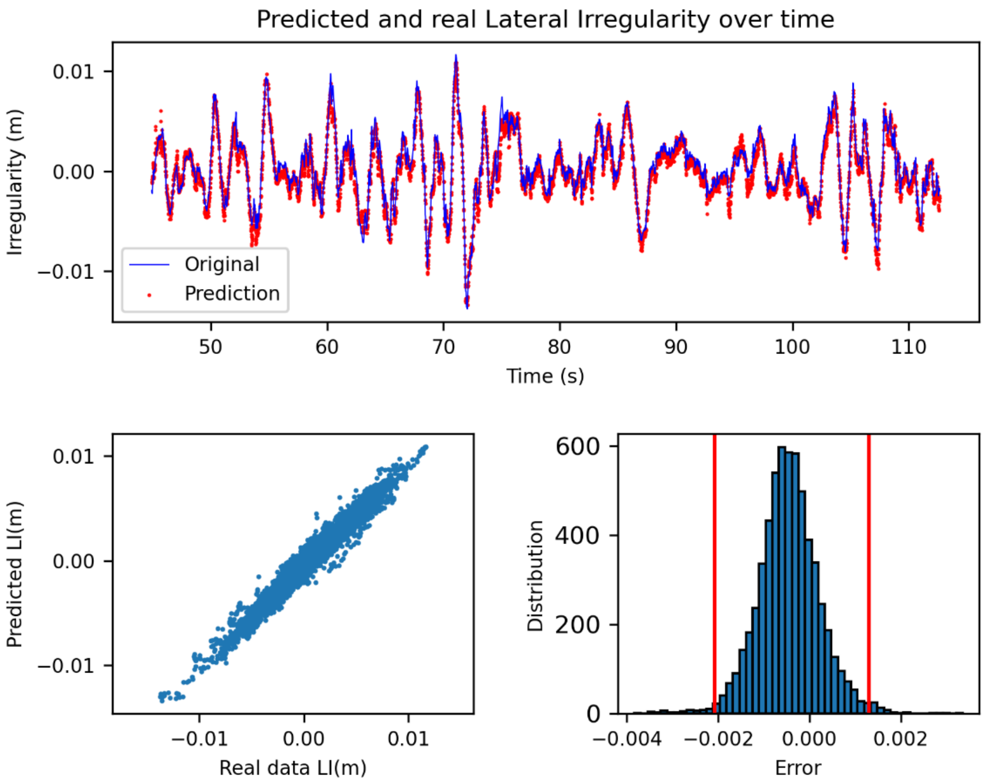


Figure 77 Predicted and real Lateral Irregularity over time and error distribution for the Curve track tested on the Velim Track with the Inclination of 1:20, 160 km/h

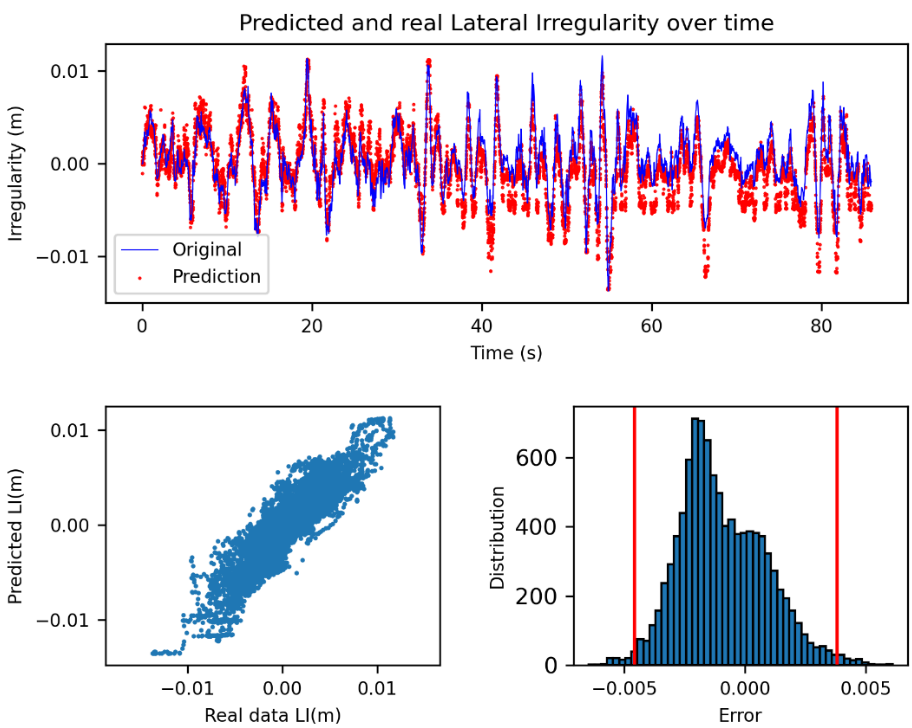


Figure 78 Predicted and real Lateral Irregularity over time and error distribution for the whole track tested on the Velim Track with the Inclination of 1:20, 210 km/h

