Artificial Intelligence Technology Using Smart Phones For Train Passenger Ride Comfort

By

Junhui Huang

A thesis submitted to the University of Birmingham for the degree of Doctoral of Philosophy

School of Engineering

Department of Civil Engineering

University of Birmingham

September 2024

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

All rights reserved. This copyright is held by the author. No part of this thesis may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the author.

Abstract

This thesis presents a novel approach to evaluating Ride Comfort (RC) in railway systems by integrating Machine Learning (ML) algorithms with data collected from smartphone sensors. Traditional methods for assessing RC, such as questionnaires and vibration-based measurements, have significant limitations, including reliance on subjective feedback, and complex logistics. These challenges are addressed by developing a scalable, cost-effective, and accurate framework that leverages modern smartphones' widespread availability and advanced capabilities.

The research begins by exploring the potential of ML algorithms in processing large datasets generated from both synthesised data using D-track simulations and real-world data collected via smartphones. The study adheres to the International Organisation for Standardisation (ISO) 2631 standard and the International Union of Railways (UIC) 513 for vibration measurement, ensuring the reliability and accuracy of the data. The ML models developed, including Convolutional Neural Network (CNN), K-means clustering, and ensemble techniques, are employed to quantify dynamic track stiffness, classify various train motions (such as roll, yaw, pitch, and bounce), and evaluate RC at individual points within the train.

A key innovation of this work is using crowd-sourced data from multiple smartphones with Graph Attention Network (GAT) to perform a comprehensive assessment of RC at the train level rather than merely at the individual passenger level. This approach facilitates the transmission of real-time data to a processing centre for subsequent analysis, generating valuable insights to improve passenger experience and operational efficiency, all while ensuring driver focus and maintaining safety standards. The study also

demonstrates that this methodology significantly reduces the financial and logistical burdens typically associated with traditional methods while offering a more holistic and nuanced RC evaluation.

The findings of this research have broad implications for the railway industry, particularly in improving infrastructure maintenance, optimising train operations, and enhancing overall RC. By providing a robust framework for RC assessment, this thesis contributes to advancing intelligent transportation systems and supports the future development of more comfortable and efficient railway services. The use of advanced ML models, such as CNN and GAT, in conjunction with crowd-sourced data from multiple smartphones, enables a comprehensive and dynamic assessment of RC at both individual and train-wide levels. This approach allows for more accurate quantification of crucial factors such as track stiffness and train motion, which are critical to understanding and enhancing RC. The cost-effectiveness of the methodology, driven by the use of widely available smartphones, significantly reduces the financial and logistical burdens associated with traditional methods, making it accessible to a wide range of railway services. The research also aligns with international standards, ensuring its applicability across diverse operational contexts. Beyond the railway industry, the findings have broader implications for other modes of transport, such as buses and trams, where similar methodologies could be applied to enhance comfort and operational performance. The adoption of this ML-driven framework not only promises to improve passenger satisfaction and safety but also supports the development of sustainable, passenger-centred transportation systems.

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Dr Sakdirat Kaewunruen, for his unwavering support and insightful guidance throughout my PhD journey. His expertise and wisdom have been invaluable to me.

My heartfelt thanks go to my friends in the TOFU research group for their friendship and the unforgettable moments we shared during my time in Birmingham. I am especially grateful to Jesseda Sresakoolchai, Chayut Ngamkhanong, Panrawee Rungskunroch, Dan Li, Xu Huang, Hao Fu, Xia Qin, Rucheng Liu, and Pasakorn Sengsri.

Lastly, I would like to thank my family for their constant encouragement and support throughout this challenging journey. Their unwavering belief in me has been a continuous source of strength and inspiration.

Contents

List of Figures

List of Tables

Abbreviations

Term	Explanation / Meaning / Definition
AI	Artificial Intelligence
ANN	Artificial Neural Network
ADHD	Attention Deficit Hyperactivity Disorder
CNN	Convolutional Neural Network
CF	Crest Factor
DT	Decision Tree
DFT	Discrete Fourier Transform
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyography
EN	European Standard
FFT	Fast Fourier Transform
FE	Finite Element
GRU	Gated Recurrent Unit
GAT	Graph Attention Network
HRV	Heart Rate Variability
HST	High-Speed Train
IMU	Inertial Measurement Unit
ISO	International Organisation for Standardisation
UIC	International Union of Railways
KCV	K-fold Cross Validation
KNN	K-Nearest Neighbours
LSTM	Long Short-Term Memory

LP	Lumped Parameter
ML	Machine Learning
MTVV	Maximum Transient Vibration Value
MB	Multi-Body
MLP	Multilayer Perception
PCA	Principal Component Analysis
RF	Random Forest
RNN	Recurrent Neural Network
RC	Ride Comfort
R.M.S	Root Mean Square
RSS	Root Sum of Square
SVM	Support Vector Machine
SVR	Support Vector Regressor
R^2	The coefficient of determination
V	Vector Sum
VDV	Vibration Dose Value
WBV	Whole-Body Vibration

Chapter 1

Introduction

1.1 Background and Research Challenges

Train RC is vital to determining public transportation users' overall satisfaction and experience, making it a key factor in attracting passengers to rail travel, particularly for short- and medium-distance journeys [1, 2]. Enhanced comfort can position rail as a more appealing alternative to air travel and private vehicles, which is increasingly relevant given the growing demand for sustainable transportation. Studies indicated that a 40% rise in High-Speed Train (HST) users over the past decade is primarily attributed to improvements in RC [1], leading to a shift away from cars and flights, thus contributing to a reduction in carbon emissions, as trains are more environmentally sustainable than other modes of transport [3]. Mouwen [4] suggested that emissions from urban passenger transport could decrease by as much as 40% by 2050 if commuters adopted alternative modes of transport, such as cycling, walking, or using rail services in urban areas.

At the same time, the growing number of railway users attracted by enhanced train RC brings additional benefits across various aspects. Soehodho [5] confirmed that more public transport travellers would reduce the number of traffic accidents. Many countries have prioritised the development of efficient public rail services to reduce reliance on private cars. It is now widely recognised that car dependency has resulted in numerous negative consequences, including increased traffic congestion, environmental degradation, noise pollution, accidents, and air contamination in cities worldwide [6].

Good RC is achieved through well-maintained infrastructure and high-quality rolling stock [7], where poor comfort can often signal suboptimal track conditions. Therefore, estimating RC is critical for achieving operational efficiency, environmental sustainability, and meeting contemporary transportation demands.

Estimating RC is not novel and has been studied for decades. For example, various methods include questionnaires, calculations based on vibrations, or estimation using biological features. Traditional methods pose several challenges that limit their effectiveness, scalability and applicability in modern railway systems. The questionnaire often depends on passengers' subjective feelings, which leads to inconsistent and unreliable data. At the same time, potential pitfalls may arise in the interpretation of the data [8]. Vibration-based methods [9], which rely on accelerometers, involve complex logistics, demand expert knowledge, and are often time-consuming. Additionally, they are susceptible to human error, further complicating the process.

The logistics issue is predominant if a full-scale assessment adhering to ISO 2631 leads to a growing number of sensors needed to cover different passenger postures and measuring locations. Traditional methods typically use post-journey data analysis, which

fails to exhibit live assessment and make immediate adjustments. Only one attribute – vibration is sometimes considered not to fully capture passengers' comprehensive comfort experience. Some studies have extended the scope by introducing additional features such as lighting and ventilation in [9], heat rate variability in [10] and physiological factors (body pressure distribution, electroencephalogram (EEG), and electromyography (EMG)) in [11]. However, individual differences, intrusiveness, cost, and lack of a universally applicable evaluation standard hinder the wide and day-to-day application of the additional features apart from the vibration feature.

1.2 The Role of Machine Learning and Smartphone in Railway Systems

Integrating ML and smartphones presents many benefits to overcoming traditional methods' limitations. The advancement of smartphone sensors underscores the integrity of data collection, and the ubiquitous smartphone can expand the area to be assessed, providing a more intact picture of the whole train RC analysis. This also eases the burden of investment and offers opportunities to monitor every train’s RC, which was impossible before. With the advent of ML, which is cut out for large datasets, the real-time estimation of RC can be delivered. ML algorithms can automate the processing and analysis of large volumes of data so that the manual effort and time to process and analyse the data is reduced. The prevalence of smartphones and ML’s data processing efficacy also allows for frequent and comprehensive assessments.

The synergy of ML and smartphones provides a robust alternative to the limitations of traditional methods. With the comprehensive, real-time, and scalable data collection and analysis the proposed method can offer, it can enhance the accuracy, efficiency, and practicality of large-scale railway RC assessments.

1.3 Aims and Objectives

To resolve the limitations of the traditional methods, this study aims to integrate ML and smartphones to estimate RC to ensure an accurate, cost-effective and scalable alternative.

To deliver the aim, the objectives are defined as follows:

1. To explore the feasibility of vibration features and ML algorithms in railway:
    - To design and implement ML that can quantify track stiffness using synthesised data from D-track.
    - To evaluate the performance of the ML models that can quantify track stiffness.
2. Implement comprehensive data collection using smartphones:
    - To follow the ISO 2631 standard and UIC 513 to deploy smartphones onboard a train to acquire accelerations.
    - To ensure data collection's reliability and integrity through calibration and validation of smartphone sensors.
3. Understand different types of train motion using ML:
    - To extract and label different motions from the data obtained from the smartphones.
    - To develop unsupervised ML models that can classify different train motions (roll, yaw, pitch, bounce, and hunting) based on smartphone data.
4. To determine RC at a singular location for an individual passenger:
    - To develop ML models to estimate the RC collected based on each smartphone.

5. Leverage crowd-sourced data for overall comfort assessment:
    - To explore the feasibility of using data from multiple smartphones
    - To test the performance using RC and the topology of every four phones to predict the RC of the rest of the phones.

1.4 Scope of the Research

The scope of work falls into understanding the train dynamic and, finally predicting the RC. This thesis is limited to one feature – vibration influencing the RC- the most widely used and comprehensively studied. Other attributes, such as temperature and lighting, are neglected. The smartphone is selected for the data collection tool, and the RC calculation and data collection deployment are guided thoroughly by ISO 2631 and UIC 513. ML models include two main categories – unsupervised learning and supervised learning.

1.5 Innovations and Contributions

The list below summarises the novelties and contributions of this thesis:

Novelties:

1. This thesis innovatively combines ML and smartphone sensors for railway RC assessment, replacing traditional accelerometers with widely available smartphones. By using smartphones, the method offers a cost-effective, scalable solution.
2. This thesis introduces using multiple smartphones to assess comfort at the train level, providing a scalable and comprehensive evaluation method.

Contributions:

1. The method reduces costs and logistics using smartphones, offering a scalable solution.
2. It provides a more thorough and detailed RC evaluation through comprehensive data collection.
3. Real-world data demonstrates the method's practical effectiveness.
4. Crowd-sourced smartphone data enables real-time, ISO 2631-compliant measurements.
5. The system estimates train operations by transmitting real-time acceleration data to the operations centre, allowing for post-analysis without distracting the driver.
6. RC monitoring offers insights into track and train conditions.

1.6 Thesis Structure

This thesis is organised into eight chapters, each designed to explore the various aspects of the research problem, methodologies, and findings related to the use of ML and smartphone sensors in railway RC analysis. The structure of the thesis is as follows:

- **Introduction**

  The first chapter introduces the research problem, highlighting the importance of the RC in railway systems and the limitations of traditional methods. It outlines the existing challenges in assessing the RC using traditional methods, such as questionnaires and vibration-based assessments, and introduces the potential of integrating ML with smartphone technology as a novel approach to overcome

these limitations. The chapter clearly states the research aim: to develop a scalable, cost-effective, and accurate method for real-time RC evaluation. It also defines the specific objectives, including the design of ML models, data collection via smartphones, and the overall goal of enhancing the reliability of RC assessments. The chapter concludes by highlighting the novelties and contributions of this work, such as pioneering the use of crowd-sourced smartphone data in RC evaluation and the potential implications for railway operations and passenger experience.

- **Literature Review**

  This chapter provides an in-depth review of the existing knowledge related to railway RC. It begins by discussing the fundamental concepts of vibration and its effects on human comfort, followed by a detailed analysis of the standards used in evaluating the RC, such as ISO 2631 and UIC 513. The review then shifts focus to traditional methods for assessing the RC, including the limitations of questionnaires, physiology-based methods, and vibration-based assessments. It also explores recent advancements in the application of ML and smartphones in the context of railway systems, highlighting their potential to revolutionise RC evaluation. The chapter identifies research gaps, particularly in integrating ML with real-time data collection via smartphones, setting the stage for the methodologies proposed in this thesis.

- **Methodologies**

  The methodologies chapter is central to the thesis, detailing the processes and techniques used to achieve the research objectives. It begins by describing the

data collection methods, including using D-track simulations to generate synthesised vibration data and deploying smartphones in real-world settings to collect field data. The chapter then covers data processing techniques, such as outlier removal, normalisation, and feature extraction, which are crucial for preparing the data for ML model training. It provides an overview of the ML models developed for this research, including supervised and unsupervised learning approaches, and explains how they are used to quantify track stiffness, classify train motions, and assess RC at individual and collective levels. The chapter concludes with a discussion on the validation and evaluation of the models, ensuring their accuracy and reliability in real-world applications.

- **Quantification of dynamic track stiffness using machine learning**

  This chapter delves into the specific application of ML to quantify dynamic track stiffness, a critical factor in assessing and improving railway infrastructure. It begins with introducing the problem and explaining the importance of track stiffness in maintaining RC and operational safety. The methodologies section outlines the use of D-track for data simulation and the development of ML models, such as Convolutional Neural Network (CNN), to analyse the correlation between track characteristics and stiffness. The chapter presents the model's results, discussing their accuracy and potential implications for railway maintenance and optimisation. The findings are compared with existing methods, highlighting the advantages of using ML for more precise and scalable assessments.

- **Train motion classification**

  This chapter uses ML models to classify train motions (e.g., roll, yaw, pitch, bounce, and hunting). The chapter begins with an introduction to the significance of understanding train motions concerning RC, as different motions can have varying impacts on passenger experience. The methodologies section details the data acquisition process, including the deployment of smartphones to capture motion data and the feature extraction techniques used to identify relevant motion patterns. The development and training of ML models are explained in detail. The chapter then presents the results, demonstrating the models' ability to accurately classify train motions and their potential to provide real-time insights for improving RC. The discussion explores how these classifications can inform train design and operational adjustments.

- **Point-specific train ride comfort quantification**

  This chapter addresses the challenge of assessing RC at specific points within the train. The chapter commences with an introduction to the problem, emphasising the need for localised comfort assessments to provide a more nuanced understanding of RC. The methodologies section describes the data collection process using smartphones placed at various points within the train, followed by the development of ensemble ML models to provide point-specific comfort evaluations. The chapter presents the results, showcasing the model's ability to accurately predict RC at individual points. This chapter provides a firm foundation for the later chapter, which uses the RC from different locations on a train to predict the overall RC for the train.

- **Overall train ride comfort using crowd-sensing data**

  This chapter presents an innovative approach to assessing overall RC by integrating data from multiple smartphones using crowd-sensing techniques. The chapter introduces the concept of crowd-sensing and its potential to provide comprehensive, real-time comfort assessments at the train level. The methodologies section details the data collection strategy, where multiple passengers' smartphones are used to gather vibration data simultaneously, and the subsequent data processing techniques to ensure consistency and accuracy across devices. The development of ML models that utilise this crowd-sourced data to predict RC across the entire train is explained. The chapter presents the results, demonstrating the robustness of the ML models in aggregating and analysing data from diverse sources. The discussion highlights the scalability of this approach, its potential to revolutionise RC assessment, and its implications for improving overall passenger experience and railway operations.

- **Conclusions and recommendations**

  The final chapter summarises the key findings of the research, reflecting on how the integration of ML and smartphones can enhance the assessment and improvement of railway RC. The chapter discusses the broader implications of the findings for the railway industry. The chapter concludes with suggestions for future research, such as exploring additional comfort factors (e.g., noise, temperature) and expanding the use of crowd sensing in other areas of transportation and urban planning.

Chapter 2

Literature Review

2.1 Introduction and Aims of the Literature Review

RC is a paramount determinant of overall satisfaction for users of railway systems. Growing RC is helpful for the attractiveness of rail transit as a preferred mode of transportation. Over the past decade, many improvements have been seen in RC, which leads to increasing train users and reducing carbon emissions by shifting reliance away from less sustainable transport options [12]. The traditional methods for assessing RC, such as questionnaires and vibration-based measurements, have notable limitations that affect their widespread application. Questionnaires depend on individual perceptions, leading to inconsistent data, while vibration-based methods requiring complex logistics are time-consuming and often lack real-time data processing capabilities.

Recent developments in ML and smartphone technologies offer a new platform for these challenges. Smartphones with high-performance sensors can collect extensive real-time

data, while ML can efficiently process and analyse large datasets to make real-time prediction possible.

This literature review aims to explore and summarise the existing research on railway RC, commencing with fundamental knowledge to support this thesis on how vibration can affect RC (Section 2.2) and the technical standards that define RC from vibrations (Section 2.3). Subsequently, it focuses on the evolution and limitations of assessment methods (Section 2.4), the potential of ML in RC (Section 2.5), and the potential of smartphones in RC (Section 2.6). By synthesising the findings from various studies, this review lays the foundation for understanding the field's current state and recognises the gaps (Section 2.7) that the proposed method can address.

2.2 Understanding of Vibration

2.2.1 Definition of Vibration in Humans

RC is a complicated and multidimensional subject that results from a wide range of entities. The main feature widely used to measure RC relies on vibrations transferred to the human body. Figure 2-1 shows examples of various vibrational waveforms – sinusoidal, multi-sinusoidal, transient, shock, stationary random, and non-stationary random.

A single-frequency waveform characterises the sinusoidal wave as a most straightforward form of vibration. The multi-sinusoidal waveform synthesises different sinusoidal waves with different frequencies and amplitudes. The transient wave starts with a high amplitude followed by a gradual decay. It usually can be found on a system responding to an abrupt, non-repetitive disturbance and returning to equilibrium. It is

observable that the following waveform is a shock that is characterised by a sudden spike lasting a short duration. This can be seen in a system impacted by a sudden force such as a hammer strike or collision. The main difference between the stationary random and non-stationary random waveforms is the mean and variance of the waveform. It is noted that the stationary waveform presents a constant mean and variance, while the non-stationary random wave's mean and variance vary over time.

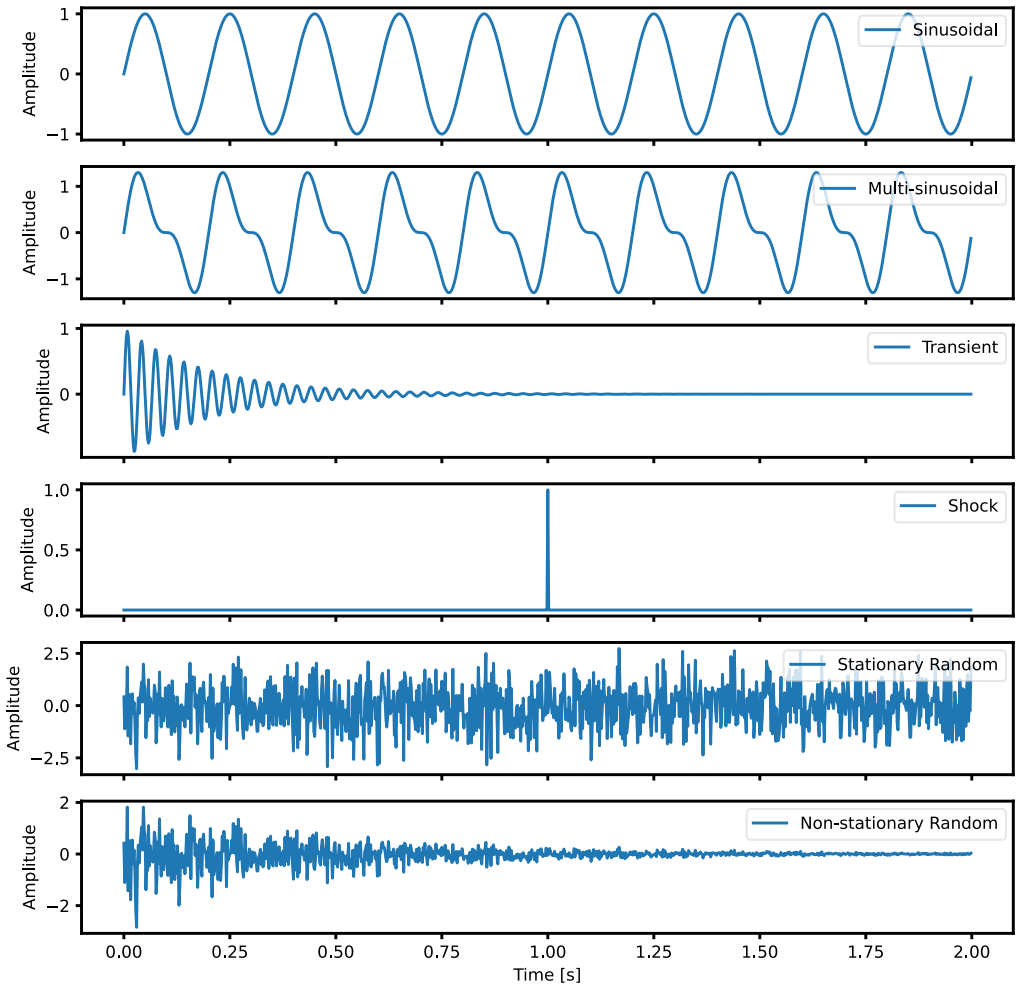


Figure 2-1 Different types of vibration

2.2.2 Impact of Vibration on the Human Body

Understanding the coordinate systems is a prerequisite to estimating the effect of vibration as a human response to vibration is direction-dependent, which means up-and-down vibrations result in different impacts from side-to-side vibrations on the human body. Basicentric coordinate systems defined in ISO2631 are employed to study human Whole-Body Vibration (WBV). Figure 2-2 gives an example of the basicentric systems, showing that the systems originate from the points where vibrations transfer to the human body.