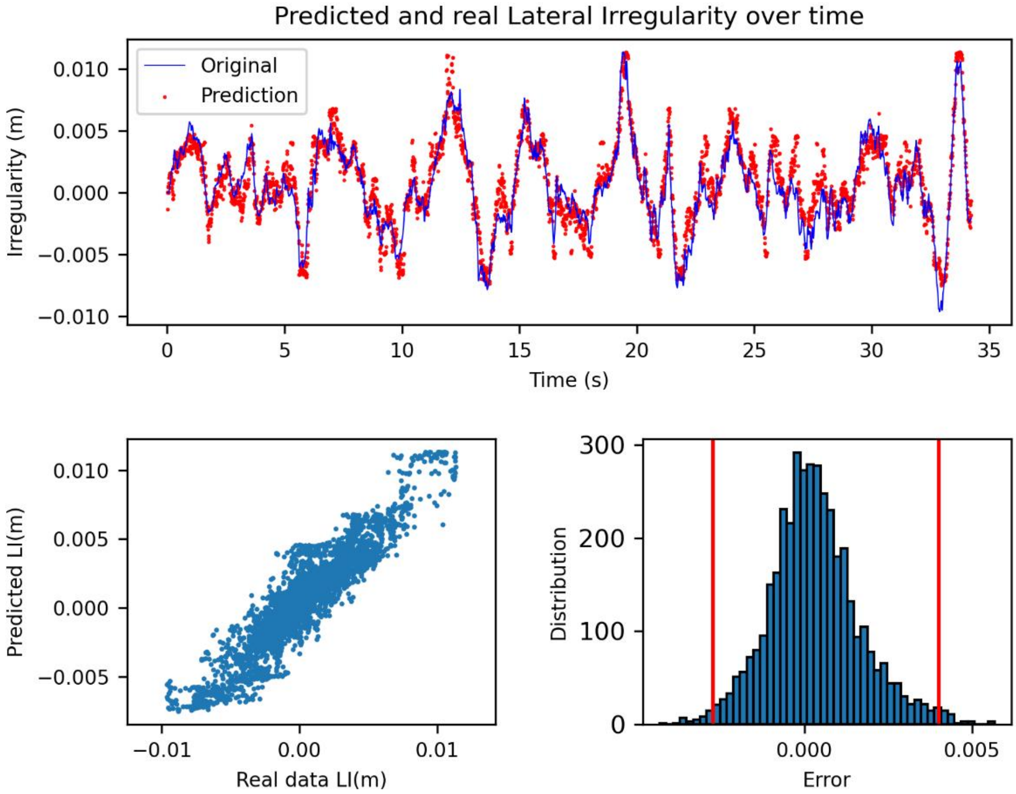


Figure 79 Predicted and real Lateral Irregularity over time and error distribution for the Straight track tested on the Velim Track with the Inclination of 1:20, 210 km/h

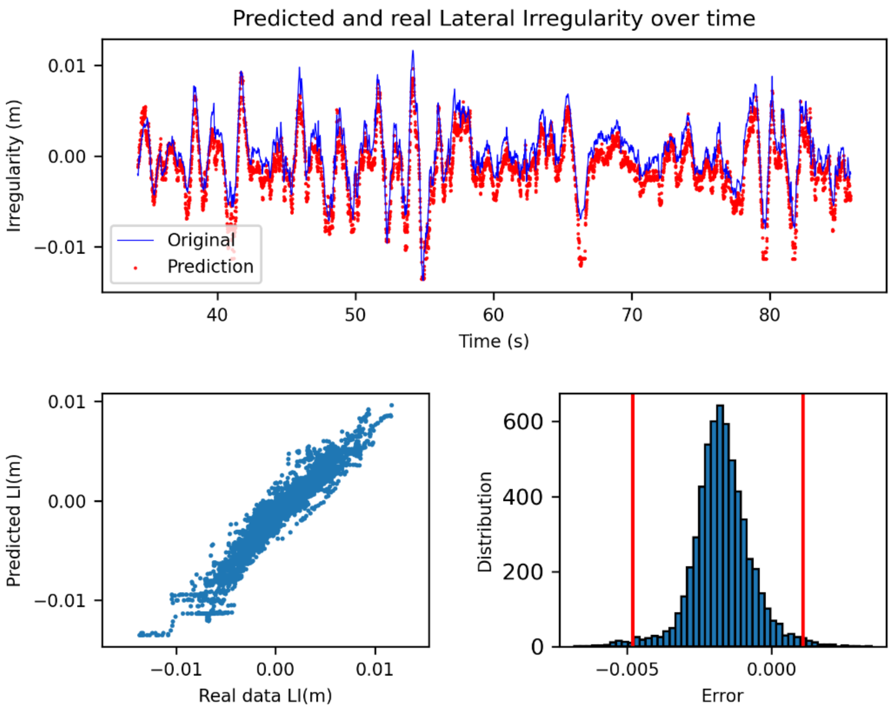


Figure 80 Predicted and real Lateral Irregularity over time and error distribution for the Curve track tested on the Velim Track with the Inclination of 1:20, 210 km/h

The data analysis presented in the figures above revealed the successful identification of lateral irregularity for group 1, which is the one with the inclination of 1 in 20; as it has the best accuracy for speed of 160 km/h, with R2 scores of 0.92, 0.90, and 0.94 for the entire track, straight track, and curve track, respectively. The accuracy levels were found to be high, but as the speed of the train increased, the accuracy of the curve began to drop due to greater lateral displacement. The accuracy of the entire track also decreased with higher speeds, while the accuracy of the straight track remained almost constant. Therefore, it is recommended to focus on overall track accuracy while considering the speed fluctuation.

Similar accuracy-speed trends were observed for the other groups with inclinations of 1:30 and 1:40, suggesting that the speed limit for the functionality of this study should be around 210 km/h. Among the tested models, the ST1 with a speed of 160 km/h exhibited the best accuracy with R2 of 0.92 (92% prediction accuracy), and an RMSE of 0.00093. The other tested models in this group of 1 in 20 inclinations, also showed good accuracy levels, except for the ST1 with a speed of 210 km/h, which had the worst scenario with R2 of 0.80 (80% accuracy) and an RMSE of 0.001514. The ST2 had a maximum accuracy of R2=0.91 at a speed of 50 km/h and a minimum accuracy of R2=0.78 at a speed of 120 km/h.

Analysing all the error distribution of the prediction on the histograms with the percentiles of 1 and 99, is illustrating that the 99% of the errors are in average distributed between ±2 mm, which is acceptable for our case study.

The convergency issues was also into consideration as an answer to the lower accuracy of higher speeds, so it was checked, and the result was the same. Therefore, the results are countable.

After aggregating all data and effective features, the main model was created and trained with a test size of 0.07. The results were satisfactory for both Random Forest Regression (RFR) and Polynomial Regression (PR). Here below, in the Figure 81, one example of the error distribution of each model could be seen, and still the error is distributed with the same values as before for both different algorithms. The total average accuracy is 0.921 (R2 Score) for RFR and 0.926 for PR.

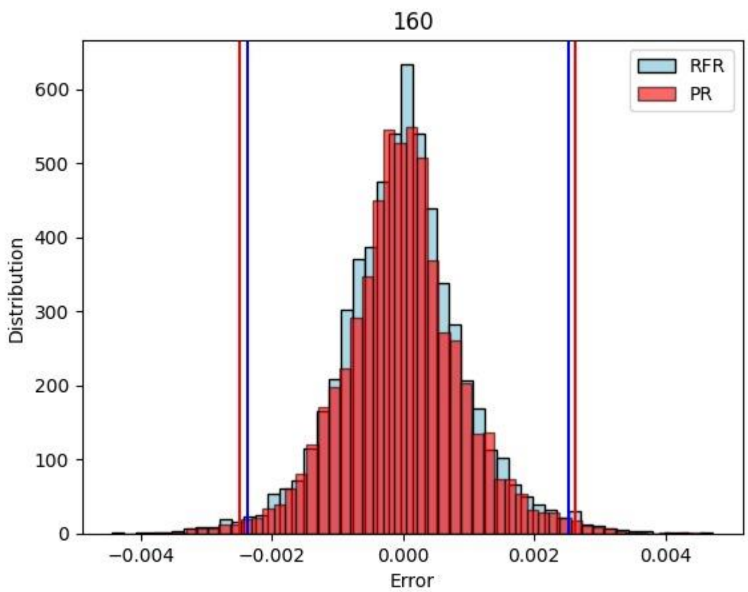
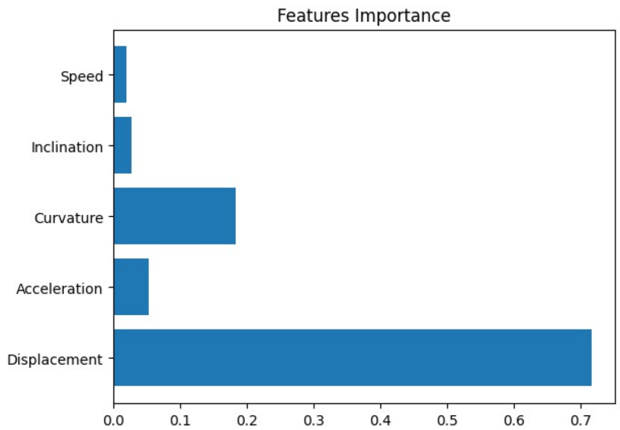


Figure 81 Error Distribution for the speed of 160 km/h tested RFR and PR

4.4) Sensitivity Analysis Result

For sensitivity analysis, feature importance analysis was conducted to comprehend the relative significance and contribution of each input variable in the model's predictive performance. This analysis results for the random forest model reveal that the most significant impact on the model is attributed to lateral displacement, while the least crucial feature among the considered variables is associated with lateral acceleration. This outcome, though surprising, aligns with the initial notion that there is no discernible relationship between lateral acceleration and lateral irregularity in the track due to the variables between the wheel and the rail in the lateral direction. Following lateral displacement with wheel rotation (LDWR), curvature emerges as the next most influential parameter, with the remaining variables contributing to a lesser extent.

It is significant to note that the tests were conducted both with and without consideration of conicity. Interestingly, in the absence of conicity, acceleration assumes a more prominent role, taking the contribution that would otherwise be attributed to conicity. Strikingly, even without measuring conicity, the detection of irregularities remains feasible. The impact of conicity is observed as an enhancement in the performance of random forest regression (RFR), while its effect on polynomial regression (PR) is comparatively minimal. The figures below provide visual representations of these findings.



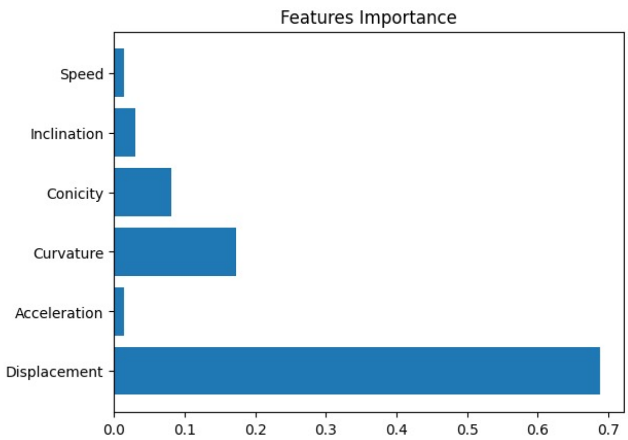


Figure 82 Feature Analysis of Random Forest model without Conicity (Left), with Conicity(Right)

4.5) Accuracy Analysis Considering Rail Inclination, Speed and Algorithmic Impact

The comparison of accuracies between Random Forest Regression (RFR) and Polynomial Regression (PR) was conducted based on speed and rail inclination, as illustrated in Figure 83, and Figure 84. The findings reveal a trend where increasing speed corresponds to a decrease in accuracy. Moreover, the rail inclination also influences accuracy, with higher inclinations contributing to greater stability in wheel displacement on the rail.

Specifically, the inclination of 1:40 exhibits the highest accuracy ranges, followed by 1:30, while 1:20 demonstrates the lowest accuracy. Although the accuracy drops at inclinations of 1:40 and 1:30 remains acceptable, exceeding 80%, the inclination of 1:20 falls below 75%, warranting further investigation in future research. The overall average R2 Score for both algorithms is approximately 0.88 in the absence of conicity. In the presence of conicity, the R2 Score is 0.96 for RFR and 0.86 for PR."