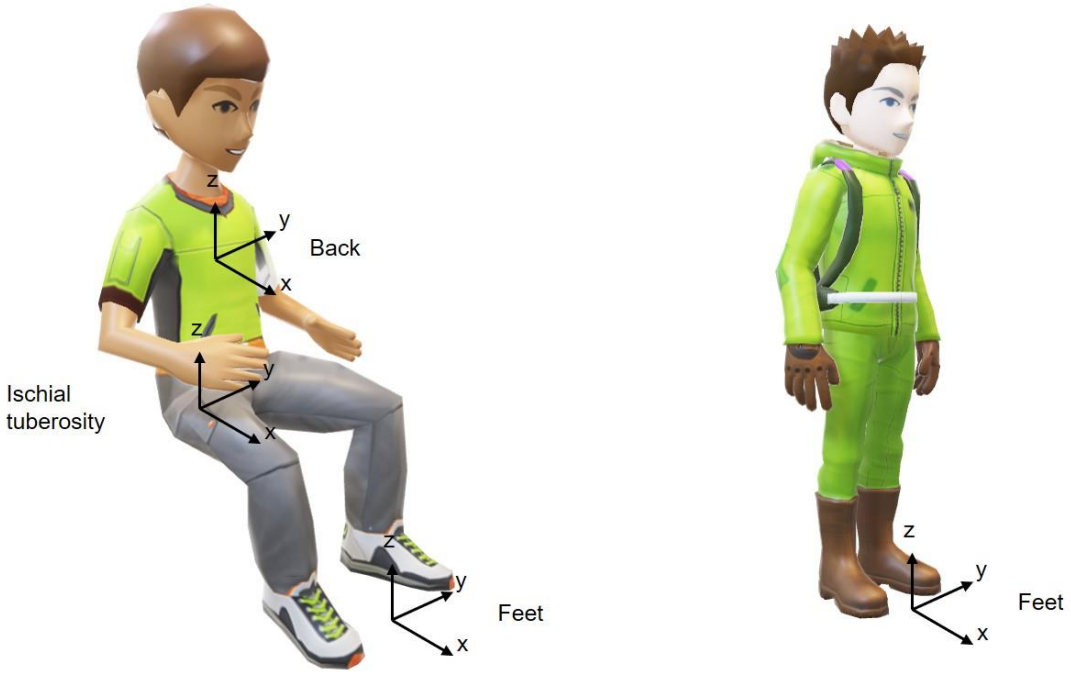


Figure 2-2 Basicentric coordinate systems

It is noted that people in different postures lead to different origins – a seated person sees three origins while a standing person has one origin since there are three and one contact points for the seated and standing persons, respectively. The nature of the basicentric coordinate systems, depending on the contact point, makes this system more accessible for coping with the assessment of environmental vibration.

The effect of vibration on different contact points is not unified as the different parts of the human body are sensitive to vibrations at different frequencies, with [13] underscoring the vibration at resonance frequency, which can amplify the motion, leading to discomfort. As illustrated in Figure 2-3, most parts of the human body are subject to vibrations at frequencies ranging from 2 – 50 Hz, except the hand grip, which is affected by a frequency span from 50 – 200 Hz.

The effect of vibration varies with vibration magnitude, frequency and direction. In general, the effect of magnitude is proportional to the magnitude of vibration [13], while the sensation of vibration magnitude can be described by Stevens' power law defined in Equation 1 [14].

$$\psi = k\,\varphi^{n} \qquad \text{Equation 1}$$

Where ψ is the psychophysical magnitude, φ is the physical magnitude, *k* is the constant dependent on the units adopted, and *n* is the exponent to determine the growth in sensation. The exponent *n* in Equation 1 for whole-body vertical vibration is considered unity one across frequencies 1 to 80 Hz, implying that the vibration magnitude doubles and the perceived discomfort also doubles [15]. The effect of vibration on humans cannot be decided only based on the magnitude of vibration, as a constant magnitude results in different discomfort at varied frequencies. At low frequencies (< 1 or 2 Hz), the forces applying on the body are roughly proportional to the input acceleration, leading to similar movement and discomfort throughout the whole body. With slightly increasing frequencies, body resonances amplify the motion, intensifying discomfort in specific body parts. As frequencies continuously increase, the body attenuates the effect of the vibration, reducing discomfort. Unlike the high-frequency vibration causing localised

discomfort to the body parts, the human body acts almost like a rigid system at low frequencies. More details how comfort contours specify discomfort responses to the magnitude, frequency and direction of vibration are in Section 2.3.

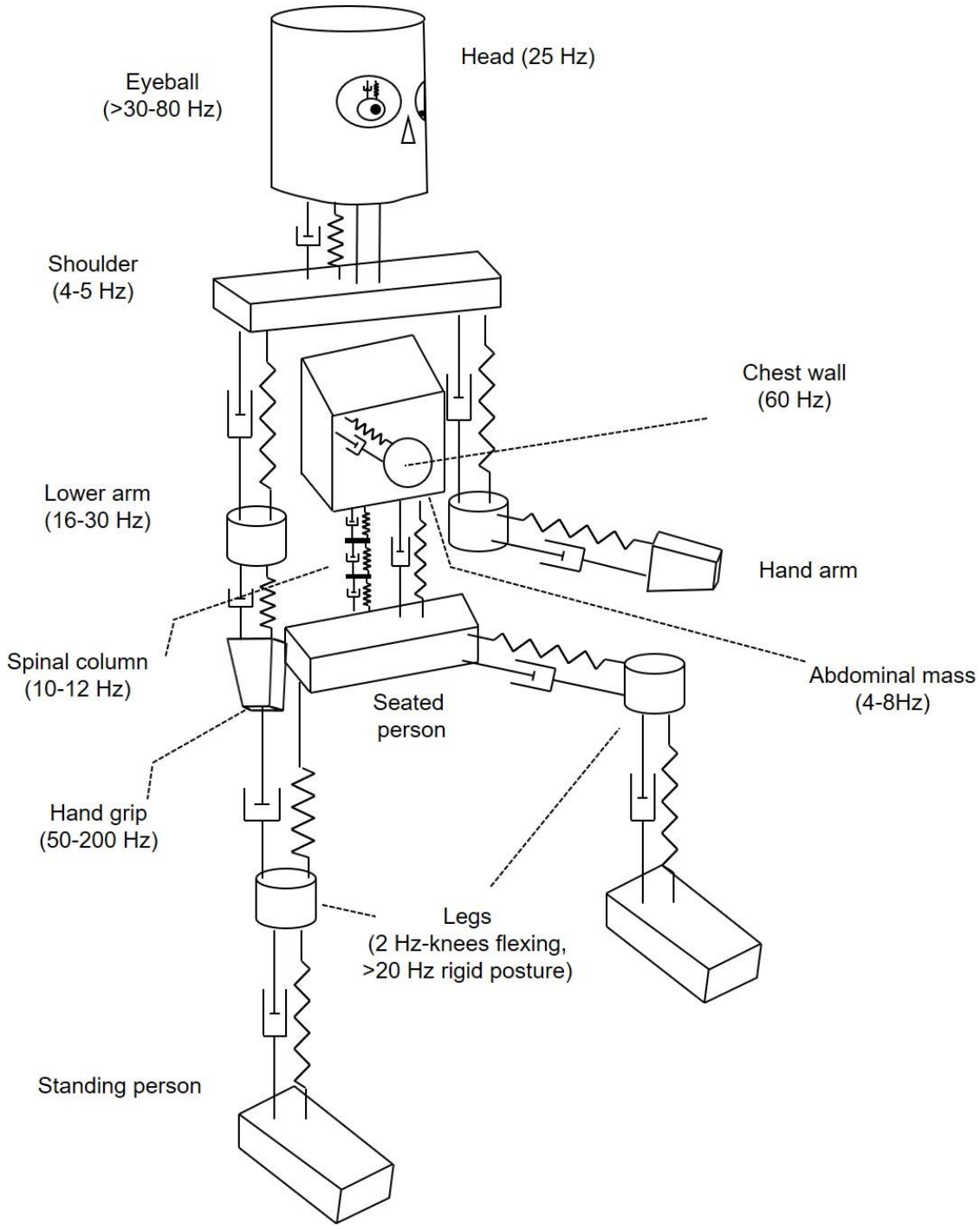


Figure 2-3 A mass-spring model of the human body [13]

2.2.3 Consequences of Vibration Exposure

The effect of vibration has been extensively studied. WBV can produce a wide range of effects from physiological (muscle, skeletal, nerve etc.), activity (vision, hearing, touch etc.), biodynamic (body impedance, head movements, body transmissibility etc) to subjective (absolute thresholds, subjective order, rating of stimuli etc.).

Table 1 Parameters that vibrations affect [13]

Subjective	**Activity**
Absolute thresholds Subjective equality Subjective order Equality of intervals Equality of ratios Rating of stimuli Cross modality judgements Differential thresholds	Vision Hearing Touch Proprioception Vestibular function Psychomotor performance Cognitive performance Vigilance
Physiological	**Biodynamic**
Skeletal Muscle Nerve Cardiovascular Respiratory Central nervous system Endocrine/metabolic	Body impedance Hand impedance Body transmissibility Head movements Hand movements Organ movements Energy absorbed

The adverse effects of vibration on the human body have been of primary interest to researchers. However, vibrations can sometimes be good. C Robin et al. stimulated the hindlimbs of adult owe using shallow magnitude and high-frequency vibrations five days

a week for one year, finding that the bone density of the sheep increased by 32% [16]. According to Fuemaier et al., WBV were applied to eighty-three healthy individuals and seventeen people diagnosed with Attention Deficit Hyperactivity Disorder (ADHD), concluding that WBV benefits the attention of both groups of people [17]. In therapeutic contexts, vibration cleared patients' lungs [14] and improved joint mobility for individuals with rheumatoid arthritis [18]. Vibration as a physiotherapy intervention has large physiological effects on human health, particularly for patients with respiratory disorders like cystic fibrosis. The manual application of fine oscillatory movements combined with chest wall compression during vibration increases peak expiratory flow rates (PEFR), assisting the clearance of secretions by moving mucus towards the oropharynx. The study found that vibration produced a higher mean PEFR than other physiotherapy techniques such as Flutter1, percussion, Acapella1, and positive expiratory pressure (PEP). Moreover, the frequency of airflow oscillation during vibration falls within the range (3-17 Hz) demonstrated to enhance mucus transport by decreasing its viscosity and stimulating ciliary movement. This makes vibration a potentially more effective method for secretion clearance in cystic fibrosis patients than other interventions, which either impede expiratory flow or lack the optimal oscillation frequency for mucociliary clearance. It is also noted that the induced oscillatory frequency of vibration aligns closely with the natural ciliary beat frequency, which could enhance mucus clearance further through resonance effects. Therefore, clinicians should consider the unique benefits of vibration when selecting physiotherapy treatments for patients with excessive respiratory secretions. Moreover, its application extends to post-amputation limb care [19], dialysis enhancement [20], and the rehabilitation of patients with paralysis [21], dramatically improving limb function and mobility. A study [19] detailed the development and

functionality of a vibrotactile device designed to enhance proprioception in persons with lower-limb amputations. Proprioception, the sense of self-movement and body position, plays an important role in balance and coordination. The loss of proprioceptive feedback in amputees can lead to instability and an increased risk of falls. The vibrotactile device developed in [19] aimed to mitigate these issues by providing tactile feedback through vibrating motors and solenoids, which simulate real-world perturbations such as dragging a foot or colliding with an object. This feedback was intended to help prosthetic users develop better control and awareness of their artificial limbs. The device's ability to produce two distinct types of feedback—vibrations and knocks—addressed different rehabilitative scenarios, potentially enhancing the motor learning and adaptation process for new prosthetic users. [20] investigated the impact of low-frequency vibrations on the efficacy of hemodialysis, specifically utilising a specially designed dialysis chair. The vibrations mimic the benefits of physical exercise, potentially improving the dialysis process by enhancing the filtration of fluids between compartments in the body. The study involved 21 patients who underwent hemodialysis while seated in a vibrating chair. Results demonstrated that vibrations significantly improved dialysis adequacy, indicated by the urea reduction ratio (URR). A notable increase in systolic blood pressure was also observed, although no adverse cardiovascular events were reported. This suggests that incorporating low-frequency vibrations during hemodialysis could be a safe and effective alternative to traditional intradialytic exercise, offering an innovative approach to improving patient outcomes. [21] demonstrated that high-frequency, low-magnitude WBV can counteract bone degradation induced by muscle paralysis via an experiment conducted on mice. Muscle-induced mechanical loading is important for maintaining bone integrity, and paralysis can lead to substantial bone loss. The study revealed that WBV

prevents trabecular bone mineral density reduction typically observed following paralysis induced by botulinum toxin A (BTX). Additionally, while insulin-like growth factor-I (IGF-I) is known to promote bone growth, its effectiveness was not observed in the absence of muscle activity. Notably, WBV combined with IGF-I did not enhance bone recovery beyond the effects of WBV alone, suggesting that WBV's mechanical stimuli are sufficient to mitigate paralysis-induced bone loss. This finding is pivotal as it underscores the potential of WBV as a non-pharmacological intervention to maintain bone health in conditions of impaired mobility.

However, vibrations can still present adverse influences on humans in various ways. WBV affects individuals exposed to occupational settings such as industry, mining, and transportation. According to [22], long-term exposure to WBV is associated with an elevated risk of developing disorders of the lumbar spine and connected nervous system, such as lower back pain, sciatic pain, and lumbar disc disorders. Additional potential effects include neck-shoulder pain, gastrointestinal issues, and disruptions in female reproductive health, although the epidemiological evidence for these is weak. Hand-transmitted vibration (HTV), typically encountered during the use of powered tools, leads to hand-arm vibration syndrome (HAVS). This syndrome encompasses vascular disorders such as vibration-induced white fingers, neurological issues like peripheral neuropathy and carpal tunnel syndrome, and musculoskeletal problems including tendinitis and joint disorders. The vascular component involves a secondary form of Raynaud’s phenomenon, where prolonged exposure causes damage to blood vessels, exacerbating sensitivity to cold and resulting in episodic finger blanching. Both WBV and HTV contribute to muscular weakness, reduced hand-grip strength, and chronic fatigue in the affected regions, potentially leading to disability. Furthermore, vibration exposure

can provoke or exacerbate other ergonomic stress-related conditions, compounding the risk of musculoskeletal disorders [23].

WBV exposure adversely affects people by impacting gastric motility. The study on healthy male volunteers revealed that WBV, particularly at frequencies of 4 Hz and 8 Hz, leads to a notable decrease in the amplitude of EEG waves and the power spectrum corresponding to slow wave components, both in fasting and post-food intake states [24]. This suppression of gastric smooth muscle activity and contraction waves is likely a result of mechanical resonance and increased neurohumoral regulation triggered by vibration stress. Consequently, such alterations in gastric motility may contribute to gastrointestinal disorders, as long-term WBV exposure has been associated with conditions like gastric neurosis and non-ulcerative dyspepsia. Additionally, the suppressive effects on gastric motility were found to be more pronounced when combined with food intake, suggesting that stomach contents amplify the adverse effects of WBV.

The vibration impacts were concerning in military and vehicular contexts, where personnel were subjected to constant mechanical vibrations from vehicle operations [25]. Vibration affects human performance, including physical, perceptual, and cognitive abilities. Physically, vibration impaired manual control and tracking tasks. For example, vertical vibrations between 4 to 6 Hz made precise tasks like writing exceedingly tricky, with performance deteriorating linearly with increased vibration magnitude. Tracking errors were most significant, with vertical vibrations around 5 Hz, while horizontal vibrations below 3 Hz also posed challenges to performance. Visually, vibration caused relative movement between the retina and visual displays, leading to blurred images and decreased visual acuity. This effect was pronounced in the 2 to 20 Hz frequency range,

where the body's primary and secondary resonances occur. The vestibulo-ocular and vestibulo-spinal reflexes, which stabilise gaze and posture, were less effective at higher vibration frequencies, leading to impaired visual performance. Cognitively, vibration negatively affected more complex tasks involving memory and problem-solving. While simple reaction time tasks were not significantly impacted by vibration, complex cognitive tasks, such as those requiring short- and long-term memory, decreased performance. Notably, vibration exposure around 16 Hz had been found to impair short-term memory tasks, with increased reaction times and errors. An important point highlighted in [25] was the significance of the nauseogenic frequencies (< 0.5 Hz), which can induce motion sickness symptoms like drowsiness, headache, nausea, and vomiting. These symptoms further exacerbated the negative impacts of vibration on performance. Therefore, measuring and mitigating vibration exposure in environments such as military vehicles is unavoidable for maintaining the health and performance of personnel.

The adverse effect of vibration can also be found in other fields, such as the transportation of goods [23], where they cause physical damage to products, alter their quality, and result in economic losses. Vibrations during transport, particularly by truck, imposed severe loads on cargo, leading to potential damage and loss. For instance, the phenomenon of 'silo quaking' in bulk solid transportation caused structural damage and significant discomfort for personnel due to intermittent high-frequency vibrations. Vibrations also influence products' sensory and quality attributes, such as texture and shelf life, especially in perishable goods like fruits and vegetables.

2.2.4 Train-Induced Vibrations

Generating mechanisms of rail vibration in railways is very complex as many components are involved, as shown in Figure 2-4.