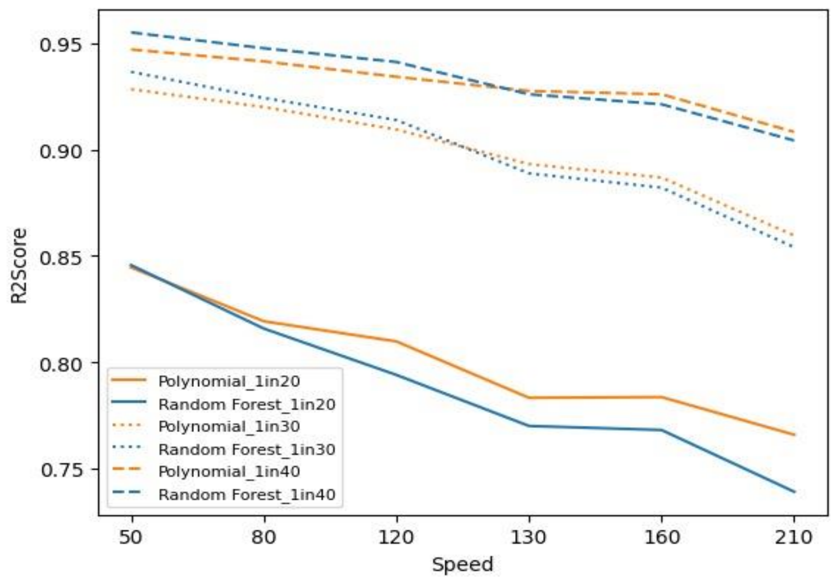


Figure 83 Accuracies based on Speed and Inclination without Conicity

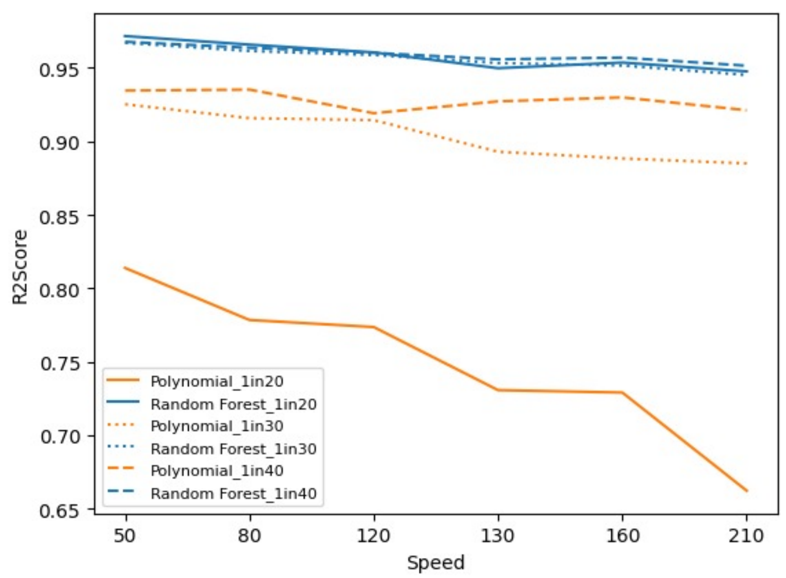


Figure 84 Accuracies based on Speed and Inclination with Conicity

Since most of the results of this thesis was focused on the random forest regression algorithm, in the following, the figure specifically for illustrating accuracies based on the rail inclination and speed considering conicity could be seen.

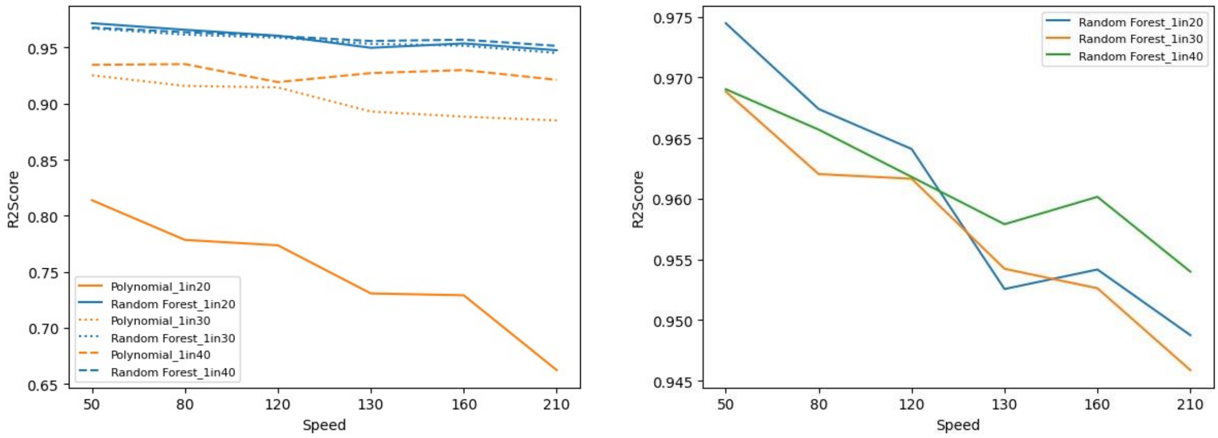


Figure 85 Accuracies based on Speed and Inclination with conicity for Random Forest Regression Algorithm

5) Conclusion

In conclusion, this doctoral thesis has demonstrated the feasibility of detecting lateral irregularities through the lateral displacement of the Wheel/Rail with their correlation of 70%. This discovery opens opportunities for the development of an on-board monitoring system, eliminating the need for inertial platforms and enabling frequent monitoring. The substantial volume of data collected daily from various commercial trains holds the potential for advancing predictive maintenance practices.

After testing several supervised learning algorithms, the results showed that random forest regression and polynomial regression both had excellent accuracy, proving their suitability for lateral irregularity detection based on the wheel's lateral displacement. Support vector machines, on the other hand, proved inadequate for this purpose.

The crucial features essential for constructing the model were ranked in the following order: Lateral displacement, curvature, conicity (if applicable), acceleration, and speed. This finding is noteworthy, considering that much of the previous research has primarily concentrated on reconstructing irregularities based on acceleration. It is imperative to highlight that the inclination played a significant role in both simulations and the development of algorithms, and its value must remain consistent for both the training and testing phases. While rail inclination might not initially appear to be a significant factor in running dynamics assessments, this study reveals its impact on the behaviour under normal running conditions, as identified by the Random Forest and Polynomial Regression models.

The reduction in prediction accuracy for lateral alignment with higher rail inclination, especially in the case of a 1:20 inclination, is attributed to heightened lateral oscillation amplitudes caused by low-frequency body-hunting motions. The exceptionally low equivalent conicity of the S1002 wheel, coupled with the 1:20 rail inclination, results in these motions potentially becoming unstable unless the model parameters are carefully selected.

As the measurements on the actual Track Recording Vehicle (TRV) running on a 1:20 rail did not exhibit signs of instability, the model underwent repeated modifications to comprehend the underlying causes. Among the contact conditions, the primary influencing factors for this instability were identified as the equivalent conicity (notably low with 1:20 rail inclination) and the creep coefficients (whose values were scrutinized and verified). The crucial suspension parameters, particularly those determining the secondary yaw stiffness, underwent thorough examination, revision, and confirmation of their effects. The introduction of two additional vehicles to the trainset, with appropriately adjusted coupling parameters, further contributed to mitigating the instability.

Given that the final model teeters on the verge of instability at various speeds with a 1:20 rail inclination, incorporating more than one vehicle in the model might be a necessary consideration if the normal running behaviour of the selected vehicle on 1:20 rails is to be accurately modelled for the preliminary training of Machine Learning algorithms for lateral track geometry detection.

This finding is expected to support improved modelling assumptions when using Multi-Body Simulation (MBS) for the generation of training datasets for condition monitoring developments. It also stresses the importance of using measured datasets for definitive ML training and limiting simulation to preliminary screening activities.

The impact of speed on accuracy was evident in the analysis of curved sections, with higher speeds resulting in decreased accuracy for both algorithms and across all inclinations. The algorithm developed in this study demonstrated reliability up to a speed of 210 km/h.

Finally, with the continuous measurement of substantial daily data through on-board monitoring systems and the utilization of the developed algorithm of this study, there exists the potential to find the temporal evolution of track defects. This comprehensive dataset allows for the identification of patterns conducive to the establishment of a predictive maintenance model, and some preliminary result to prove the need and efficiency of this work was mentioned in the previous chapter of analysis of multi-annual measurements.

Future work

The relationship between conicity and acceleration needs to be discovered more and explained in a better way.

Meanwhile, the activities of the ‘MOST’ project to understand the evolution of the track faults is ongoing, therefore the continuation of task 3.1, with implementation of the sensor system to run the test, and use of the developed model in this work for the on-board monitoring data towards predictive maintenance goal is planned.

References

A.B. Jung, K. Wada, S. Tanaka, & C. Reinder, et al. (2020). *Imgaug -https://github.com/aleju/imgaug* .

Alemazkoor, N., Ruppert, C. J., & Meidani, H. (2018). Survival analysis at multiple scales for the modeling of track geometry deterioration. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *232*(3).

An, D. , K. N. H. , & C. J. (2015). Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. *Reliability Engineering & System Safety*, *133*, 223–236.

Andrade, A. R. , T. P. F. (2011). Uncertainty in rail-track geometry degradation: Lisbon-Oporto line case study. *Andrade, A.R., Teixeira, P.F.*, *137*(3), 193–200.

Andrade, A. R., & Teixeira, P. F. (2013). Hierarchical Bayesian modeling of rail track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 364–375.

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

Andrade, A. R., & Teixeira, P. F. (2018). Assessing temporary speed restrictions and associated unavailability costs in railway infrastructure. *International Journal of Civil Engineering*, *16*(2), 219–228.

Arasteh Khouy, Schunnesson, H., Juntti, U. , N., & Larsson-Kraik, P. (2013). Evaluation of track geometry maintenance for heavy haul railroad in Sweden-A Case Study. *Evaluation of Track Geometry Maintenance for Heavy Haul Railroad in Sweden-A Case Study*, *228*(5), 496–503.

Asano, T., & Katoh, N. (1996). Variants for the Hough transform for line detection. *Computational Geometry*, *6*(4), 231–252. https://doi.org/10.1016/0925-7721(95)00023-2

Audley, M. , A. J. (2013). The effects of tamping on railway track geometry degradation. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *227*(4), 376–391.

B. Ripke et al. (n.d.). *Report on track/switch parameters and problem zones, technical report Deliverable D4.1 of the IN"SMART project*.

B. Xiao, & H. Wu, Y. W. (2018). Simple Baselines for Human Pose Estimation and Tracking. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 472–487.

Baqersad, J., Poozesh, P., Niezrecki, C., & Avitabile, P. (2017). Photogrammetry and optical methods in structural dynamics – A review. *Mechanical Systems and Signal Processing*, *86*, 17–34. https://doi.org/10.1016/j.ymssp.2016.02.011

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). *BlazePose: On-device Real-time Body Pose tracking*.

Bergquist, B., & Söderholm, P. (2015). Data analysis for condition-based railway infrastructure maintenance. *Quality and Reliability Engineering International*, *31*(5), 773–781.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification And Regression Trees*. Routledge. https://doi.org/10.1201/9781315139470

Bruni, S., Kulkarni, R., Qazizadeh, A., Berg, M., Di Gialleonardo, E., Facchinetti, A., & De Rosa, A. (2021). Monitoring of lateral and cross level track geometry irregularities through onboard vehicle dynamics measurements using machine learning classification algorithms. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *235*(1), 107–120. https://doi.org/10.1177/0954409720906649

Budai-Balke. (2009). *Operations research models for scheduling railway infrastructure maintenance*.

Caetano, L. F., & Teixeira, P. F. (2016). Predictive Maintenance Model for Ballast Tamping. *Journal of Transportation Engineering*, *142*(4). https://doi.org/10.1061/(ASCE)TE.1943-5436.0000825

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70–79. https://doi.org/10.1016/j.neucom.2017.11.077

Cárdenas-Gallo, I., S., C. A., M., G. A., B., & M. A., & A.-T. R. (2017). An ensemble classifier to predict track geometry degradation. *Reliability Engineering & System Safety*.

Cárdenas-Gallo, I., S., C.A., M., G.A., B., & M.A., A.-T. (2017). An ensemble classifier to predict track geometry degradation. *Reliability Engineering & System Safety*, *161*, 53–60.

CEN. (2010). *EN 13306: Maintenance terminology*.

CEN/TC 256/WG10. (n.d.). Standardization of requirements for running characteristics of railway vehicles on track, and test methods for the running behaviour. *VEHICLE/TRACK INTERACTION*.

Chen, R., Zhang, P., Wu, H., Wang, Z., & Zhong, Z. (2019). Prediction of shield tunneling-induced ground settlement using machine learning techniques. *Frontiers of Structural and Civil Engineering*, *13*(6), 1363–1378. https://doi.org/10.1007/s11709-019-0561-3

Choi, I. I. ; U. J. ; L. J. S. ; C. H. (2013). The influence of track irregularities on the running behavior of high-speed trains. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *227*(1), 94–102.

Commission Implementing Regulation (EU). (2019). *Commission Implementing Regulation (EU)*. 766.