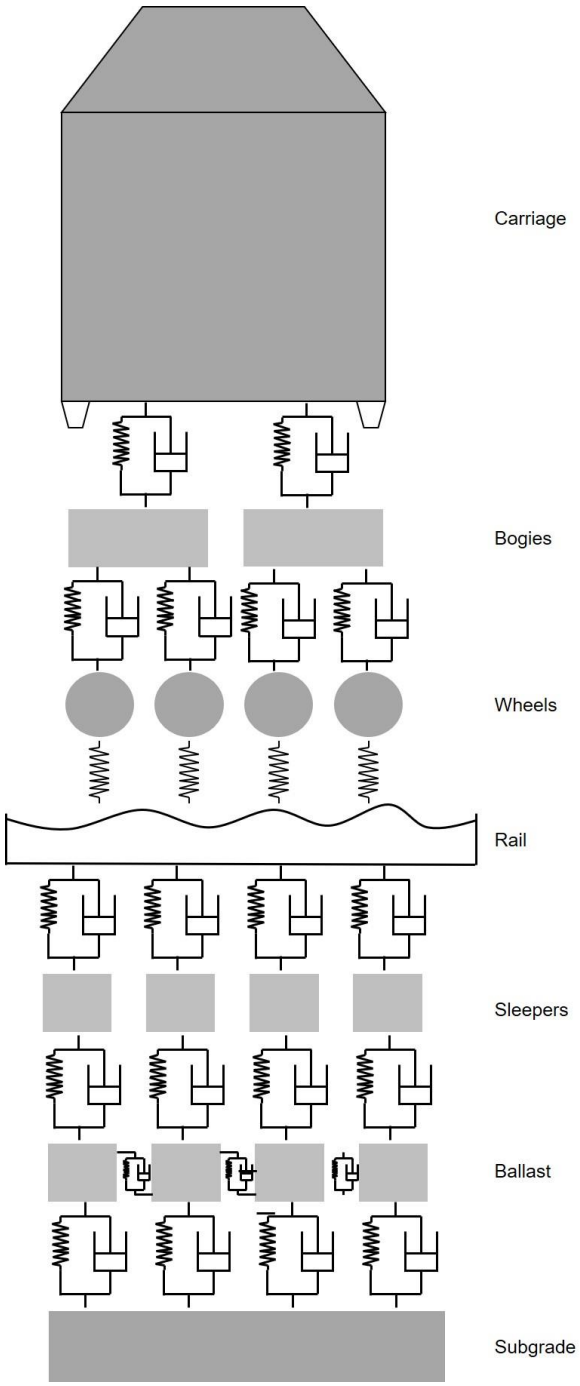


Figure 2-4 Sources of railway vibration [26]

A train moving along a track generates vibrations in both the wheels and the track [26]. The vibration of the wheels is affected by the system above them, such as the bogie, springs, dampers, and vehicle load. Meanwhile, the vibration of the rails is impacted by the system below, including the track, subsoil and soil. As the rail surface is not smooth, the wheels inevitably move in the vertical direction while the rails are not utterly stiff, so they also move vertically, leading to excitation to rail pads and sleepers. The rails are supported by discrete sleepers secured by a ballast bed. The subsoil beneath the ballast varies in composition and elasticity. Due to the unevenness (surfaces or variation in the support stiffnesses), wheels and the track are subject to a dynamic force before they vibrate. The vibration is fiercer when the exciting force at the wheels and rails resonance frequencies. There are six vibration developing mechanisms can be summarised into quasi-static and dynamic excitation in Table 2:

Table 2 Train vibration-generating mechanisms [26]

Types of excitation	**Cause**	**Description**
Quasi-static	Train mass / Moving load	It results from the mass of the moving train at a certain speed. It can cause vibrations at a shallow frequency.
Dynamic	Parametric excitation	Parametric excitation is caused by regular variations such as rail fixations and sleepers. Additional sources of parametric excitation include variations in sleeper spacing, bogie spacing, or inter-vehicle spacing.
	Unevenness of the track	It contributes to dynamic excitation with wavelengths ranging from 0.1 to 1 meter. This unevenness can stem from subsoil, ballast bed variations, or misalignments in the track.

	Rail corrugation	Rail corrugation results from period wear on the rail head with wavelengths ranging from 0.01 to 0.05 meters.
	Track singularities	Track singularities at the wheel/rail contact, such as uneven joints, switch and crossings, cause abrupt changes in the contact dynamic.
	Unevenness of the wheel surface	The unevenness of the wheel surface includes wheel out-of-roundness and wheel roughness. Out-of-round wheels induce broadband vibration, while wheel polygonisation, where wheels develop flat spots or a polygonal shape due to irregular wear, causes modular behaviour and periodic impacts.

The vibration frequencies in Table 2 are sensitive to the train's speed. Table 3 depicts how the frequencies adjust following the speeds of 40 km/h, 80 km/h, and 160 km/h.

Table 3 Frequencies of vibrations response to different speeds [26]

Speed (km/h)	**40**	**80**	**160**
Moving load (axle spacing ~1.8 m)	3 Hz	5 Hz	11 Hz
Track unevenness	1 - 100 Hz	2 - 200 Hz	4 - 400 Hz
Rail corrugation	~500 Hz	~1000 Hz	~2000 Hz
Wheel unevenness	≥ 4 Hz	≥ 4 Hz	≥ 4 Hz
Wheel polygonisation (if a wavelength of 0.1 m)	~100 Hz	~200 Hz	~400 Hz
Inter bogie spacing (if ~8m)	~1Hz	~3 Hz	~5 Hz
Sleeper spacing (0.6 m)	Multitudes of 16 Hz	Multitudes of 32 Hz	Multitudes of 64 Hz

As the train's speed increases, the frequency of vibrations resulting from factors such as moving load, track unevenness, rail corrugation, and wheel polygonisation also increases. For instance, at a speed of 40 km/h, the moving load causes vibrations at 3 Hz, whereas at 160 km/h, it causes vibrations at 11 Hz. Similar trends are observed for other factors like track unevenness, which ranges from 1-100 Hz at 40 km/h to 4-400 Hz at 160 km/h.

According to [13], exposure to vibrations at frequencies (1 – 80 Hz) can lead to discomfort and annoyance, causing degraded quality of the train journey. The long duration of the vibrations can worsen the effect of the discomfort as vibrations can cause sustained physical fatigue and mental strain [27].

2.3 Standards for Train Ride Comfort Evaluation

Vibration is recognised as the primary contributor to RC, which is also affected by other parameters such as temperature, humidity, acoustic noise, smell, design layout, visual stimuli, and service quality. Once vibrations are triggered by train motions, they are transmitted to the human body through the contact between the passengers and the train. For seated passengers, the contact areas are categorised into seat, backrest and floor, while only the floor is for standing passengers. Although the generation of vibration is discussed, the physical magnitude of vibration does not directly correspond to how people perceive it. Thus, the magnitude is adjusted to reflect actual human perception, considering factors such as magnitude, frequency, and direction. Numerous studies have investigated how to correlate the physical effects of vibration with human perception. There is no universal standard to evaluate the RC as a broad range of parameters affects the RC. Nowadays, there are many methods for evaluating railway vehicle running

stability and RC, such as the most widely used four standards: ISO 2631 [28, 29], UIC 513 [30], the Sperling index [31], and European Standard (EN) 12299 [32].

ISO 2631 is a comprehensive standard providing guidelines for evaluating the effects of WBV on human health, comfort, and perception. Using Root Mean Square (R.M.S) values and Vibration Dose Value (VDV), comfort is categorised into different levels. UIC 513 adapts IOS 2631 specifically for railway applications with the unique vibration characteristic of trains. It uses statistical analysis methods to process sampled data and correlates the frequency-weighted R.M.S accelerations and the comfort level. The Sperling index focuses on the stability of the vehicle rather than the RC. Although it does not account for the duration of the vibration exposure, it provides valuable insights for optimising damping parameters of lateral and vehicle stiffness in vehicle design, which is helpful for engineers refining vehicle suspension systems. EN 12299 is the same as the UIC 513 but has different threshold values for each comfort level. EN 12299 is broadly applied in Europe and emphasises the importance of vertical and lateral vibrations and their synergy on the RC.

This section discusses the standard for the vibration-based method only, as there are no standards for the questionnaires and the physiology-based method.

2.3.1 Application of ISO 2631 Standard

Using ISO 2631 for RC evaluation, many attributes can affect the results, such as speed, seat design and train load, as studied in [33]. Findings indicated that the speed presents the most effect on the RC and the most negligible impact from the loading. This is further supported by [34], which correlated train operation speed and RC, as calculated by ISO

2631, highlighting that the optimal speed can enhance RC and mitigate health risks associated with vibration exposure.

As illustrated in Figure 2-4, the infrastructure and vehicle conditions significantly impact vibration generation. P Zoccali et al. used ISO 2631 to assess RC in railways, enabling effective continuous monitoring and maintenance by identifying critical areas with potential discomfort due to vertical, lateral and longitudinal accelerations [35]. This finding is corroborated by [36], which conducted RC experiments in a suburban railway using an Arduino-based comfort monitoring system, unfolding that vibration levels frequently exceed ISO 2631 thresholds, indicating immense discomfort, particularly in the seat pan and backrest. A comparative study of new and old trains in RC based on ISO 2631 was conducted in [36], concluding that new trains with better vehicle conditions are more comfortable within an 8-hour exposure period. The combination of fast speeds and poor track conditions increases discomfort as measured by ISO 2631 [37]. As specified in ISO 2631-4, it is not sensible to regard a train as a rigid body, as vibration levels vary at different measuring points on board a train. L Bakinowski and B Firlik [38] measured vibration levels at various tram locations, including the floor and passenger seats, under controlled conditions at 40 km/h. They found the highest vibration levels on the floor above bogies, while the passenger seats substantially dampened these vibrations to enhance RC.

Other applications of ISO 2631 to study the impact of track irregularity and track geometry have also been conducted. Irregularities were considered a dominant contributor to excite vibrations, leading to worse RC [39]. Interestingly, the track geometry, including the

effects of cant, the rates of cant change, and the radius of curves, has a minor influence on RC [40].

2.3.2 UIC 513 & EN 12299 Standards

This section showcases the application of UIC 513 and EN 12299 combined into one section due to their similarities. In a comparative study [41], UIC 513 and the Sperling index were applied to the same dataset, considering factors such as velocity, carboy flexibility, and suspension damping. The findings revealed notable differences: UIC 513 indicated higher comfort at high velocities, whereas the Sperling index reported more discomfort due to its sensitivity to dominant vibration models like bounce and pitch.

EN 12299 was utilised to calculate RC across different trams in [38], finding a large variability even among trams of the same type, underscoring the importance of regular maintenance. Additionally, EN 12299 was employed to design a railway transition curve, comparing traditional circular arcs with a new polynomial curve of the 9^{th} and 11^{th} degrees [42]. It was concluded that the optimised curves can reduce a large amount of discomfort. Apart from factors such as track and vehicle conditions and speed, an interesting factor affecting RC – crosswind has also been evaluated based on EN 12299. It was found that wind load affects RC more than track irregularities, especially at high speeds [43]. A comparative analysis between EN 12299 and the Chinese Spering index revealed that the latter is more sensitive to vertical vibrations, while the former focuses on lateral vibrations [43].

L Leicht et al. explored the impact of reduced powertrain masses on the Flexity Zurich tram's RC using EN 12299 and the Sperling index [44]. Simulations in [44] show no correlation between the mass distribution and the RC. R Corradi et al. leveraged a

numerical method based on the EN 12299 standard to estimate RC in an HST vehicle during a design phase. Results show that the RC remains within acceptable limits across various speeds, positions, car body flexibility, and track irregularity. The research also highlighted the critical speed phenomenon, where resonance excitation affects comfort levels.

Furthermore, G Isacchi et al. investigated how the installation angle of yaw dampers affects RC in HSTs. Their multibody simulations revealed that changing the damper angles from 0° to 10° substantially redistributes accelerations perceived by passengers, shifting discomfort (based on EN 12299) areas from the rear to more central and lateral potions of the car body [45].

2.3.3 The Sperling Index

Various studies have employed the Sperling index using different methodologies and algorithms, sometimes leading to inconsistencies in its application and results. In [46], the Sperling index was used to assess a tram's comfort levels in Cuenca, Ecuador. The study utilised the Freematics One+ device to collect acceleration data, which was subsequently processed to determine RC using the EN 12299 and the Sperling index. The results from the two standards coincided and showed that the RC was comfortable. The relationship between the Sperling index and the EN 12299 was further explored in [47]. They found a significant correlation between the two methods and suggested they are reliable for comfort evaluation in vertical vibrations. Their study also reinforced the utility of the Sperling index and highlighted the need for consistent application of the calculation method. C Deng et al. investigated the effects of various calculation methods on the Sperling index [48]. They identified that the choice of algorithm and sampling time

significantly decide the results. For instance, if the index is based on the second acceleration power, the results are stable across different sampling times, whereas the third power basis showed variability. Therefore, they concluded that the algorithm based on R.M.S values is preferable for consistent and reliable application. This conclusion is also supported by [49], which is concerned about the variability in the Sperling index results depending on the calculation approach.

While the Sperling index remains a robust tool for evaluating RC in rail vehicles, its application is different by the choice of algorithms and sampling intervals. The studies in this section consistently indicate that using R.M.S.-based is more reliable. These insights underscore using R.M.S-based calculations such as ISO 2631 and UIC 513 in this thesis.

2.4 Traditional Train Ride Comfort Measurement Methods

This section presents information for RC using traditional methods: (1) questionnaires, (2) physiology-based methods and (3) vibration-based methods.

2.4.1 Questionnaire-Based Assessments

Questionnaires are widely used to assess RC by capturing subjective evaluations of factors affecting comfort. This section synthesises findings from studies on how questionnaires have been used to estimate RC. Questionnaires typically encompass sections on demographic information, trip details, and specific comfort-related questions. For instance, [50] used a 9-point Likert scale and a 1-100 magnitude estimation scale, finding that seat dimensions significantly impacted comfort while personal fatigue had a negative effect.

Many studies focused on environmental factors. For instance, [51] investigated how passengers respond to different conditions, concluding that higher temperatures and passenger density reduced comfort and caused notable discomfort above 30°C. An assessment of long-distance trains in Iran highlighted that passengers were most satisfied with lighting and least satisfied with temperature control, with elderly passengers being more sensitive to vibrations [52]. Noise and vibration were also critical factors, as studies in Kuala Lumpur's urban rail transit system unveiled differences in comfort ratings between rail systems, emphasising the importance of noise and vibration management [53]. Swedish train passengers identified rattling and squeaking noises as the most annoying, especially near entrances [54], while regular commuters' assessments further underscored that noise was an essential determinant of discomfort [54]. HST passengers' evaluations of tunnel noise in [55] suggested a potential effect of noise on pressure discomfort and provided insights into the relationship between acoustical and physical comfort factors.

Questionnaires have been evidenced to provide an all-inclusive platform that captures a wide range of data and subjective perceptions, such as comfort measures, demographic information, and some specific physical environmental factors like temperature, noise, and seat characteristics. It further allows very fine detailed feedback and assessment of psychological constructs, providing a holistic view of passenger comfort. However, the principal weakness of questionnaires is that they only represent self-reported data, which are subjective and hence are open to biases and errors based on the mood, personal preference, and situation of the participant. Likewise, the scope of questionnaires might be limited as they would only represent those factors of comfort that were pre-decided upon and might miss other relevant factors.

In contrast, the cause-effect relationship between the factors and the effect of vibration is built up by systematically studying each variable individually or simulating real-world conditions to observe their effect on actual tasks [13]. Therefore, the standard of vibration is more reliable and trustworthy.

Vibration data collected using smartphones represent a more objective and quantifiable approach. Real-time monitoring and data collection through this method presents a dynamic view of vibration levels during the journey. It is easy to scale up the data collection using smartphones to cover many trips and passengers, which is logistically difficult if one has to administer surveys. However, ML techniques can process vast amounts of data to identify patterns and allow accurate predictions about the RC, probably uncovering insights that subjective reporting might have overlooked.

2.4.2 Physiology-Based Comfort Evaluations

Referring to Table 1 under physiological parameters, some indicators, including muscle and skeleton, central nervous system, and peripheral nervous systems, are used to estimate RC in studies.

In investigations concerning seating comfort, biomechanical human models can be used to estimate internal loads, like trunk muscle forces and spine loads, which are usually difficult as they are expensive or invasive to subjects. These internal loads are relevant for assessing seating comfort and give further information on how design guidelines for products interacting with the human body can be established. Lumped parameter (LP) models [56] have been developed to evaluate the RC [56]. LP models with 1–12 degrees of freedom have been developed to provide biodynamic response and vibration transmissibility data [57, 58]. These models use lumped masses connected by massless

springs and dampers to simulate passive biodynamic responses under different conditions. This configuration of a simple spring-damper may not be enough to capture human tissues' complex viscoelastic properties. Furthermore, the spring-damper model cannot represent the complex conditions in a moving train.

EEG is a method to record the brain's electrical activity. Y Peng et al. used EEG to quantify RC in HSTs by analysing the brain activity of 20 participants during travel [59]. Participants wore 64-channel EEG caps and provided comfort ratings every ten minutes. Increased beta band activity in specific brain regions (BA6, BA13, BA20, BA24, BA31, BA40, BA47) correlated with discomfort. These regions are linked to environmental perception, emotion processing, and movement control. A lightGBM model trained on EEG features to predict comfort levels. However, EEG has low spatial resolution (a few millimetres to one centimetre) and is highly susceptible to motion artefacts.

For peripheral nervous systems, electrocardiogram (ECG) measures the heart's electrical activity through electrodes placed on the skin [60]. Heart rate variability (HRV) refers to the variations in the intervals between consecutive heartbeats as measured by an ECG. It reflects the balance of the autonomic nervous system and has been applied in fatigue detection [61], emotion recognition [61], and comfort assessment [62]. Al-Nashash et al. [63] and Masahiro et al. [63] demonstrated that vibration influences the normalised difference in power spectral density from ECG and HRV. However, research on RC related to ECG in HSTs is lacking. The advent of wearable sensors presents an opportunity to integrate ECG features with vibration analysis. Such integration provides more insights into passenger comfort. However, the widespread adoption of this

integration may be impeded by the fact that not all individuals utilise wearable devices now.

2.4.3 Vibration-Based Assessment Techniques

This section aims to show the traditional methods of acquiring vibrations. The main methods fall into simulations and using accelerometers. Studies are tabulated in the following table.

Table 4 Traditional methods to capture vibrations for RC

Title	**Methodology**	**Remarks**
Correlation between ride comfort index and Sperling's index for evaluation ride comfort in railway vehicles [47]	Simulation	The mechanical model in this study was employed to assess the RC of a passenger train travelling at a constant velocity of *V* on a track characterised by vertical irregularities. This vehicle was represented through a rigid-flexible coupled model, incorporating an elastic beam for the car body and six rigid bodies corresponding to the bogies and axles.
Study on the evaluation methods of the vertical ride comfort of railway vehicle—mean comfort method and Sperling's method [41]	Simulation	A model of a rigid-flexible coupled vehicle was utilised to assess the RC while operating at a constant velocity on a track with vertical irregularities.
Simulation for WBV to assess ride comfort of a low–medium speed railway vehicle [49]	Simulation	The car body is represented as an Euler–Bernoulli beam with a constant cross-section and uniformly distributed mass characterised by the bending modulus *EI*.
Evaluation of Sperling's index in passenger and freight trains under	Simulation	This study used Multi-Body (MB) Dynamics software to simulate train dynamics.

different speeds and track irregularities [64]		
Dynamic train vertical Sperling Index evaluation model considering wheel-rail contact loss [65]	Simulation	This study illustrated the mechanical model of a four-wheel, two-suspension vehicle, incorporating three equivalent passenger masses distributed at the midpoint and two points where the vehicle body connects with the bogies.
Ride comfort assessment of high-speed rail vehicles: influence of yaw dampers installation angle [45]	Simulation	A flexible car body has been modelled in Abaqus using a FE approach for integration into the Simpack multibody model. This model aimed to simulate the modal behaviour of a typical car body in a modern HST.
The effect of onboard passengers' seating arrangement on the vertical ride comfort of a HST [66]	Simulation	This study schematically depicted the vertical motion model of a railway vehicle with passenger seats. The rigid degrees of freedom for the car body and bogie frames include bounce and pitch, while the wheelsets have bounce. The car body was represented as a uniform Euler–Bernoulli beam with free boundary conditions to study flexural vibrations. Additionally, the regular contact forces between wheelsets and rail are modelled using linearized Hertzian theory.
Influence of modified powertrain masses on the ride comfort of a light rail vehicle [44].	Simulation	a MB model of Flexity was developed using Simpack. Subsequently, a simulation was conducted for a run.
Integrated modelling and simulation of vehicle and human multi-body dynamics for comfort assessment in railway vehicles [39]	Simulation	This study presented an integrated vehicle-biodynamic model developed using multi-domain (energy domain simulations) and the multibody simulation software ADAMS VI-Rail. These models evaluated the RC in a railway vehicle, considering vertical and lateral excitations at the wheels.
Evaluation of passenger ride comfort of Indian rail	Simulation	A coupled vertical-lateral rail vehicle model has been developed using

and road vehicles with ISO 2631-1 standards: Part 1 - mathematical modelling [67]		Lagrangian dynamics, with 37 degrees of freedom assigned to 9 rigid bodies.
Development of a reduced dynamic model for comfort evaluation of rail vehicle systems [33]	Simulation	This paper presented an analytical reduced dynamic model of a rail vehicle system with 38 degrees of freedom.
Analysis of tram comfort using the UNE EN 12299 standard Sperling method [46]	Sensors	The Freematics One+ devices, which incorporate accelerometers, were utilised to measure acceleration data. These devices are also equipped with GPS sensors to record global positioning data.
Analysis of the consistency of the Sperling index for rail vehicles based on different algorithms [48]	Sensors	Accelerometers were used to measure vibrations at different points on the rail vehicle. The sensors were mounted on steel plates to ensure a stable and accurate measurement environment.
Momentary discomfort of HSTs passing through complex terrain sections under strong wind conditions [68]	Sensors	Accelerometers were installed on the floor at the centre and both ends of the passenger compartment. This placement was in accordance with the vibration comfort measurement method specified in EN-12299.
An investigation into evaluation methods for ride comfort of railway vehicles in the case of car body hunting instability [69]	Sensors	A sophisticated onboard data acquisition system was utilised to record acceleration data. This system comprises several high-precision accelerometers to capture vibrations in multiple directions.
Evaluating methods of whole-body-vibration exposure in trains [70]	Sensors	The instruments used for vibration measurement were triaxial accelerometers. These devices were mounted on the railway vehicle's seat pan, backrest, and floor to capture the vibrations experienced by seated and standing passengers. The accelerometers used in this study included MAESTRO vibrometers and

		01dB triaxial seat-pad accelerometers.
Effect of train speed and track geometry on the ride comfort in high-speed railways based on ISO 2631-1 [35, 40, 71, 72]	Sensors	Three piezo-electric accelerometers were mounted at the top corner of an aluminium block (15 cm × 15 cm × 15 cm, 6.23 kg) and placed on the floor carpet by the subject's feet. Sit pads with tri-axial accelerometers were placed at the seat pan and backrest. Inertial Measurement Unit (IMU) were Placed on the floor of the trailing cab.
A factorial analysis for the determination of an optimal train speed with a desired ride comfort [34].	Sensors	Three B&K accelerometers, type 4370, were employed for vibration acquisition.
Developing a Ride Comfort Monitoring System from Scratch: An Experience in a Suburban Railway [36]	Sensors	The MPU 6050 GY-521 accelerometers were used to acquire vibrations.