Commission Regulation of Rail system in the European Union. (n.d.). *Commission Regulation (EU) No 1302/2014 of 18 November 2014 concerning a technical specification for interoperability relating to the 'rolling stock locomotives and passenger rolling stock' subsystem of the rail system in the European Union.*

D. Luvizon, D. Picard, & H. Tabia. (2018). 2D/3D Pose Estimation and Action Recognition using Multitask Deep Learning. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5137–5146.

Daniel Pyke. (2019). Rail Inclination. *Rail 101*.

diiselrong. (2018). *Train wheel on a rail 2*. Https://Youtu.Be/6oEkVbhT_T8.

EN 13848-1. (2019). *EN 13848-1: Railway applications - Track - Track geometry quality - Part 1: Characterization of track geometry*.

EN 14363. (2005). *Railway applications – testing for the acceptance of running characteristics of railway vehicles – testing of running behaviour*.

EN 14363:2016. (2016). *"Railway applications - Testing and Simulation for the acceptance of running characteristics of railway vehicles - Running Behaviour and stationary tests."*

Entezami, M., Authors: Roberts, C., Stewart, E., Yeo, G. , Peinado Gonzalo, A., & Hayward, M. , A. P. (2019). Continuous railway track monitoring using passenger trains. *World Congress on Railway Research*.

Entezami, M., Gonzalo, A. P., Roberts, C., Weston, P., Yeo, G., Horridge, R., Stewart, E., Hayward, M., & Rippin, M. (2022). Observations on track evolution from onboard inertial measurements. *2022 UKACC 13th International Conference on Control (CONTROL)*, 18–23. https://doi.org/10.1109/Control55989.2022.9781456

Esveld, C. (2001). *Modern Railway Track*.

European Committee for Standardization (CEN). (2019). Railway applications - Track - Track geometry quality - Part 1: Characterization of track geometry. *EN 13848-1*.

Gao, W., Hu, L., Zhang, P., & He, J. (2018). Feature selection considering the composition of feature relevancy. *Pattern Recognition Letters*, *112*, 70–74. https://doi.org/10.1016/j.patrec.2018.06.005

Gensys. (n.d.). *Gensys simulation for Rail Road applications*. Retrieved June 20, 2022, from https://www.gensys.se/doc_html/analyse_sum.html#Mainmenu)

Gerlici, J., Domin, R., Cherniak, G., & Lack, T. (2018). Calculated estimation of railway wheels equivalent conicity influence on critical speed of railway passenger car. *MATEC Web of Conferences*, *157*. https://doi.org/10.1051/matecconf/201815703006

H. Abbasian, J. Park, S. S., & S. Rella. (2020). *Speeding Up Deep Learning Inference Using TensorRT*. https://developer.nvidia.com/blog/speeding-up-deep-learning-inference-using-tensorrt/

Hakan Guler. (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.

He, Q., Li, H. , B., D., P. D. P., & Hampapur, A. (n.d.). Track geometry defect rectification based on track deterioration modelling and derailment risk assessment. *Journal of the Operational Research Society*, *66*(3), 392–404.

Hendrycks, D., & Dietterich, T. (2019). *Benchmarking Neural Network Robustness to Common Corruptions and Perturbations*.

Hosseini, E. S., & Moattar, M. H. (2019). Evolutionary feature subsets selection based on interaction information for high dimensional imbalanced data classification. *Applied Soft Computing*, *82*, 105581. https://doi.org/10.1016/j.asoc.2019.105581

Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L.-C., Tan, M., Chu, G., Vasudevan, V., Zhu, Y., Pang, R., Adam, H., & Le, Q. (2019). Searching for MobileNetV3. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 1314–1324. https://doi.org/10.1109/ICCV.2019.00140

Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2017). *Squeeze-and-Excitation Networks*.

Iman Soleimani. (n.d.). *Predictive Models for Railway Track Geometry Degradation*. Luleå University of Technology.

International Heavy Haul Association. (2001). Guidelines to Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues. *Library of Congress Control No.: 2001-131901*.

Iwnicki, S. (2006). *Handbook of railway vehicle dynamics*. CRC press.

J. Redmon, A. F. (2018). *YOLOv3: An Incremental Improvement - https://arxiv.org/abs/1804.02767v1*. ArXiv.Org.

Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, *20*(7), 1483–1510.

Jiang, T., Frøseth, G. T., Rønnquist, A., & Fagerholt, E. (2020a). A robust line-tracking photogrammetry method for uplift measurements of railway catenary systems in noisy backgrounds. *Mechanical Systems and Signal Processing*, *144*, 106888. https://doi.org/10.1016/j.ymssp.2020.106888

Jiang, T., Frøseth, G. T., Rønnquist, A., & Fagerholt, E. (2020b). A robust line-tracking photogrammetry method for uplift measurements of railway catenary systems in noisy backgrounds. *Mechanical Systems and Signal Processing*, *144*, 106888. https://doi.org/10.1016/j.ymssp.2020.106888

Jovanovic, S. (2004). Railway track quality assessment and related decision making. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, *6*, 5038–5043. https://doi.org/10.1109/ICSMC.2004.1400992

K. Karttunen, E. K. A. E. (2012). A numerical study of the influence of lateral geometry irregularities on mechanical deterioration of freight tracks. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *226*(6), 575–586.

K. Sun, B. Xiao, D. Liu, & J. Wang. (2019). Deep High-Resolution Representation Learning for Human Pose Estimation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5686–5696.

'Kaviani, N. (2020). *Identification of the Most Suitable Sensor Technology to Measure the Transversal Position of the Wheel in Relation to Rail*. Sapienza University of Rome.

Kaviani, N., Antognoli, M., Bureika, G., Ricci, S., Rizzetto, L., & Skrickij, V. (n.d.). *Development of a contactless sensor system to support rail track geometry on-board monitoring*.

Kaviani, N., Licciardello, R., & Shahidzadeh Sina. (2023). Effects of rail vehicle dynamics modelling choices on machine learning analysis. *28th IAVSD Symposium on Dynamics of Vehicles on Roads and Tracks (IAVSD 2023)*.

Khouzani, Golroo, A.H.E, & A., B. M. (2016). Railway maintenance management using a stochastic geometrical degradation model. *Journal of Transportation Engineering, Part A: Systems*, *143*(1).