It is observable that many studies are favourable to the use of simulations. Simulations can provide a controlled environment where variables (track geometry, vehicle dynamics, environmental factors, etc.) can be precisely controlled. This control allows for detailed analysis using complex models that incorporate multiple degrees of freedom, and non-linear behaviours can provide a large amount of data of interest, which can be expensive, time-consuming and labour-intensive to collect in the real world. Safety and scalability are also attractive strengths provided by simulation tools. Simulation is free of risk from installing equipment and is scalable to explore different settings like speeds, track conditions, and vehicle designs. However, simulations also bring some limitations. The accuracy of simulations is subject to the simplifications and assumptions to model real-

world conditions. Accurate simulations demand substantial computational resources and are limited by the capabilities of the software used. Despite the usefulness of simulations, real-world testing remains essential for validating results, as some environmental factors, such as weather conditions, track wear, and passenger movements, are challenging to simulate accurately. Meanwhile, the accuracy of simulations is susceptible to data quality.

Dedicated sensors, such as piezo-electric and triaxial accelerometers, provide high sensitivity and broad frequency response, essential for capturing detailed and precise vibration data. These sensors are often calibrated against known standards to ensure their measurements are accurate and reliable over long periods and under various environmental conditions. Additionally, sophisticated onboard data acquisition systems allow high-frequency sampling rates and real-time data processing to capture vibrations in multiple directions for comprehensive analysis. However, the high cost of the sensors and the complexity of their setup and maintenance can be prohibitive. Installation requires permanent or semi-permanent modifications to the vehicle, which poses logistical challenges.

In contrast, smartphones offer a highly accessible and cost-effective alternative for vibration data collection. Their ubiquity and ease of use enable widespread data collection with minimal technical expertise. Numerous applications facilitate real-time data visualisation, GPS integration, and cloud storage, making smartphones a versatile tool for quick, large-scale deployments.

Studies have validated the reliability of smartphone accelerometers. Meng et al. [73] conducted calibration tests comparing an iPhone's MEMS accelerometer with a high-sensitivity piezoelectric accelerometer, finding strong agreement in both time and

frequency domains. Similarly, [74] tested a smartphone and a piezoelectric accelerometer on a shaking table regarding X-, Y-, and Z- axes accelerations. The results showed remarkable consistency between the devices, confirming the accuracy and reliability of the smartphone sensor.

According to [35], passengers are most susceptible to vibrational discomfort at low frequencies. Empirical studies reveal that humans are susceptible to vibrations in the vertical axis within the 4-16.5 Hz range and in the longitudinal and lateral axes within the 0.6-2 Hz range. Complementary findings from [75] indicate that vibration levels are notably below 20 Hz in high-speed and conventional rail systems. Thus, using a smartphone to capture vibration in the context of a train RC is sensible.

2.5 Machine Learning Applications in Train Ride Comfort

ML is a subset of Artificial Intelligence (AI) that can identify patterns and make decisions based on large datasets. It enables systems to learn from data and predict using new data without explicit programming.

ML is particularly well-suited for predicting railway RC due to its capabilities in handling large datasets, real-time prediction, scalability, and computational efficiency. The application of ML in RC is well-documented across various domains, such as railway and automotive, demonstrating its versatility and effectiveness. In railway systems, train vibration data can be efficiently processed using ML algorithms. For instance, the Gated Recurrent Unit (GRU) neural network employed by Z. Yu et al. demonstrated superior performance and computational efficiency over other models like Multilayer Perception

(MLP) and non-gated Recurrent Neural Network (RNN) [76]. This efficiency is critical for real-time prediction, where quick and accurate RC assessments are essential.

Additionally, ML models such as Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM) have shown high accuracy in predicting comfort based on various input features, including physiological signals such as EEG [59, 77] but more features such as compartment seating, ambient temperature were used in [77]. The researchers in [78] correlated different sitting postures into different comfort levels with the gradient boosting machine algorithm, achieving an accuracy of 89.5%. ML for comfort analysis is not just limited to railway passengers. G Kyung et al. employed ML on car seats to predict comfort and discomfort using Principal Component Analysis (PCA) and regression on interface pressure data [79]. By analysing 36 pressure variables, the study identified that comfort is more closely linked to pressure distribution.

The ability of ML to handle different input features, its versatility in different contexts, and the nature of saving computational cost make ML a suitable tool to process large datasets, provide real-time predictions and maintain computational efficiency. Its proven performance across different contexts and input features underscores its suitability for enhancing passenger experiences in high-speed rail systems.

2.6 The Role of Smartphones in the Train Ride Comfort

Smartphones have been used to acquire data in many studies because of their real-time collection, ubiquitousness, and cost-effectiveness, emphasised by the following studies.

The feasibility of using smartphones to acquire vibration data was studied by [80] recording vibration data using the phone's accelerometer and comparing it to standard

measurements from traditional accelerometers. With the development of an appropriate application and an Artificial Neural Network (ANN) for data analysis, smartphones can provide a reliable and cost-effective method for passengers to monitor RC. The feasibility is further confirmed by N Tien Do et al., validating smartphone sensors against conventional wireless sensors, demonstrating that smartphones can provide reliable data for assessing ride quality [81]. They deployed smartphones in several ways: attached to the floor, seat level, and passenger pockets. These various placements allowed us to capture data from different perspectives within the train. The study also highlighted the utility of using smartphones due to their geolocation capabilities, which help identify track components such as rail joints, switches, curves, bridge transitions, and crossings that adversely affect ride quality.

Rodriguez et al. explored the innovative use of smartphones and tablets in railway track monitoring and RC assessment [82]. These mobile devices were used to measure vertical accelerations within train carriages. This approach significantly reduced costs and simplified the traditionally complex and expensive process of railway monitoring. One of the key novelties of this study was the application of everyday mobile devices for continuous data collection and analysis, which enabled the quick identification of problematic track sections. Moreover, the study demonstrated that smartphones can effectively measure RC by providing real-time data on vertical accelerations directly related to RC.

Apart from vibrations, other features that smartphones can capture are vehicle acceleration, velocity, and location [82], as well as humidity, temperature, noise, and lighting level [83]. The real-time data was processed to evaluate indices that enhance the

ability to ensure RC and operational safety, making them valuable in intelligent transportation systems.

Smartphones were utilised to investigate the discomfort of railway standing passengers in Singapore [84] and bus passengers in Italy [85]. Passengers in [84] used a stopwatch to log instances of discomfort, which were then correlated with the acceleration data acquired by smartphones using logistic regression, highlighting the advantages of smartphones due to their portability, availability, and precision. In assessing the comfort of standing passengers on a bus [85], smartphones recorded GPS coordinates and accelerations along various axes. A comprehensive real-time comfort scale was developed by integrating objective data with subjective comfort ratings from passengers.

2.7 Research Gaps

While traditional methods for assessing RC in railway systems, such as questionnaires and vibration-based assessments, have provided some insights, they present several critical limitations. Questionnaires, for instance, rely heavily on subjective passenger feedback, resulting in inconsistent and unreliable data that often lack objectivity. Furthermore, these surveys are time-consuming, difficult to apply on a large scale, and fail to capture real-time ride comfort, making them unsuitable for immediate comfort adjustments during a journey. Vibration-based methods, which use specialised accelerometers, require complex setups, expert knowledge, and significant numbers of sensors to capture different passenger postures and locations. This not only makes them costly and time-consuming but also restricts their scalability and feasibility for continuous or widespread use across entire train systems. Moreover, traditional methods typically employ post-journey data analysis, limiting the ability to offer real-time feedback or make

operational adjustments during a journey. This significantly reduces their effectiveness in enhancing immediate passenger satisfaction.

Existing approaches are subjected to handle the vast, complex datasets generated by dynamic systems like moving trains, requiring substantial manual effort and time for data processing. These systems cannot provide real-time or continuous monitoring of RC, which is essential for modern railway operations aiming to proactively improve the overall passenger experience. Financial and logistical barriers also persist, with high costs and the logistical complexities of deploying and maintaining specialised hardware making large-scale implementation difficult.

This research introduces smartphone sensors in conjunction with ML algorithms for real-time RC assessment to address these gaps. Already widely available and equipped with advanced sensing capabilities, smartphones provide a scalable and cost-effective alternative to traditional accelerometers. By leveraging crowd-sourced data from multiple smartphones, the proposed approach enables continuous, real-time monitoring across entire train systems, something previous studies have not accomplished at scale. This research also integrates various ML techniques to automatically process large datasets, reducing the need for manual data handling while offering more accurate and comprehensive assessments of RC. This system's real-time processing and scalability allow railway operators to make immediate adjustments during a journey, improving both passenger satisfaction and operational efficiency in ways that traditional methods could not.

Chapter 3

Methodologies

3.1 Introduction

Figure 3-1 provides an overview of the methodology employed, divided into three stages: data collection, data processing, and modelling. In the data collection phase, a combination of synthesised and field data is gathered, with a strong emphasis on maintaining data integrity to ensure the reliability of subsequent analysis. The collected data is then passed to the data processing phase, which undergoes essential pre-processing steps, such as cleaning and transformation, to prepare it for further analysis. Additionally, a comfort index is calculated at this stage as an output for the model. The processed data is subsequently fed into the modelling phase, where AI models are

developed. This stage also includes performance evaluation to measure the model's accuracy and effectiveness.

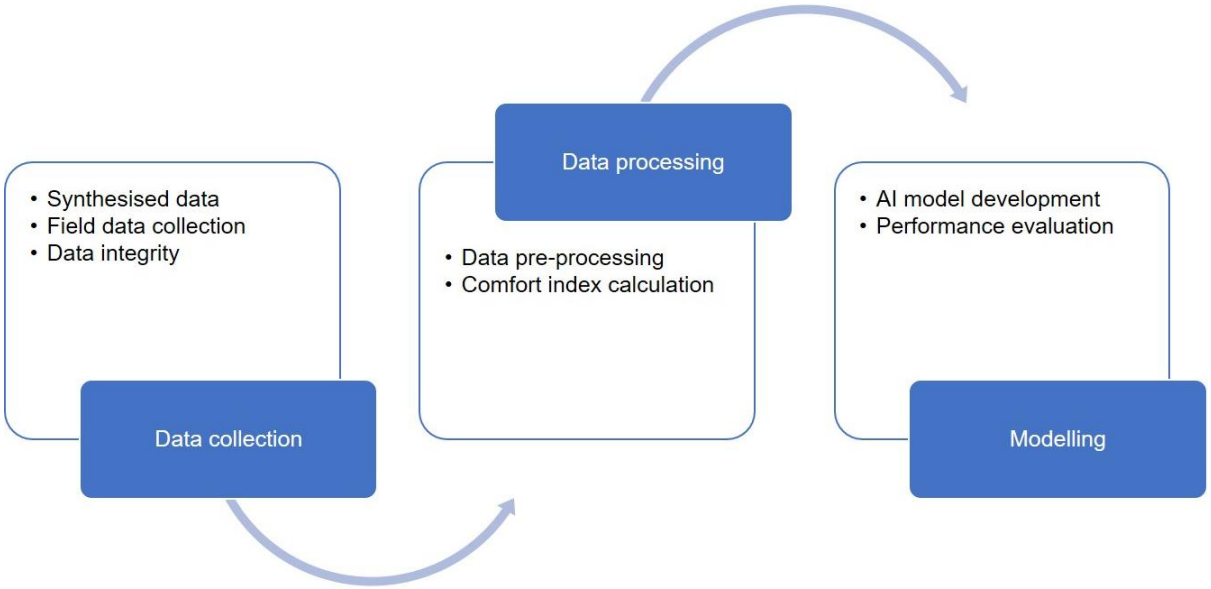


Figure 3-1 The flowchart for methodologies

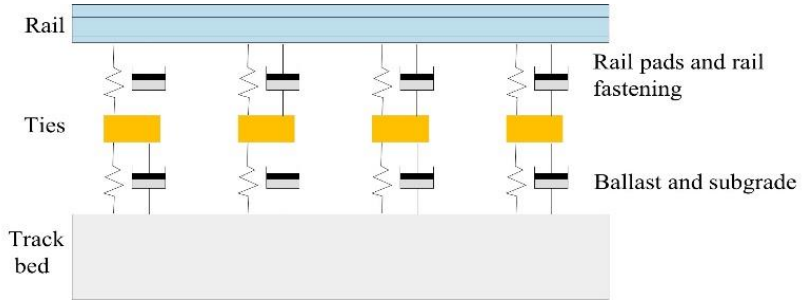
3.2 Data Collection Techniques

3.2.1 D-track Simulations

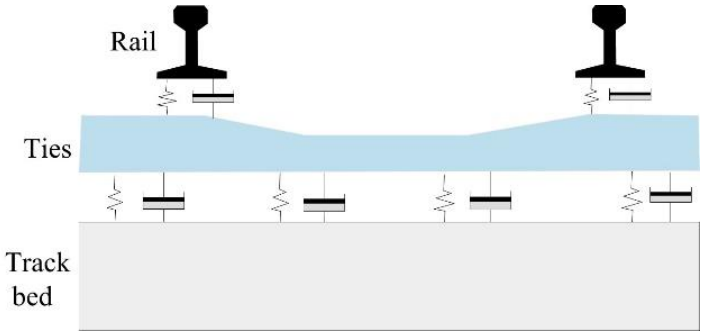
D-track is the MB simulation package used in this thesis to generate vibration data to quantify the correlation between one of the track characteristics – track stiffness and other factors of track and operation. This section commences with D-track validation to lay the profound foundation and trustworthiness for the simulated data.

Initially developed by Cai [86], the D-track is a nonlinear FE simulation package designed based on an ideal rail track model depicted in Figure 3-2(a). This model conceptualises the rails and ties as two elastic beams, with the rails supported by the ties through rail

pads and fastening mechanisms and the ties resting on the ballast and subgrade. Figure 3-2(b) illustrates non-uniform concrete ties.



(a)



(b)

Figure 3-2 Idealised track vibration model (a) longitudinal view of the track (b) cross-section of the track

To enhance the stability of the rail elements, two spring stiffness at each end of the rail span are utilised to represent the inertial, flexural, and resilient properties of the rail support components, as shown in Figure 3-3.

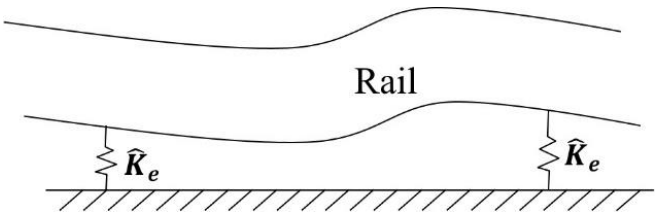


Figure 3-3 Uniform rail beam segment

Cai [86] simplified one span of the rail track into a uniform rail beam segment supported by two spring coefficients $\widehat{K}_e$ shown in Figure 3-3 and defined in Equation 2.

$$\widehat{\boldsymbol{K}}_{\boldsymbol{e}} = \frac{\boldsymbol{k_p}}{\mathbf{1} + \boldsymbol{k_p} \sum_{\boldsymbol{n=1}}^{\boldsymbol{z}} \frac{\mathbf{2}[\boldsymbol{z_{n(d_r)}}]^{\mathbf{2}}}{(\boldsymbol{w_n^2} - \boldsymbol{\Omega^2})\boldsymbol{M_n}}} \qquad \text{Equation 2}$$

Where k_p is the contact stiffness between two beams in Figure 3-2, z_n represents the n_{th} mode of the tie beam, d_r is the distance between the end of the tie and the rail seat, Ω is the vibration frequency of the track, w_n is the natural frequency of the tie, and M_n stands for the tie mass.

The spring coefficient defined $\widehat{K}_e$ in Equation 2 is integrated with the exact dynamic stiffness matrix of the rail span, described by the subsequent equation. Detailed explanations of each element in Equation 3 are provided in [86], resulting in displacement responses to the nodal forces as expressed in Equation 4.

$$[\boldsymbol{K_r}] = \begin{matrix} \boldsymbol{k_{11}} + \widehat{\boldsymbol{K}}_{\boldsymbol{e}} & \boldsymbol{k_{12}} & \boldsymbol{k_{13}} & \boldsymbol{k_{14}} \\ \boldsymbol{k_{21}} & \boldsymbol{k_{22}} & \boldsymbol{k_{23}} & \boldsymbol{k_{24}} \\ \boldsymbol{k_{31}} & \boldsymbol{k_{32}} & \boldsymbol{k_{33}} + \widehat{\boldsymbol{K}}_{\boldsymbol{e}} & \boldsymbol{k_{34}} \\ \boldsymbol{k_{41}} & \boldsymbol{k_{42}} & \boldsymbol{k_{43}} & \boldsymbol{k_{44}} \end{matrix} \qquad \text{Equation 3}$$

$$\{F\} = [K_r\,(\Omega)]\{\delta_r\} \qquad \text{Equation 4}$$

Using Equation 2 - Equation 4, D-track assembles the track elements and adjacent rail span elements. Steffens [87] has introduced the dynamic analysis of rail track structures and an interface to D-track. However, D-track accuracy remained uncertain due to significant discrepancies between field and simulated data. Leong [88] conducted an in-depth study focusing on sleeper pad reactions, wheel-track interactions, and sleeper

bending assessments to enhance the D-track's performance. Validation through comparative analysis between site data from Melbourne to Geelong, Australia, and simulated data demonstrated a substantial improvement, achieving an error margin of less than 10%.

Table 5 tabulates the available parameters that can be adjusted on D-track. It is possible to export the outputs to a file in CSV format. If the output is the acceleration, the correlation between the acceleration and one of the tuneable parameters of interest can be established using AI.

Table 5 Tuneable parameters on D-track

Tuneable parameters				
Track	Vehicle	Irregularity	Variety	Output
Rail type Gauge Axial force Pad stiffness Pad damping Sleeper type Spacing Track bed stiffness Track bed damping	Speed Tare mass Carry mass Primary suspension stiffness Primary suspension damping Wheel dimension Hertzian contact coefficient	Corrugation Dipped joint Dipped weld Peaked weld Arbitrary profile Wheel flat	Speed Irregularity length Irregularity depth Centre of irregularity Rail analysis position Sleeper analysis position	Acceleration Force and pressure Moment and shear Bending moment Displacement

3.2.2 Field Data Acquisition

3.2.2.1 Calibration

Before deployment, each smartphone will be calibrated using a standardised vibration source to ensure measurement accuracy. Regular checks and recalibrations will be conducted throughout the data collection period to maintain data integrity. Blue tack is used to secure the iPhone for vibration measurements due to its practical advantages in field settings where rigid mounts are unavailable. Its adhesive properties keep the device stable on varied surfaces, minimising slippage during measurement. The app Phyphox is employed for data collection due to its high versatility and compatibility with smartphone sensors, offering precise control over data acquisition parameters and real-time data visualisation.

3.2.2.2 Collection

A robust data collection methodology conforming to ISO 2631-1 and ISO 2631-4 is employed to evaluate the RC and train vibrations systematically. This standard provides guidelines for assessing human exposure to WBV to ensure the rigour and validity of the collected data.

It is worth noting that passenger weight is not considered in this thesis. In assessing overall RC, the focus lies on the vibration characteristics of the entire train rather than the influence of individual passenger weight. The train’s overall vibration response reflects the combined dynamics of the train structure itself, not the variations introduced by individual occupants. For individual comfort assessments, seats are generally designed to provide consistent damping across different passenger weights, ensuring stability in

core vibration characteristics. Furthermore, standard comfort assessment frameworks typically do not account for passenger weight, as its variability adds complexity without substantially impacting generalised results. Consequently, excluding passenger weight enables a more consistent analysis.

Smartphones equipped with high-precision accelerometers are utilised as data collection devices. These devices are placed at multiple locations within the train (the floor and seating areas) to capture a comprehensive range of vibration data. Each smartphone is securely mounted with the X-axis of the phone aligning to the train running direction using blue tack to ensure consistent contact with the surface, thereby minimising extraneous movements that might compromise data accuracy.

Timestamping is used to synchronise data from multiple smartphones during the train journey, with each device recording data alongside precise local timestamps. This approach allows for accurate alignment of datasets during post-processing, even in conditions of intermittent network connectivity common on train routes, as the timestamps are generated locally and do not depend on continuous internet access. This method ensures that data from all devices can be accurately synchronised, facilitating robust analysis of the collective dataset. The data is collected at a high sampling rate of 100 Hz to capture the detailed dynamics of train vibrations. This high-frequency sampling ensures that low-frequency and high-frequency vibrations relevant to the RC are recorded. The recorded data includes the longitudinal (X-axis), lateral (Y-axis), and vertical (Z-axis) accelerations. These measurements are critical as they represent the multi-dimensional nature of train vibrations affecting passengers.

Data must be diverse enough to represent all the possible situations so that data can be collected across various operational scenarios to ensure a comprehensive analysis. These scenarios cover different train speeds, track conditions (e.g., straight tracks, curves, junctions), and operational states (e.g., acceleration, deceleration, steady-state running). By covering a wide range of conditions, the collected data can provide a thorough understanding of the vibrational environment experienced by passengers. The collected data is stored in a secure, centralised database, making it readily accessible for subsequent analysis. Metadata, including the time of measurement, start and destination of the journey, and any abnormalities, are logged alongside the vibration data to provide context for the analysis.

3.2.3 Data Integrity

Data integrity refers to data's accuracy, consistency, and trustworthiness throughout its lifecycle [88]. It ensures that data remains trustworthy and reflects its true form [89].

Assessing data integrity involves systematically evaluating several key indicators to ensure that data remains accurate, consistent, and reliable. Firstly, data validation is employed to verify that information adheres to expected formats, values, and ranges, ensuring initial accuracy. Consistency checks are conducted across systems and databases to confirm that identical data remains synchronised without discrepancies. Audit trails and change logs are scrutinised to track authorised modifications and detect unauthorised alterations. Furthermore, the completeness of the data is assessed to ensure no information is missing or truncated, while regular verification against original sources or trusted benchmarks confirms its ongoing accuracy. Finally, the timeliness and relevance of the data are considered, ensuring it remains applicable and up-to-date. Data

integrity can be effectively maintained and assessed by applying these rigorous measures.

In this thesis, data validation is ensured by calibrating smartphone sensors against known standards, guaranteeing that the collected data adheres to the expected formats. Consistency checks are performed by synchronising data from multiple devices, ensuring an identical timeframe is recorded without discrepancies across different smartphones used in the study. Logging of metadata such as timestamps and sensor details is used to provide an audit-like mechanism to track any potential modifications. Data completeness is addressed through extensive data collection to cover as many operational conditions as possible.

3.3 Data Processing Methods

Figure 3-4 illustrates data processing methods divided into three main phases: data pre-processing**,** data segmentation**,** and data fusion. In the data pre-processing stage, data cleaning, transformation, and reduction are employed to refine and optimise raw data. The data segmentation phase includes windowing techniques that partition data into meaningful subsets. Finally, data fusion combines features and applies techniques like GAT to enhance data integration for advanced analysis.

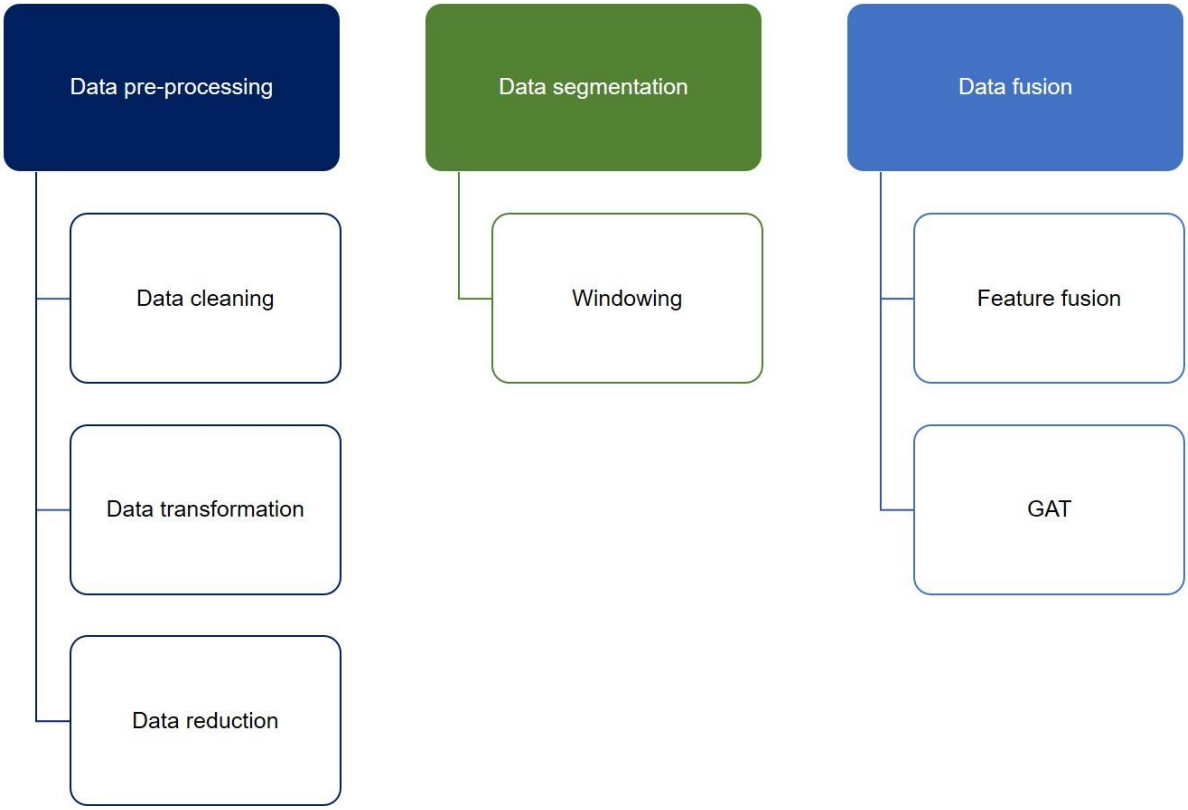


Figure 3-4 Overview of data processing

3.3.1 Data Pre-processing

Data pre-processing is a fundamental step in the ML pipeline. It plays an essential role in the quality and reliability of the data used for model training and evaluation. Sensory data is often incomplete, noisy, and inconsistent, which can significantly impair the performance of ML algorithms. Pre-processing techniques, such as data cleaning and normalisation, are helpful to mitigate these issues by addressing missing values, correcting inconsistencies, and removing outliers.

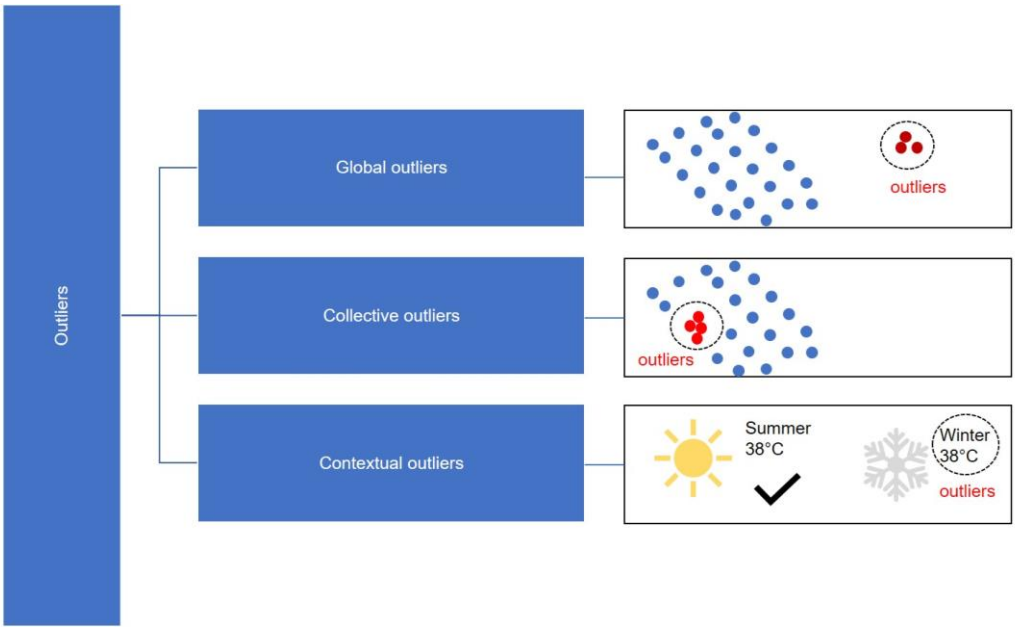
Noisy data can distort the learning process and lead to inaccurate models. Pre-processing techniques identify and denoise to ensure that the data fed into the ML algorithms is as clean and representative as possible. Vibrations are typically high-dimensional data, which can lead to the curse of dimensionality, where the performance

of ML models deteriorates due to the sparsity of data points in high-dimensional spaces. Techniques like feature extraction can improve the efficiency and effectiveness of the learning process by focusing on the most relevant attributes. Different ML algorithms require data to be in specific formats or scales. Data transformation processes convert raw data into formats suitable for analysis and algorithms, optimising the learning process.

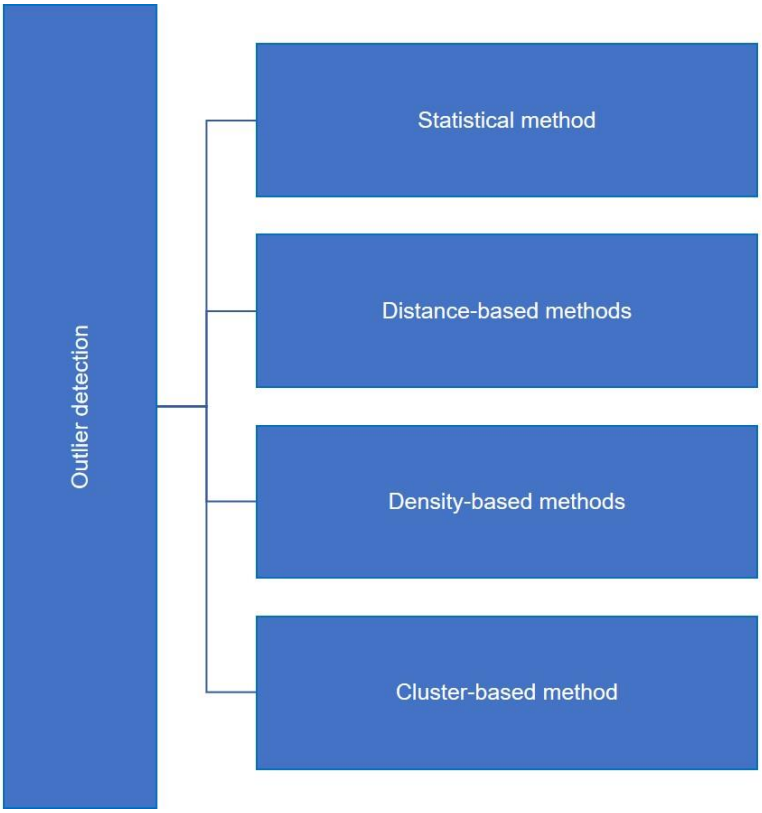
3.3.1.1 Data Cleaning

Outlier Detection

An outlier is a data point that significantly deviates from the rest of the data in a dataset, which can be caused by various factors such as measurement errors or sampling errors [90]. Figure 3-5(a) categorises outliers into three types [91]. Global outliers refer to data points exceptionally far from the entire dataset, while collective outliers are a group of data points that collectively deviate from the expected pattern, even though individual points might not seem anomalous. Contextual outliers are data points unusual only within a specific context, as seen in Figure 3-5(a), which considers a temperature of 38°C might be an outlier in winter but usual in summer.



(a)



(b)

Figure 3-5 Overview of outliers (a) outlier types (b) outlier detection methods

Outlier detection techniques [92] can be classified into four categories, as illustrated in Figure 3-5(b). Statistical methods rely on known data distribution models and work well when these models are clearly defined. However, they struggle to produce accurate results when the data distribution is unknown or complicated. Distance-based methods find outliers by measuring the distance between data points to identify those far from the rest. Although effective in many situations, these methods become slow and resource-intensive when used on high-dimensional data. Density-based methods focus on the local density of data points, identifying those with much lower densities as outliers. While useful for medium-sized datasets, they become inefficient for larger datasets. Finally, cluster-based methods group data points into clusters, marking those that do not fit into any cluster as outliers. These methods can handle various types of data but often rely on careful parameter settings and may struggle with data that has varying densities or many dimensions.

3.3.1.2 Data Transformation

Data discretisation and normalisation

Data discretisation and normalisation are crucial pre-processing techniques in ML that significantly enhance the quality and suitability of data for analysis. Discretisation involves transforming continuous variables into discrete intervals, simplifying the data, and making it easier for various algorithms to handle. The binariser method converts numerical features into binary ones. Similarly, the bucketiser divides continuous features into specified buckets. On the other hand, normalisation adjusts the distribution of data values to standardise scales across features. The normaliser technique ensures that each row has a consistent unit norm across data points. The “StandardScaler” adjusts features to

follow a normal distribution, while the “MinMaxScaler” confines feature values within a specified range using lower and upper bounds.

Fast Fourier Transform

Fast Fourier Transform (FFT) is an essential tool used in this thesis to reveal the in-depth pattern of vibrations. Setting up boundaries and filtering out specific train motions is possible with FFT. In general, FFT can reveal features that are not available in time domains. Using FFT, time-domain data can be divided into many periodic waves, as shown in Figure 3-6.

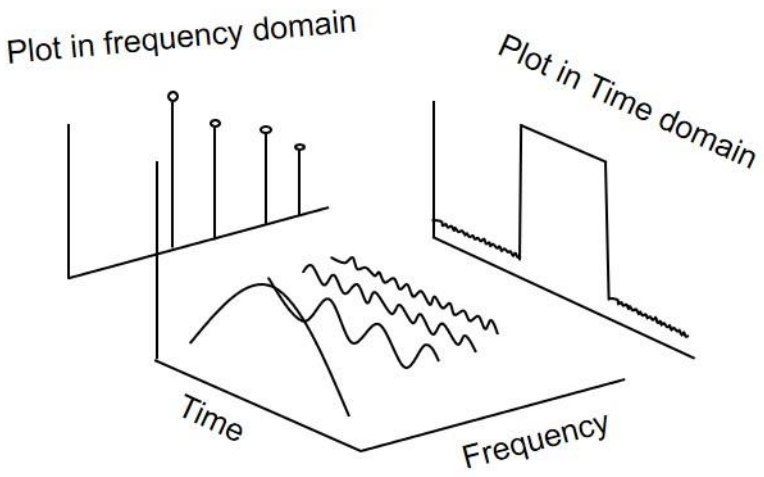


Figure 3-6 Fast Fourier Transform

The FFT is instrumental in analysing vibration data because it can efficiently convert time-domain signals into their frequency-domain counterparts. It is observable in Figure 3-6 that FFT allows for the identification and analysis of the individual frequency components present in a complex signal, which can detect the specific patterns or anomalies indicative of mechanical issues.

The application of FFT in vibration analysis is well-documented in academic literature. As a comparative study in [93], the FFT enhances the resolution of spectral data and significantly reduces the computational time compared to traditional Discrete Fourier Transform (DFT) methods. Moreover, FFT's capability to handle large datasets efficiently makes it indispensable in real-time monitoring systems as it is a powerful tool in dimensionality reduction in vector databases [94].

These attributes make FFT a robust tool for processing vibration data for early fault detection [91], demonstrating its effectiveness in diagnosing complex machinery faults through advanced signal processing techniques.

3.3.1.3 Data Reduction

Feature Extraction

Feature extraction constitutes a fundamental procedure in ML, wherein raw data is transformed into a set of variables, or features, that are more amenable to computational analysis. The rationale behind feature extraction is rooted in real-world data's multidimensional and often unstructured nature, which can impede the performance of learning algorithms if not adequately pre-processed. The process involves identifying and selecting features most indicative of the underlying patterns within the data, thereby improving the algorithm's ability to learn from the data. The efficacy of ML models is significantly dependent on the quality of the extracted features, as they serve as the primary inputs that the models use to generate predictive outcomes. Hence, rigorous feature extraction is essential for optimising model performance and ensuring the robustness of the analytical results.

Effective feature extraction reduces the data's dimensionality, alleviating the curse of dimensionality, preventing overfitting, and improving computational efficiency. It also plays a crucial role in enhancing the model's interpretability and ability to generalise across diverse datasets. The success of an ML model is linked to the quality of the features extracted, as these features encapsulate the most relevant and informative aspects of the data, thereby determining the model's predictive performance.

3.3.2 Data Segmentation

Data segmentation is a process of raw data divided into smaller, more manageable segments or chunks. This practice is essential when dealing with time-series data, as shown in Figure 3-7, where continuous data streams, such as smartphone accelerometer readings, are divided into fixed-size windows.

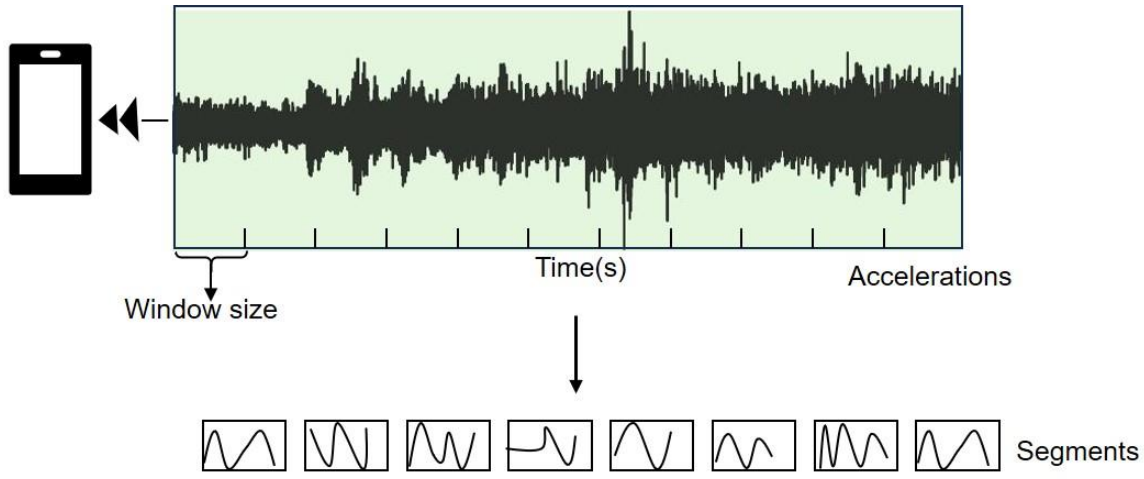


Figure 3-7 Data segmentation

The primary purpose of data segmentation is to transform raw, often unstructured, data into structured inputs that ML algorithms can effectively use. By breaking down the data into smaller segments, this process enables the extraction of meaningful patterns, facilitates the detection of significant events, and improves the overall accuracy of

predictive models. Without this step, the algorithms might struggle to find patterns with the noise of large, unsegmented datasets, leading to suboptimal performance and potentially erroneous predictions. Thus, data segmentation is not just a technical requirement but a foundational aspect of building reliable and interpretable ML systems.

3.3.3 Data Fusion

Data fusion is the process of integrating data from multiple sensors or sources to produce more accurate, reliable, and valuable information than any individual source [95]. This thesis mainly uses two types of data fusion techniques, feature fusion and GAT. Feature fusion combines extracted features from multiple techniques (time-domain and frequency domain here) to create a more robust and informative data representation [96].

GAT can be used to fuse crowd-sourced data from multiple smartphones. Each phone is treated as a node in a graph, with edges based on spatial relationships between phones. The GAT model assigns varying importance to each phone's data using an attention mechanism, enabling it to adapt to changes in the graph structure as passengers move. This allows the system to effectively combine data from different phones for accurate and dynamic comfort predictions across the train.