Kotsiantis, S. B. (2007). Supervised machine learning: a review of classification techniques. *Informatica*, *31*, 3–24.

Lee, J. S., Hwang, S. H., Choi, I. Y., & Kim, I. K. (2018). Prediction of track deterioration using maintenance data and machine learning schemes. *Journal of Transportation Engineering, Part A: Systems*, *144*(9).

Lewis, R. B., & Richards, A. N. (1988). A compensated accelerometer for the measurement of railway track crosslevel. *IEEE Transactions on Vehicular Technology*, *37*(3), 174–178. https://doi.org/10.1109/25.16544

Li, Z.-W., He, Y.-L., Liu, X.-Z., & Zhou, Y.-L. (2020). Long-Term Monitoring for Track Slab in High-Speed Rail via Vision Sensing. *IEEE Access*, *8*, 156043–156052. https://doi.org/10.1109/ACCESS.2020.3017125

Lu, S. , T. Y. , L. H. (2007). Predictive condition-based maintenance for continuously deteriorating systems. *Quality and Reliability Engineering International*, *23*(1), 71–81.

M. Kim. (2012). Measurement of the Wheel-rail Relative Displacement using the Image Processing Algorithm for the Active Steering Wheelsets. *International Journal of Systems Applications, Engineering & Development*, *6*.

Martey, E. N., & Attoh-Okine, N. (2018). Modeling tamping recovery of track geometry using the copula-based approach. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *232*(8), 2079–2096.

MoniRail. (2019). *commercialising an in-service train data analytics platform coupled with an inertial measurement unit (IMU)*. https://monirail.co.uk/

NVIDIA. (n.d.). *NVIDIA Jetson TX2: High Performance AI at the Edge*. Https://Www.Nvidia.Com/En-Us/Autonomous-Machines/Embedded-Systems/Jetson-Tx2/.

Pedregosa et al. (2011). *Permutation feature importance*. https://scikit-learn.org/stable/modules/permutation_importance.html

Quiroga, L., & Schnieder, E. (2012). *Heuristic Forecasting of Geometry Deterioration of High Speed Railway Tracks* (pp. 609–616). https://doi.org/10.1007/978-3-642-27579-1_78

Ranjha, S. A. (2013). *Effect of head wear on rail underhead radius stresses and fracture under high axle load conditions*. https://doi.org/10.13140/2.1.2380.0644

Reeves, J., Remenyte-Prescott, R., & Andrews, J. (2019). Sensor selection for fault diagnostics using performance metric. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, *233*(4), 537–552. https://doi.org/10.1177/1748006X18804690

RIVAS. (2013). *Overview of Methods for Measurement of Track Irregularities Important for Ground-Borne Vibration - Deliverable D2.5*.

S. Ioffe, & C. Szegedy. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning*, *37*, 448–456.

Sharma, S., Cui, Y., He, Q., Mohammadi, R., & Li, Z. (2018). Data-driven optimization of Railway maintenance for track geometry. *Transportation Research Part C: Emerging Technologies*, *90*, 34–58.

Shi, D., Kaviani, N., Rizzetto, L., Skrickij, V., Oliverio, R., Šabanovič, E., Ye, Y., Bureika, G., Ricci, S., & Hecht, M. (2022). Deep learning based virtual point tracking for real-time target-less dynamic displacement measurement in railway applications. *Mechanical Systems and Signal Processing*, *166*, 108482. https://doi.org/10.1016/j.ymssp.2021.108482

Shi, D., Ye, Y., Gillwald, M., & Hecht, M. (2021). Designing a lightweight 1D convolutional neural network with Bayesian optimization for wheel flat detection using carbody accelerations. *International Journal of Rail Transportation*, *9*(4), 311–341. https://doi.org/10.1080/23248378.2020.1795942

SIST EN 13848-5:2008. (2008). *Railway applications - Track - Track geometry quality - Part 5: Geometric quality levels*.

Soleimanmeigouni, I., Ahmadi, A., Khajehei, H., & Nissen, A. (2020). Investigation of the effect of the inspection intervals on the track geometry condition. *Structure and Infrastructure Engineering*, *16*(8), 1138–1146. https://doi.org/10.1080/15732479.2019.1687528

SS-EN 13848-1: 2004+A1, 2008. (2004). *Railway applications - Track - Track geometry quality - Part 1: Characterisation of track geometry*.

Stereolabs. (2019). *Datasheet ZED2 Nov 2019 rev6*. Https://Www.Stereolabs.Com/Assets/Datasheets/Zed2-Cameradatasheet.Pdf.

Tao, Z., Huiling, L., Wenwen, W., & Xia, Y. (2019). GA-SVM based feature selection and parameter optimization in hospitalization expense modeling. *Applied Soft Computing*, *75*, 323–332. https://doi.org/10.1016/j.asoc.2018.11.001

TRAFIKVERKET. (2017). *The Swedish Transport Administration Annual Report*.

TSUNASHIMA, H., MORI, H., KOJIMA, T., MATSUMOTO, A., & MIZUMA, T. (2012). Condition Monitoring of Railway Track Using In-service Vehicle. *Journal of Mechanical Systems for Transportation and Logistics*, *3*(1), 154–165. https://doi.org/10.1299/jmtl.3.154

T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, & D. Ramanan, et al. (2018). *Microsoft COCO: Common Objects in Context*. European Conference on Computer Vision. https://www.microsoft.com/en-us/research/publication/microsoft-coco-common-objects-in-context/

UIC. (2008). *Best practice guide for optimum track geometry durability* (UIC (International Union of Railways), Ed.). ETF - Railway Technical Publications.

UIC 518: 4ED. (2009). *Testing and approval of railway vehicles from the point of view of their dynamic behaviour - safety - track fatigue - ride quality*.

Vladimir Lyashenko. (n.d.). *How to use random forest for regression*. Retrieved June 20, 2022, from https://cnvrg.io/random-forest-regression/

VUZ. (n.d.). *VUZ Test Centre Velim*. Výzkumný Ústav Železniční. Retrieved April 10, 2022, from https://www.cdvuz.cz/en/page/zkusebni-centrum-vuz-velim

Weston PF, L. C. G. C. et al. (2007). Monitoring lateral track irregularity from in-service railway vehicles. *Proc IMechE, Part F: J Rail Rapid Transit*, *221*, 89–100.