3.3.4 Comfort Level Calculation

3.3.4.1 ISO 2631

This thesis considers both ISO 2631-1 – general requirements for evaluating human exposure to WBV applicable to many conditions like transportation, workplace environments and general public exposure, and ISO 2631-4 – tailored explicitly to evaluating the effects of vibration and shock on passenger and crew comfort in fixed-

guideway transport systems like trains and trams. UIC 513, a variant of ISO 2631, is also employed in Chapter 6 to prove the proposed method's generality.

ISO 2631-1 focuses on weighted R.M.S acceleration for general measurement before the raw acceleration is weighted by weighting filters that account for frequency-dependent human sensitivity to vibration. Crest Factor (CF) is used to assess the severity of vibration peaks if the effect of vibration shock is unneglectable. Meanwhile, the VDV for evaluating the cumulative effects of vibration over time is introduced. It is worth mentioning that ISO 2361-4 adheres to the same calculation as ISO 2631-1 but with adjusted weighting filters applicable to trains or trams, as shown in Figure 3-8.

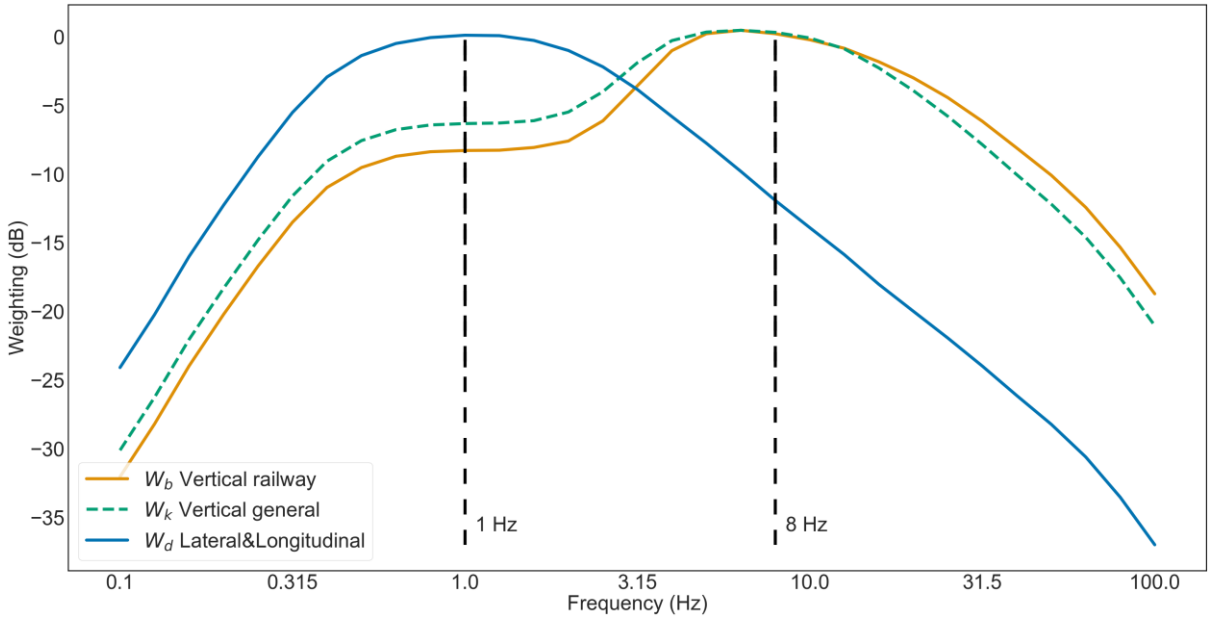


Figure 3-8 Frequency curves for W_d, W_k, and W_b

W_d is the frequency curve for the X- and Y- axes. It peaks at around 1 Hz, implying that vibrations have the most impact on the lateral and longitudinal directions at 1 Hz in the assessment of RC. It is noted that there are two frequency curves for vertical direction but for different uses. W_k is designed to estimate the effects of vertical vibrations on health

and comfort over a wide range of environments specified in ISO 2631-1, not limited to railways. Meanwhile, the W_b is preferred for railway applications as specified in ISO 2631-4, accounting for the unique vibration characteristics in the vertical direction specific to railway environments. The differences between W_k and W_b are notable below and above 8 Hz. Below 8 Hz, the W_b weighting curve yields a smaller weighting value by a factor of 0.8 compared to the W_k. In contrast, The W_b provide a more considerable weighting value by a factor of 1.2 compared to W_k.

Figure 3-9 highlights a complete methodology specified by ISO 2631-1 and ISO 2631-4 for calculating RC from vibrations. The assessment commences with vibration measurement and the application of frequency weighting filters before the R.M.S value is calculated in Equation 5.

$$a_w = [\frac{1}{T}\sum_{0}^{T} a_w^2\,(t)dt]^{\frac{1}{2}} \qquad \text{Equation 5}$$

Where $a_w(t)$ is the frequency-weighted accelerations at time t, with T denoting the measurement duration.

The process is then bifurcated according to the CF defined in Equation 6.

$$\mathrm{CF} = \frac{\max(\mathrm{a_w(t)})}{\mathrm{a_w^{r.m.s}}} \qquad \text{Equation 6}$$

Where $max(a_w(t))$ is the instantaneous peak value throughout the measurement and $a_w^{r.m.s}$ refers to the R.M.S value for the same duration.

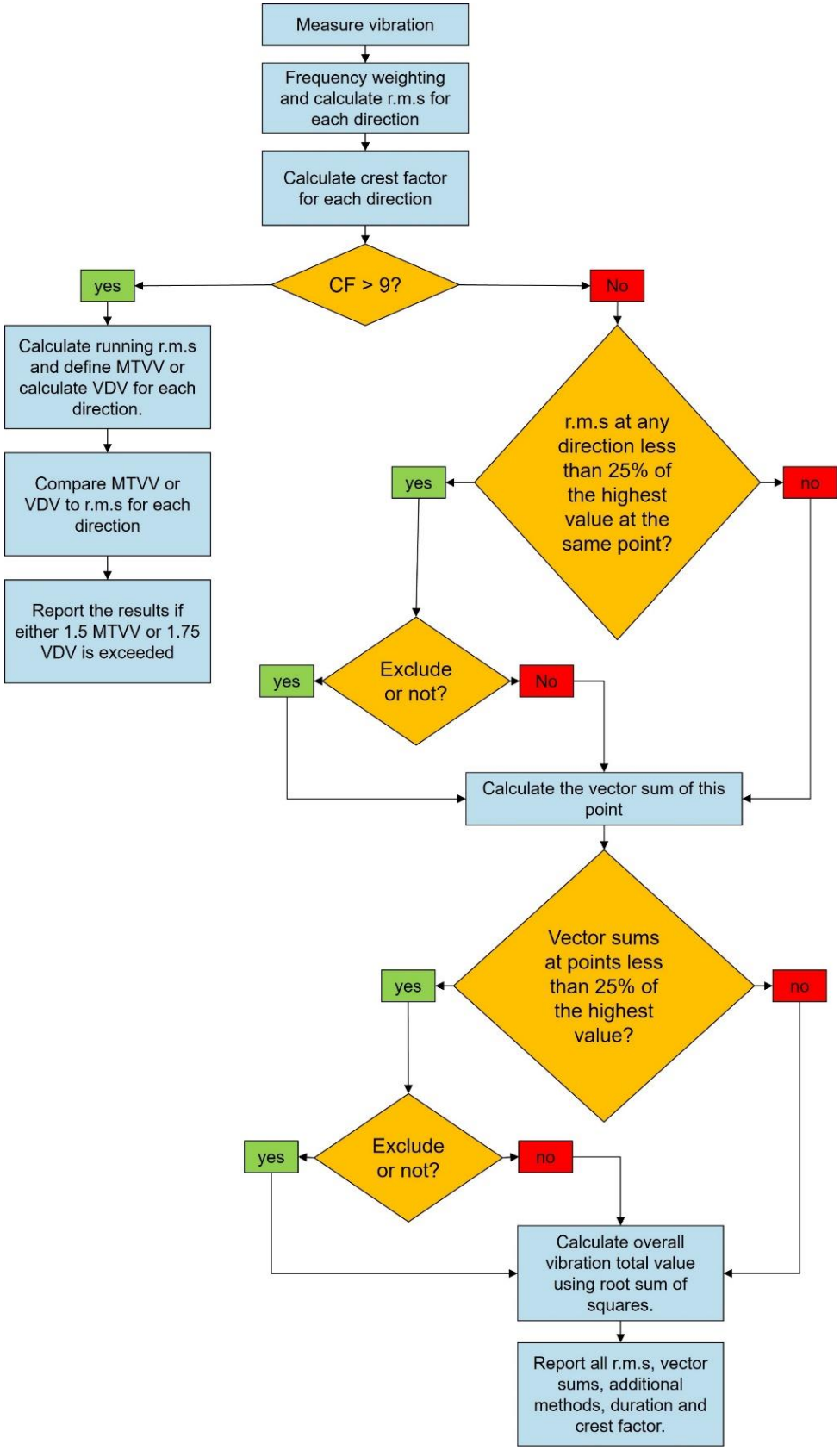


Figure 3-9 Ride index calculation using ISO 2631-1 and ISO 2631-4

If the CF is less than 9, the calculation remains at the R.M.S value and the overall vibration value is calculated in terms of different points. At each measuring point, if the R.M.S value in any direction is less than 25% of the highest R.M.S value at the same point, the less impacted value can be either excluded or included. The vector sum *V* can be calculated using the Root Sum of Square (RSS) method:

$$V = \sqrt{(r.m.s_x)^2 + (r.m.s_y)^2 + (r.m.s_z)^2} \qquad \text{Equation 7}$$

For the overall vibration value, the *V* at any point needs to be checked to see if it is less than 25% of the highest *V*, and then the overall value can be aggregated by Equation 7.

If the CF equals or exceeds 9, the VDV accounting for transient vibrational peaks or maximum transient vibration value (MTVV) is involved. The VDV is calculated below:

$$VDV = \left\{\int_0^T [a_w(t)]^4 \, dt\right\}^{\frac{1}{4}} \qquad \text{Equation 8}$$

The MTVV is defined as:

$$\mathrm{MTVV} = \max[\mathrm{a_w(t_0)}] \qquad \text{Equation 9}$$

Where the highest magnitude of $a_w(t_0)$ is read during the measurement period.

After the above procedures gain the vibration magnitude for comfort, Table 6 provides approximate indications of the likely reactions to various magnitudes of overall vibration values on a train.

Table 6 Comfort reactions to vibrations based on ISO 2631

Vibration (m/s^2)	**Comfort Scale Rating**
< 0.315	Not uncomfortable
0.315 to 0.63	A little uncomfortable
0.5 to 1	Fairly uncomfortable
0.8 to 1.6	Uncomfortable
1.25 to 2.5	Very uncomfortable
> 2.5	Extremely uncomfortable

3.3.4.2 UIC 513

UIC 513 inherits many principles from ISO 2631. Both standards share several similarities in their measurement configurations and frequency weighting functions. Figure 3-10 shows two types of measurement setups specified in UIC 513. The simplified measurement configuration includes three measuring points on the floor at the centre and both ends of the vehicle. In contrast, the full-scale measurement encompasses all 12 locations, as shown in Figure 3-10.

For the simplified measurement, parameters measured include accelerations in Y- and Z- directions in three locations with the X-axis vibration measured at the middle point since the three points can share the same vibration in the X-axis. Full-scale measurement is conducted at an additional 9 points at three seats.

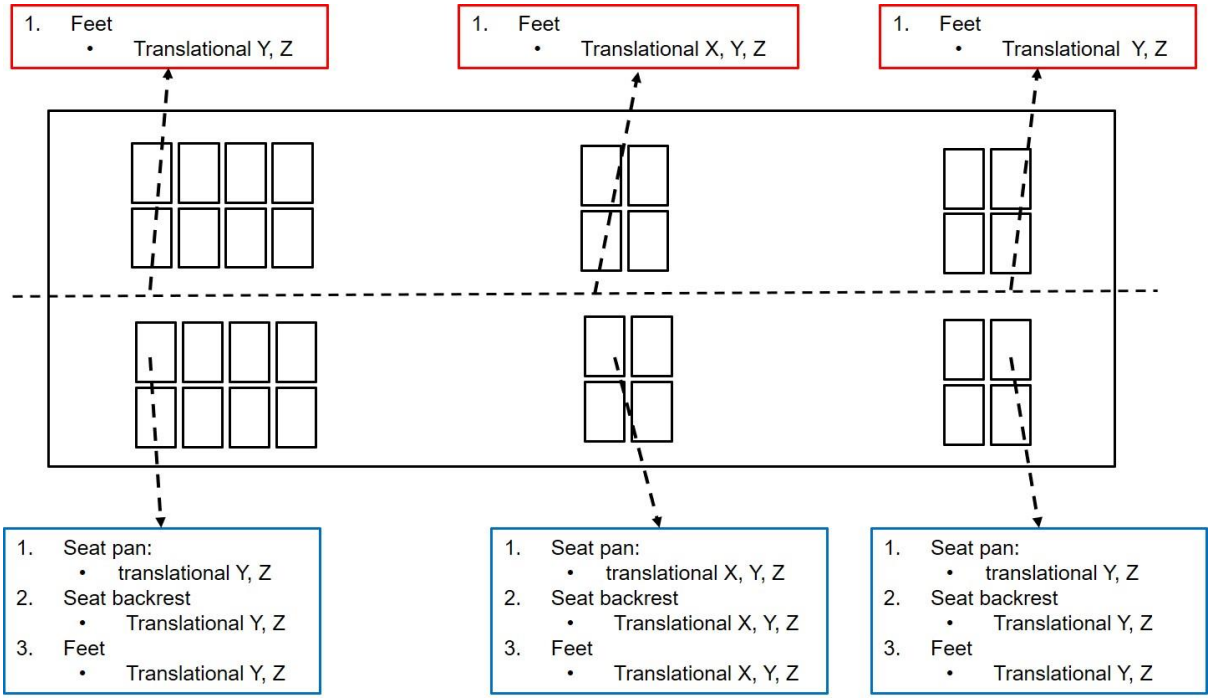


Figure 3-10 UIC 513 measuring points

The comfort level is determined by calculating the frequency-weighted R.M.S values of acceleration collected every 5 seconds over a five-minute duration. The comfort level indices are then derived from the following R.M.S values depending on the scale of measurement:

Simplified measurements:

$$N_{mv} = 6 \cdot \sqrt{(a_{XP95}^{w_d})^2 + (a_{YP95}^{w_d})^2 + (a_{ZP95}^{w_b})^2} \qquad \text{Equation 10}$$

Where:

- N_{mv} is the comfort level for both seated and standing passengers.
- *a* denotes accelerations.
- *X*, *Y*, and *Z* are the directions for the accelerations.
- *P* represents the data collection conducted on the vehicle floor.

- w_d is the frequency weighting function for the *X*- and *Y*- axes, and w_b is the frequency weighting function for the *Z*-axis.
- *95* is the confidence probability.

Full-scale measurement comfort calculation for seated passengers:

$$N_{VA} = 4 \cdot \left(a_{ZP95}^{w_b}\right) + 2 \cdot \sqrt{(a_{YA95}^{w_d})^2 + (a_{ZA95}^{w_b})^2} + 4 \cdot (a_{XD95}^{w_d}) \qquad \text{Equation 11}$$

Where:

- *A* is the seat pan.
- *D* is the seat backrest.

Full-scale measurement comfort calculation for standing passengers:

$$N_{VA} = 3 \cdot \sqrt{16 \cdot (a_{XP50}^{w_d})^2 + 4 \cdot (a_{YP50}^{w_d})^2 + (a_{ZP50}^{w_b})^2} + 5 \cdot (a_{YP95}^{w_d}) \qquad \text{Equation 12}$$

After the comfort index is determined, the comfort level can be referred to in Table 7, categorising comfort level into five levels, from very comfortable to very poor comfortable.

Table 7 Comfort index for UIC 513

Comfort level	**Interval**	**Description**
1	N < 1	Very comfortable
2	1 ≤ N < 2	Comfortable
3	2 ≤ N < 4	Moderate comfortable
4	4 ≤ N < 5	Poor comfortable
5	N > 5	Very poor comfortable

3.4 Machine Learning Models

In computer science, AI models mimic humans who can perceive their environments and execute actions to achieve some goals [97]. AI models leverage large datasets and mathematical frameworks to recognise patterns, make decisions, and predict outcomes. The importance and application of AI models have been proven in many areas, such as healthcare [98], where they aid in diagnostics and personalised medicine; finance [99], where they enhance fraud detection and algorithmic trading; and industry [100, 101], where they improve predictive maintenance and robotics.

3.4.1 Overview of Artificial Intelligence Model Types

AI, ML and DL are interconnected yet distinct concepts in computer science. It is fundamental to understand the definitions of them and the differences between them.

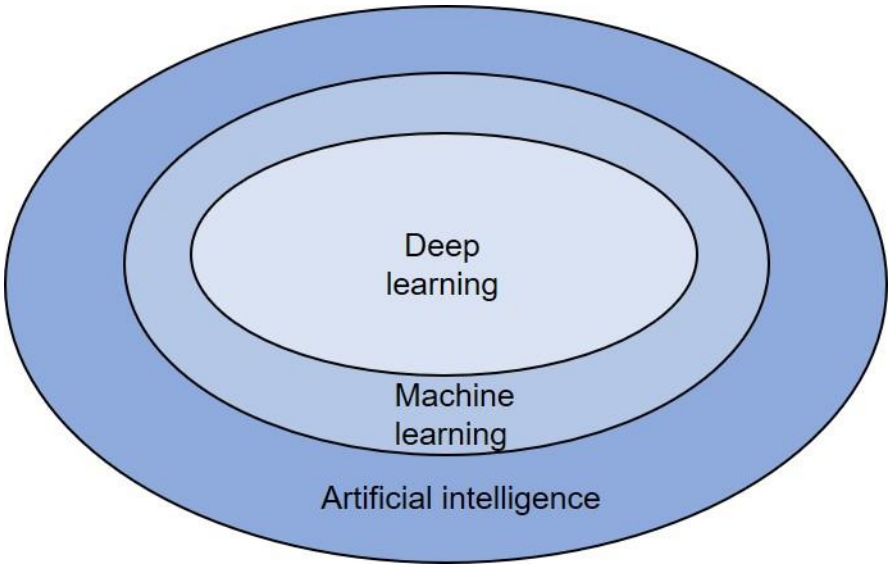
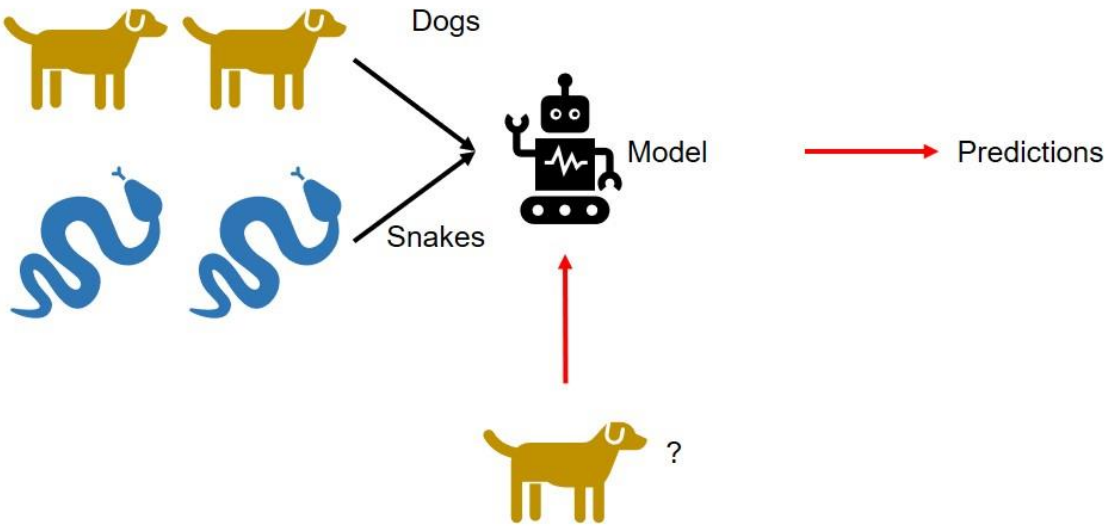


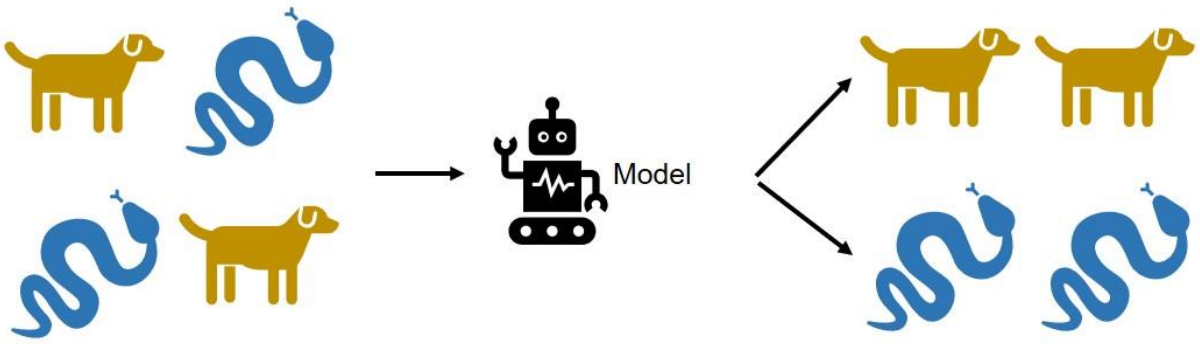
Figure 3-11 Difference between AI, DL and ML

As can be seen in Figure 3-11, ML is a subset of AI, while DL is a subset of ML. AI is a field that encompasses the development of systems or machines capable of performing

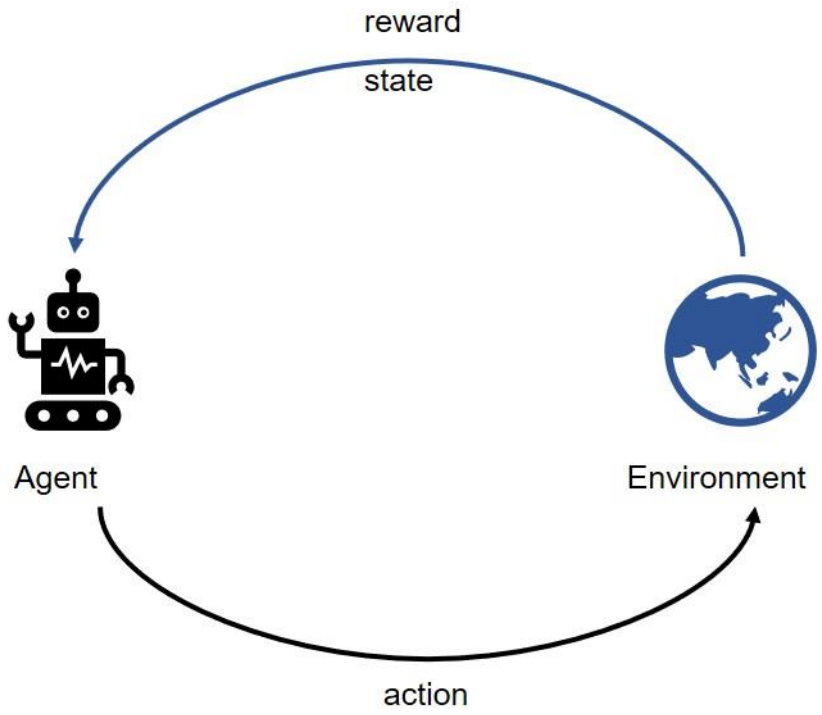
tasks that require human intelligence. At the same time, ML focuses on developing algorithms that allow machines to learn from data and predict using unseen data instead of explicit programming. Machines can learn from data in several ways: supervised, unsupervised, and reinforcement learning.



(a)



(b)



(c)

Figure 3-12 Machine learning (a) unsupervised learning (b) supervised learning (c) reinforcement learning

Figure 3-12(a) presents the mechanism for supervised learning [102]. Inputs and the corresponding labels are inputted into the model, which can map the input and label during the training process. After the training process, the model can make a prediction based on an unseen instance. Unsupervised learning [103] can find the hidden patterns or structures of the input data using unlabelled data, as shown in Figure 3-12(b). Figure 3-12(c) unfolds the reinforcement learning process [104], which learns to make decisions by receiving rewards or penalties for actions taken to maximise the cumulative reward.

3.4.2 Deep Learning Models

3.4.2.1 Neural Networks

Artificial neural network (ANN), as shown in Figure 3-13, were inspired by how the human brain processes information.

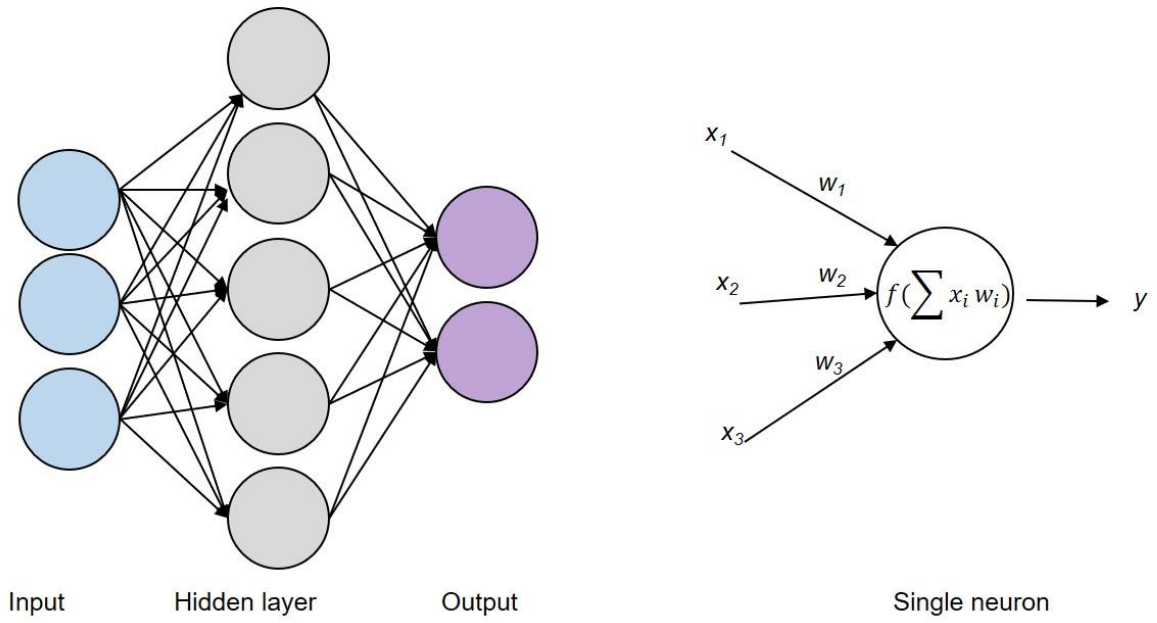


Figure 3-13 Artificial neural network

An ANN consists of interconnected groups of artificial neurons, which process information using a connectionist approach to computation [105, 106]. Each neuron in the network receives input, processes it through a series of weights, and generates an output based on a transfer function. The fundamental operation of an artificial neuron involves the calculation of a weighted sum of inputs, followed by applying a non-linear activation function to produce the neuron's output, as seen in Figure 3-13. This output can either be a binary value or a continuous value between 0 and 1, depending on the nature of the activation function used. The network learns and adapts by adjusting the weights of the connections through a process called training, which typically involves minimising the

error between the actual output and the desired target output using algorithms like back-propagation. This iterative adjustment process is similar to how synapses in a brain strengthen or weaken, allowing the network to learn complex patterns and make predictions or classifications based on new information.

Deep neural networks have a deep architecture consisting of multiple hidden layers (typically more than three) between the input and output layers. These additional layers enable the network to learn hierarchical representations of data.

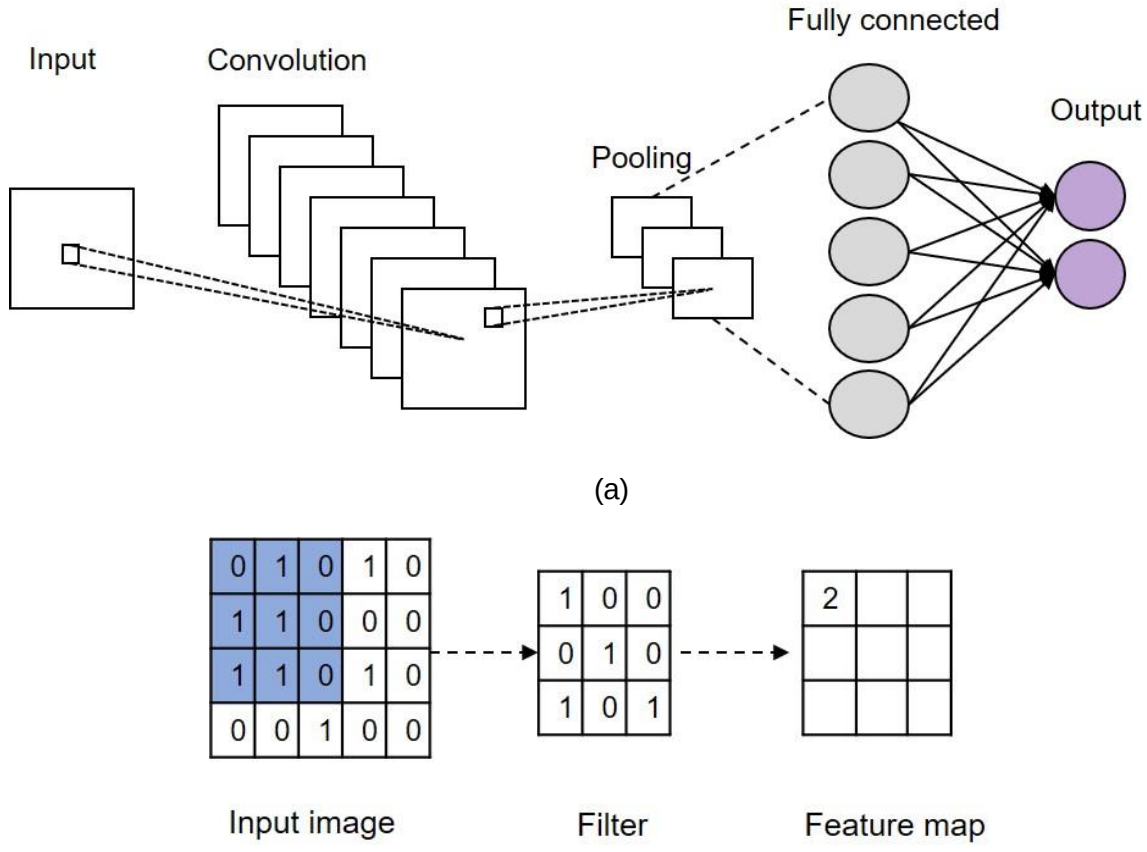


Figure 3-14 Convolutional neural network (a) architecture (b) convolutional function

A CNN is a specialised type of ANN that can process data with grid-like topology, such as images. CNN is structured to take advantage of the spatial relationships between

pixels using three main types of layers: convolutional layers, pooling layers, and fully connected layers, as shown in Figure 3-14(a). The convolutional layers apply various filters to the input image, as depicted in Figure 3-14(b), to generate activation maps highlighting specific feature maps like edges or textures. Pooling layers reduce the spatial dimensions of these activation maps to improve computational efficiency and reduce overfitting. Fully connected layers then interpret these features to make the final prediction. The hierarchical and deep nature of CNN allows them to learn complex patterns through multiple layers, where each layer captures increasingly abstract features of the input data.

3.4.3 Model Evaluation and Validation Techniques

3.4.3.1 Metrics

Metrics used to estimate ML performance depend on the problem type. For classification problems, accuracy, precision, recall, F1 score, and confusion matrix are widely considered. In this thesis, the prediction of the RC index is treated as a regression problem.

In the evaluation of regression analysis, the coefficient of determination, commonly known as R-squared (R^2) [107], stands out as a more informative metric compared to other commonly used methods such as the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE). This 0 to 1 nature of R^2 allows it to effectively communicate the proportion of variance explained by the model so that a clearer picture of the model's predictive power. R^2 is less affected by extreme values and thus provides a more stable performance measurement. To complement the performance measurement, the visualisation between the actual and predicted values is also provided as a

straightforward way to see the performance at the sample level and avoid any potential bias of the metrics. Therefore, due to R^2 robustness, ease of interpretation, and the straightforwardness of the contrast plot between the actual and predicted values, the plot and R^2 are recommended as the standard protocol for assessing regression analyses in this thesis.

Given a dataset with labels (y_1, …, y_n) and predicted values (y'_1, …, y'_n,), the R^2 is defined as follows:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad \text{Equation 13}$$

$$SS_{res} = \sum_i (y_i - y'_i)^2 \quad \text{Equation 14}$$

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \quad \text{Equation 15}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{Equation 16}$$

Where

- SS_{res} is the sum of squares of residuals between the actual value and the predicted value.
- SS_{tot} is the total sum of squares calculated by the mean value of the actual label and the label.

As can be seen from Equation 13, a small SS_{res} results from a small gap between the actual and predicted values lead to an R^2 value close to 1. However, there is an extreme

case that $\frac{SS_{res}}{SS_{tot}}$ is larger than 1 making a negative R^2 value. This means that the model's performance is worse than that of a model that does nothing but give the actual label's mean value every time.

While metrics like the R^2 provide valuable insights into the proportion of variance explained by the models, MAE and RMSE offer additional perspectives on model performance that are crucial for understanding the practical implications of prediction errors. MAE provides a transparent and interpretable measure of the typical magnitude of errors, directly reflecting the model's precision, by averaging the absolute differences between predicted and actual values. RMSE is particularly useful for identifying cases where the model's performance may be significantly impacted by outliers, which penalises more significant errors due to its quadratic nature. Therefore, some chapters also use MAE or RMSE to provide more insights. MAE and RMSE are defined as follows:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - y_i'|}{n} \qquad \text{Equation 17}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \|y_i - y_i'\|^2}{n}} \qquad \text{Equation 18}$$

For the cluster problems, the Silhouette score and Davies-Bouldin index (DBI) are used. The Silhouette score and DBI are metrics for evaluating the performance of clustering algorithms. The Silhouette score measures how similar an object is to its cluster compared to other clusters, ranging from -1 to 1 as defined:

$$s = \frac{b - a}{\max(a, b)} \qquad \text{Equation 19}$$

Where *a* is the distance between sample 'x' and other samples within the same group while *b* is the average distance between sample 'x' and all samples in the closest cluster, which is b_0 in Figure 3-15.

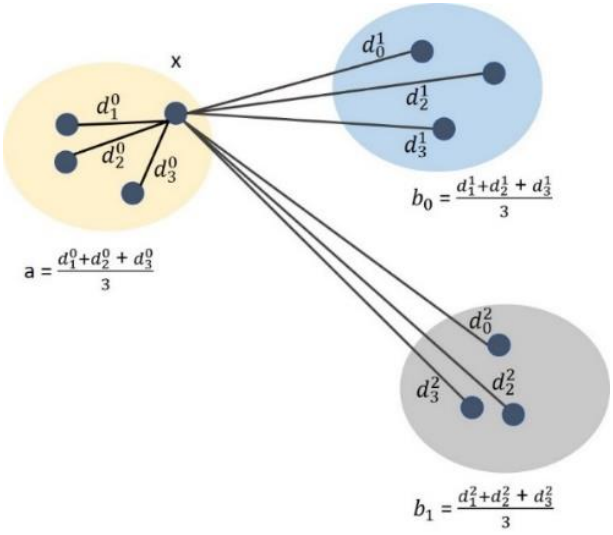


Figure 3-15 Silhouette coefficient

A higher Silhouette score indicates well-separated and cohesive clusters, making it a helpful tool for determining the optimal number of clusters and validating the clustering quality [108].

The DBI introduced by [109] measured clustering performance considering characteristics within each cluster and inter-cluster. The DBI evaluates the average similarity ratio of each cluster with its most similar cluster, where lower values signify better clustering. The metric is defined as:

Defining the average distance within a cluster.

$$d_i = \frac{1}{n_i}\sum_{k=1}^{n_i}\|x_k - C_i\| \qquad \text{Equation 20}$$

Where d_i is the average distance within cluster *i*; n_i is the number of samples in cluster *i*. x_k is the *k* sample while C_i is the centroid of cluster *i*.

The distance between the two groups' centroids is defined by:

$$M_{ij} = \|C_i - C_j\| \quad \text{Equation 21}$$

Where C_j is the centroid of *j* cluster.

For clusters *i* and *j*, the similarity R_{ij} is defined by:

$$R_{ij} = \frac{d_i + d_j}{M_{ij}} \quad \text{Equation 22}$$

Finally, the DBI is computed as:

$$DBI = \frac{1}{N}\sum_{i=1}^{N} \max_{j \neq i} R_{ij} \quad \text{Equation 23}$$

Where *N* represent the number of groups.

3.4.3.2 Cross-validation

K-fold Cross Validation (KCV) is one of the most prevalent methods in ML for evaluating the model's reliability and robustness. KCV is a resampling technique used to evaluate the performance of a model by dividing the original data set into k equally sized folds or subsets [110]. Each fold acts as a testing set at some point, while the remaining k-1 folds are used as the training set. This process is repeated k times, with each fold being used exactly once as the testing data. The k results from the folds can then be averaged to produce a single estimation schematically explained in Figure 3-16 using an example of 10-fold cross-valuation.

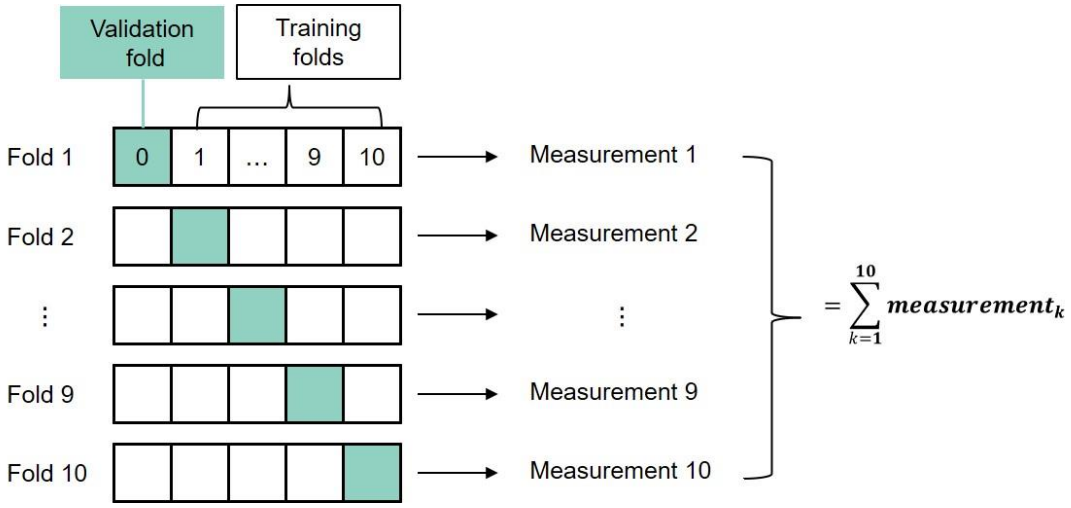


Figure 3-16 K-fold cross-validation

KCV mitigates overfitting by validating the model on different subsets of data, thereby providing a more accurate reflection of its performance on unseen data. When the data is scarce, KCV optimises the available data, allowing each instance to be used for validation exactly once and for training k−1 times. This approach improves the stability and reliability of performance metrics by averaging results across k iterations to make the estimates less sensitive to the specific subset. Additionally, the choice of k impacts the bias-variance trade-off, with higher k values reducing bias but increasing variance, whereas lower k values increase bias and reduce variance. K is commonly chosen as five or ten to balance these extremes effectively and ensure robust performance estimates.

Chapter 4

Quantification of Dynamic Track Stiffness Using Machine Learning

4.1 Introduction to Track Stiffness Quantification

Railway track stiffness, which measures track deflection in response to wheel load, is vital for design and maintenance. High track stiffness typically supports larger loads and reduces track deflection but accelerates wear and fatigue in other track components like rails, sleepers, and ballast. To enhance railway competitiveness and appeal, achieving higher running speeds and reducing maintenance costs is essential, as evidenced by the Madrid-Barcelona high-speed line, which operates at over 300 km/h in some sections, and the UK's ongoing high-speed railway project. From a design perspective, there is a significant interest in optimising track stiffness to minimise track deterioration and maintenance costs, as explored in [111]. Regarding maintenance, track stiffness is a crucial indicator for diagnosing track geometry issues [112]. At the same time, optimising

the track stiffness also contributes to a good RC, as the track conditions impact the vibration significantly. Thus, the chapter aims to preliminary test the feasibility of the vibration feature and usefulness of the ML model in the railway sector with a synthesised dataset from D-track.