Weston PF, R. C. Y. G. et al. (2015). Perspectives on railway track geometry condition monitoring from in-service railway vehicles. *Veh Syst Dyn*, *53*, 1063–1091.

Wheel Conicity. (n.d.).

X. Zhang, K. He, S. Ren, & J. Sun. (2016). Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Xin, T., Famurewa, S. M., Gao, L., Kumar, U., & Zhang, Q. (2016). Greysystem-theory-based model for the prediction of track geometry quality. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *230*(7), 1735–1744.

Xu, Y., & Brownjohn, J. M. W. (2018). Review of machine-vision based methodologies for displacement measurement in civil structures. *Journal of Civil Structural Health Monitoring*, *8*(1), 91–110. https://doi.org/10.1007/s13349-017-0261-4

Zach. (2021, June). *RMSE vs. R-Squared: Which Metric Should You Use?* https://www.statology.org/rmse-vs-r-squared

Zhan, D., Jing, D., Wu, M., Zhang, D., Yu, L., & Chen, T. (2018). An Accurate and Efficient Vision Measurement Approach for Railway Catenary Geometry Parameters. *IEEE Transactions on Instrumentation and Measurement*, *67*(12), 2841–2853. https://doi.org/10.1109/TIM.2018.2830862

Zhang, L., Wu, X., Ji, W., & AbouRizk, S. M. (2017). Intelligent Approach to Estimation of Tunnel-Induced Ground Settlement Using Wavelet Packet and Support Vector Machines. *Journal of Computing in Civil Engineering*, *31*(2). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000621

Zhang, P. (2019). A novel feature selection method based on global sensitivity analysis with application in machine learning-based prediction model. *Applied Soft Computing*, *85*, 105859. https://doi.org/10.1016/j.asoc.2019.105859

Zmigrod. (n.d.). *Zmigrod Track Test Centre*. Retrieved April 10, 2022, from http://www.ikolej.pl/en/units/test-track-centre-near-zmigrod

Appendix

Annex I

Figure 86 Examples of static images taken on different bogies

Figure 87 Examples of images from a YouTube video

Annex II

Table 17 LightPointNet_large architecture

Block	Type	NL	SE	Exp size	Filters	Size/Stride	Output
0	Conv	HS	false	-	16	$3 \times 3/2$	$128 \times 128 \times 16$
1	Bneck	RE	false	16	16	$3 \times 3/2$	$64 \times 64 \times 16$
2	Bneck	RE	false	72	24	$3 \times 3/2$	$32 \times 32 \times 24$
3	Bneck	RE	false	72	24	$3 \times 3/1$	$32 \times 32 \times 24$
4	Bneck	RE	true	72	40	$5 \times 5/2$	$16 \times 16 \times 40$
5	Bneck	RE	true	120	40	$5 \times 5/1$	$16 \times 16 \times 40$
6	Bneck	RE	true	120	40	$5 \times 5/1$	$16 \times 16 \times 40$
7	Bneck	HS	false	120	48	$5 \times 5/1$	$16 \times 16 \times 48$
8	Bneck	HS	false	144	48	$5 \times 5/1$	$16 \times 16 \times 48$
9	Bneck	HS	false	144	48	$5 \times 5/1$	$16 \times 16 \times 48$
10	Bneck	HS	false	144	48	$5 \times 5/1$	$16 \times 16 \times 48$
11	Bneck	HS	true	192	64	$5 \times 5/2$	$8 \times 8 \times 64$
12	Bneck	HS	true	192	64	$5 \times 5/1$	$8 \times 8 \times 64$
13	Bneck	HS	true	384	128	$5 \times 5/1$	$8 \times 8 \times 128$
14	Bneck	HS	true	384	128	$5 \times 5/1$	$8 \times 8 \times 128$
15	Bneck	HS	true	384	128	$5 \times 5/1$	$8 \times 8 \times 128$
16	ConvTranspose	RE	false	-	256	$4 \times 4/2$	$16 \times 16 \times 256$
17	ConvTranspose	RE	false	-	256	$4 \times 4/2$	$32 \times 32 \times 256$
18	ConvTranspose	RE	false	-	256	$4 \times 4/2$	$64 \times 64 \times 256$
19	Conv	-	false	-	3	$1 \times 1/1$	$64 \times 64 \times 3$

Table 18 LightPointNet_small architecture

Block	Type	NL	SE	Exp size	Filters	Size/Stride	Output
0	Conv	HS	false	-	16	$3 \times 3/2$	$128 \times 128 \times 16$
1	Bneck	RE	false	16	16	$3 \times 3/2$	$64 \times 64 \times 16$
2	Bneck	RE	true	72	24	$3 \times 3/2$	$32 \times 32 \times 24$
3	Bneck	RE	false	88	24	$3 \times 3/1$	$32 \times 32 \times 24$
4	Bneck	HS	false	96	40	$5 \times 5/2$	$16 \times 16 \times 40$
5	Bneck	HS	true	240	40	$5 \times 5/1$	$16 \times 16 \times 40$
6	Bneck	HS	true	240	40	$5 \times 5/1$	$16 \times 16 \times 40$
7	Bneck	HS	true	120	48	$5 \times 5/1$	$16 \times 16 \times 48$
8	Bneck	HS	true	144	48	$5 \times 5/1$	$16 \times 16 \times 48$
9	Bneck	HS	true	192	64	$5 \times 5/2$	$8 \times 8 \times 64$
10	Bneck	HS	true	384	64	$5 \times 5/1$	$8 \times 8 \times 64$
11	Bneck	HS	true	384	64	$5 \times 5/1$	$8 \times 8 \times 64$
12	ConvTranspose	RE	false	-	64	$4 \times 4/2$	$16 \times 16 \times 64$
13	ConvTranspose	RE	false	-	128	$4 \times 4/2$	$32 \times 32 \times 128$
14	ConvTranspose	RE	false	-	256	$4 \times 4/2$	$64 \times 64 \times 256$
15	Conv	-	false	-	3	$1 \times 1/1$	$64 \times 64 \times 3$