Over the past three decades, numerous studies have highlighted the correlation between track stiffness and conditions [113, 114]. Wang et al. have provided a comprehensive summary of track stiffness interpretation, focusing on low track stiffness, varying stiffness in transition zones, virtual track stiffness, and assortative stiffness [115].

Track stiffness measurement typically refers to the global track stiffness, encompassing all layers of the track structure [116]. Component-level stiffness, such as ballast and fasteners, can be assessed in laboratory conditions [117]. The methods for measuring track stiffness have significantly evolved over time. Various systematic reviews have been conducted on these methods, such as rolling stiffness measurement vehicles [118], the hammer method [119, 120], portancemeter [121], and FWD (Falling Weight Deflectometer) [122]. Numerous parametric studies have been carried out to evaluate the performance of these measurements, focusing on factors like obstructiveness and measurement speed. Most of these methods operate at relatively low speeds, with the China Academy of Railway Sciences (CARS) achieving the highest measurement speed of 60 km/h [123]. This low measurement speed can impose additional burdens on track capacity.

To address the need for faster track stiffness measurement, a new solution that provides continuous, wide-coverage, non-intrusive measurements at realistic speeds and loads is required. Advances in sensor and wireless communication technologies have made it

possible to deploy such solutions using in-service trains. Recent research has utilised vehicle-track dynamic interaction and cross-entropy optimisation to determine track stiffness [124]. Numerical validation was performed using a half-bogie model and a beam-on-elastic-foundation track without irregularities. However, it is important to consider various vehicle types, track conditions, and track irregularities to enhance the robustness of this framework. Mehrali et al. proposed a measuring vehicle equipped with cameras and lasers to determine track stiffness at speeds up to 120 km/h [125]. While this high speed allows for deployment on in-service trains, the complexity of using cameras and lasers could be reduced for more practical applications.

To address the issues specified above, this section employs dilated convolutions to evaluate global track stiffness using train axle box vibrations. This technique is particularly suitable for handling noisy and large sensory data. Dilated convolutions allow the model to capture patterns from a broader field, reducing the impact of noise and detail misinterpretation with the additional benefit of being computationally cost-friendly. This section uses the validated D-track simulation to generate vibration inputs for models. The core contributions of this framework are:

- CNN Architecture: a dilated convolutional layer is integrated to process vibration data. This layer provides a broad view of the input without additional parameters. A comparative analysis of traditional CNN and dilated CNN is presented in terms of performance and computational cost.
- Robust performance: the proposed model shows robust performance with realistic vibration data, including track irregularities. The synthesised dataset includes three rail irregularities (dipped joint, dipped weld, and corrugation) under different

track stiffness conditions (100 MN/m, 600 MN/m, and 800 MN/m) to test the model's behaviour.

- The typical range of track stiffness on the UK railway network is between 45 to 80 MN/m per sleeper end, balancing load support and flexibility to maintain track stability and passenger RC. However, in simulations, higher values—such as 600 MN/m or 800 MN/m—are used to diversify the dataset, allowing the model to train on a wider range of conditions. This approach improves the model's ability to generalise, preparing it for both standard and extreme scenarios. Additionally, using high stiffness values allows for the testing of extreme conditions, revealing potential stress points, track wear, and structural responses that may arise under unusual track configurations or with heavy rail traffic.

4.2 Methodologies

4.2.1 D-track

D-track allows for extensive parameter tuning from track properties (global track damping and stiffness) and vehicle characteristics (models, speed, weight, and wheel radius) to irregularities (dipped joint, corrugation, etc.). Figure 4-1 illustrates the primary procedure for generating axle box accelerations using D-track, and Table 8 details the three key inputs adjusted in the simulation.

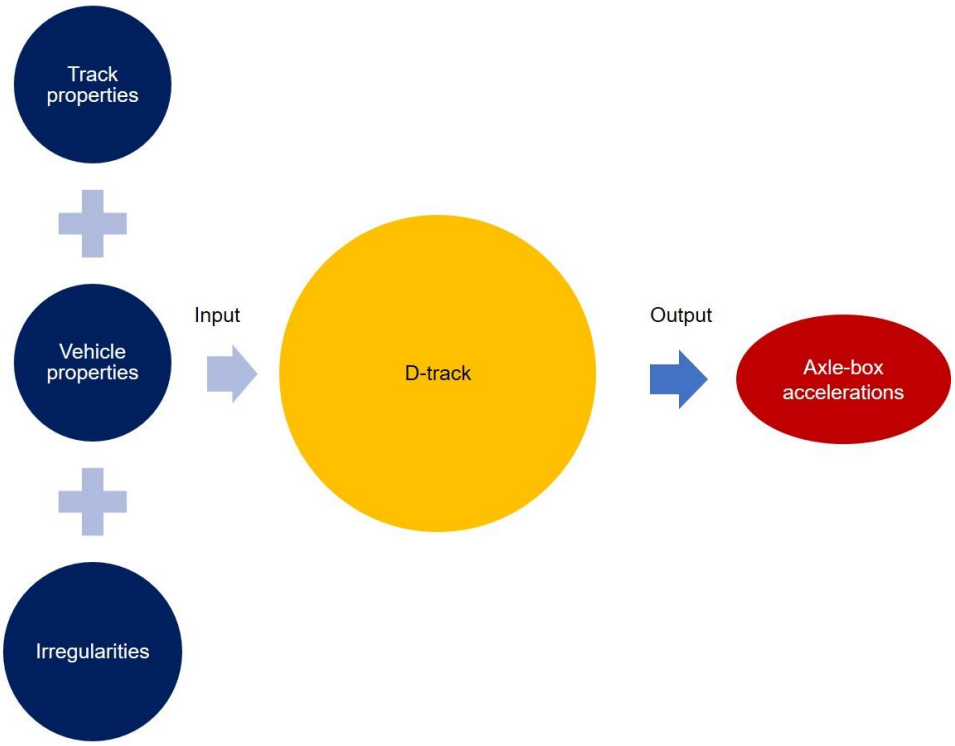


Figure 4-1 The process of data simulation

The input parameters are categorised into three main groups: track properties settings, vehicle properties settings, and track irregularities. These inputs are input to the D-track, which outputs the axle box accelerations.

Table 8 Parameters tuned in D-track

Parameters	**Value**
Global Stiffness	0 - 900 MN/m
Global Damping	0 - 500 kN/m
Rail Type	AS60
Gauge	Standard (1435mm)
Pad/Plate Type	HDPE
Sleeper Type	Prestressed
Vehicle Speed	60-120 km/h

Vehicle Type	106t Coal Wagon
Bogie Type	QR56
Wheel Type	QR
Centre of Irregularity	Midspan Before Sleeper
Rail Analysis Position	Above Sleeper
Sleeper Analysis Position	Rail Seat
Track Modulus	12.8-41.4 m/s^2
Track irregularities	1000 mm length
	1-10 mm depth

In the simulations, the vehicle speed ranges from 60 to 120 km/h. Three types of track irregularities—corrugation, dipped weld, and dipped joint—are incorporated to enhance realism. Prestressed sleepers are considered due to their longer life cycle and lower maintenance cost than reinforced concrete sleepers [126]. Track stiffness varies from 0 to 900 MN/m, increasing in 100 MN/m increments.

4.2.2 Application of Dilated Convolutional Neural Network

CNN has emerged as a ground-breaking architecture in image recognition and pattern analysis. In chapter 3.4.2.1, basic information on CNN is provided. However, that is not enough to understand how CNN works. Here, it delves into the details of CNN, exploring their architecture, the functionality of individual layers, and the methods that contribute to their robustness and efficiency. Mathematical equations and models are included to provide a comprehensive understanding of how these networks function, making

significant contributions to various fields, including computer vision, medical image analysis, and autonomous systems.

The CNN architecture is composed of multiple layers. The three primary types of layers in a CNN are convolutional layers, pooling layers, and fully connected layers (can be seen in Figure 3-14(a)). These layers are stacked sequentially to form the complete architecture of the network, which can gradually transform the input data into an easily classified form. The convolutional layer is the most important part of a CNN, where the main processing of the input data happens. This layer utilises learnable filters, or kernels convolved with the input data, to produce activation maps. Each activation map represents the presence of specific features within the input data. The calculation of a convolutional layer can be mathematically defined as:

$$A_{i,j,k} = \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{d=1}^{D} I_{i+m-1,j+n-1,d} \cdot K_{m,n,d,k} + b_k \qquad \text{Equation 24}$$

Where:

- $A_{i,j,k}$ is the activation at position (*i*, *j*) in the *kth* feature map.
- *I* is the input image or feature map with dimensions *H * W * D.*
- *K* is the size of a convolutional kernel *M * N * D.*
- *M* and *N* are the height and width of the kernel.
- *D* is the depth of the input matching the depth of the kernel.
- b_k is the bias term associated with the *kth* filter.

The calculation performed by Equation 24 is through the whole input, extracting many feature maps that store different information in the input data. One of the critical

advantages of convolutional layers is their ability to share parameters, significantly reducing the number of learnable parameters in the network. This parameter sharing assumes that features detected at one spatial location in the input will likely be helpful in other locations. This not only reduces the computational complexity of the network but also helps prevent overfitting. The convolution operation computes a weighted sum of the input values within a local region (determined by the kernel size) and outputs this sum as a single value in the feature map. The feature map size depends on the input dimensions and the convolution operation's hyperparameters, such as stride and padding. The spatial dimensions of the output feature map can be calculated using the following formula:

$$O = \frac{H - M + 2P}{S} + 1 \qquad \text{Equation 25}$$

where:

- O is the size of the output feature map.
- P is the amount of zero padding.
- S is the stride of the convolution.

Stride refers to the step size with which the convolution filter moves across the input data. A larger stride reduces the spatial dimensions of the output feature map, while a stride of 1 maintains a higher resolution.

Padding (as depicted in Figure 4-2) involves adding zeros around the input image's border to control the output's spatial dimensions. The padding ensures that the output feature map retains the same spatial dimensions as the input, which is especially useful when stacking multiple convolutional layers.

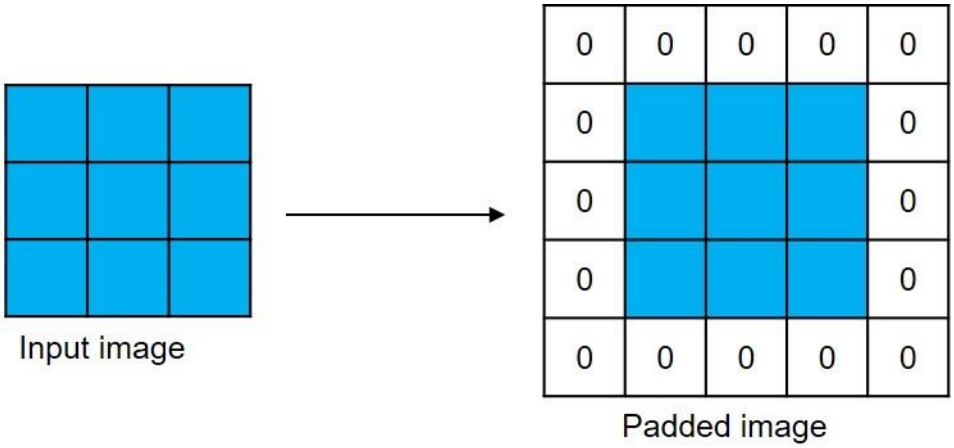


Figure 4-2 Padding

The pooling layer (as observed in Figure 4-3), also known as the subsampling or down-sampling layer, is responsible for reducing the spatial dimensions of the activation maps generated by the convolutional layers.

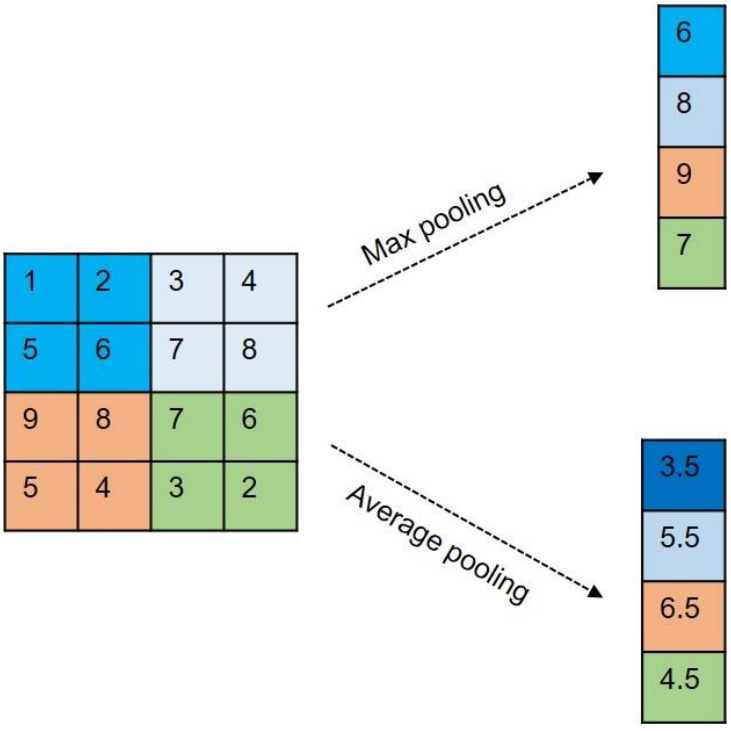


Figure 4-3 Pooling

This dimensionality reduction decreases the computational load and prevents overfitting by summarising the presence of features in a given region of the input data.

The most commonly used pooling operation is max-pooling, mathematically defined as:

$$P_{i,j,k} = max_{m,n}\,(A_{i+m,j+n,k}) \qquad \text{Equation 26}$$

Where:

- $P_{i,j,k}$ is the pooled output at position (*i*, *j*) in the *kth* feature map.
- The *max* operation is applied on a region of the activation map *A* defined by the size *m* and *n* pooling window, typically of size 2 * 2 with a stride of 2.

In addition to max-pooling, average pooling is another common approach where the average value of the elements in the pooling window is computed. This method can be beneficial in tasks where the exact value of the features is more important than their presence. The average pooling operation is defined as:

$$P_{i,j,k} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} A_{i+m-1,j+n-1,k} \qquad \text{Equation 27}$$

The fully connected layer is the final layer in a CNN and is analogous to the layers found in traditional ANN. In this layer, each neuron is connected to every neuron in the previous layer, allowing for the integration of all the features extracted by the convolutional and pooling layers. If a CNN is developed for a regression problem, the number of neurons in this layer can be set to 1 (in the case of one output) to get the final output. If it is for a classification problem, the output of the fully connected layer is typically fed into a

SoftMax function, which produces a probability distribution over the possible output classes:

$$\hat{y}_k = \frac{\exp(z_k)}{\sum_{j=1}^{C} \exp(z_j)} \quad \text{Equation 28}$$

Where:

- $\hat{y}_k$ is the predicted probability of class *k*.
- z_k is the output of the *kth* neuron in the fully connected layer.
- *C* is the number of output classes.

The fully connected layer aims to combine the features extracted by the previous layers and produce the final decision. The weights in this layer are learned through backpropagation, where the error between the predicted output and the actual label is minimised.

Training a CNN involves adjusting the network's parameters (weights and biases) to minimise errors between the predicted output and the actual label. This is typically achieved using gradient-based optimization techniques, such as Stochastic Gradient Descent (SGD), defined as follows:

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta J(\theta) \quad \text{Equation 29}$$

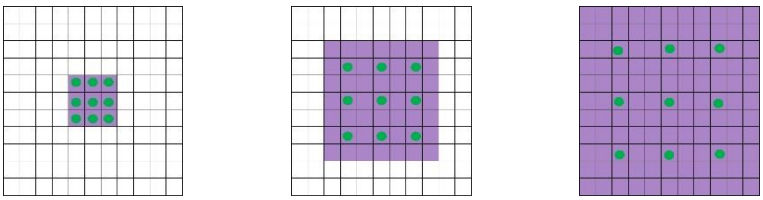
where:

- θ represents the network's parameters (weights and biases).
- η is the learning rate, a hyperparameter controlling the step size of each update.

- $J(\theta)$ is the cost function, often the cross-entropy loss in classification tasks and mean squared error for regression problems.
- $\nabla\theta J(\theta)$ is the gradient of the cost function for the parameters.

Dilated convolution, also known as atrous convolution, effectively enlarges the filter's receptive field without increasing computation or parameters. This technique is widely used in semantic image segmentation [127]. It expands the traditional convolutional layer by inserting voids between adjacent elements. Beyond image segmentation, dilated convolution has been applied to various domains, including speech emotion recognition [128] and image classification [129].

In Figure 4-4(a)-(c), the kernel size remains 3 × 3 (green dots), but the receptive field varies with the dilation rate. A 1-dilated convolution maintains a 3 × 3 receptive field that matches the kernel size. Inserting one zero hole between elements enlarges the receptive field to 7 × 7, as shown in Figure 4-4(b). A 3-dilated convolution further expands the receptive field to 11 × 11, as illustrated in Figure 4-4(c). This demonstrates how the receptive field grows exponentially with the dilation rate.



(a) (b) (c)

Figure 4-4 The expanded receptive field of a 3×3 kernel using different dilated rate (a) dilated rate = 1 (traditional CNN), (b) dilated rate = 2, (c) dilated rate = 3

4.3 Results and Discussions

4.3.1 Dataset Overview

For the dataset, 401 simulations are conducted, leading to 401 Excel files storing the vibrations. Before splitting the dataset, the raw data is segmented into samples of 250, 500, or 1,000-time steps. The dataset is then divided into 85% for training and 15% for testing. From the training set, 20% is set aside as the validation set to ensure unbiased model selection, keeping the test set untouched during training and selection. Subsequently, min-max normalisation is applied, as its effectiveness has been demonstrated by [130]. This normalisation uses the maximum and minimum values from the training set rather than the entire dataset to prevent information leakage from the test set. The summary of the dataset responding to different window sizes is tabulated in Table 9. Larger window sizes can lead to a prediction of track stiffness spanning a longer distance.

Table 9 The summary of the dataset

Window size	The number of samples
250	1604
500	802
1,000	401

Figure 4-5 illustrates the vertical acceleration for a track stiffness of 300 MN/m at 60 km/h and 80 km/h speeds. The figure highlights two different vibration patterns resulting from the different speeds but corresponding to the same track stiffness.

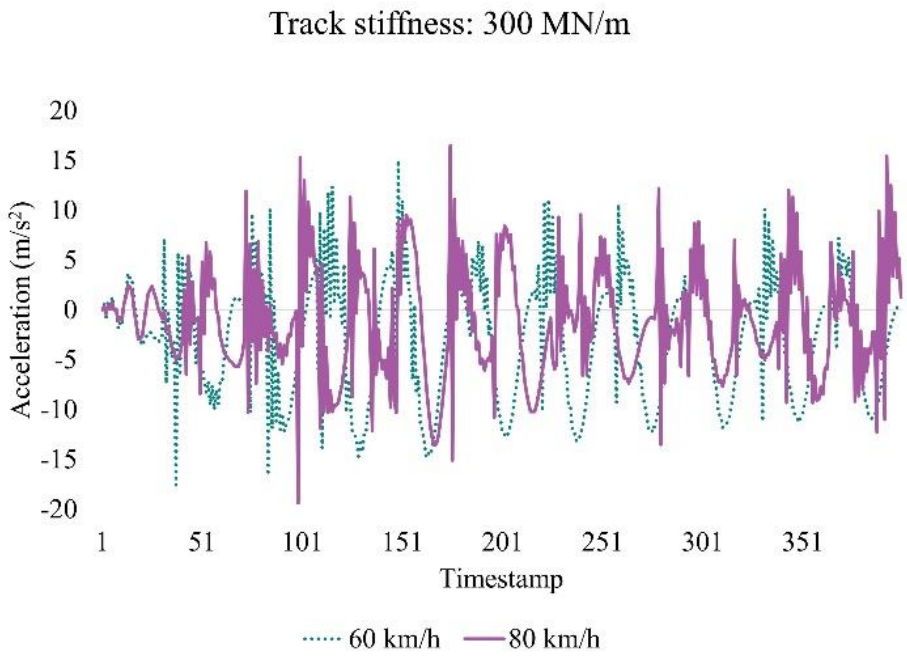


Figure 4-5 A sample of the raw data

4.3.2 Performance of Various Models

This section presents the optimal results for both traditional CNN and dilated CNN models, evaluated using test sets from three different datasets based on sample size. Model performance is assessed using MAE and the R^2. Figure 4-6 summarises the performance and computation time for both models.

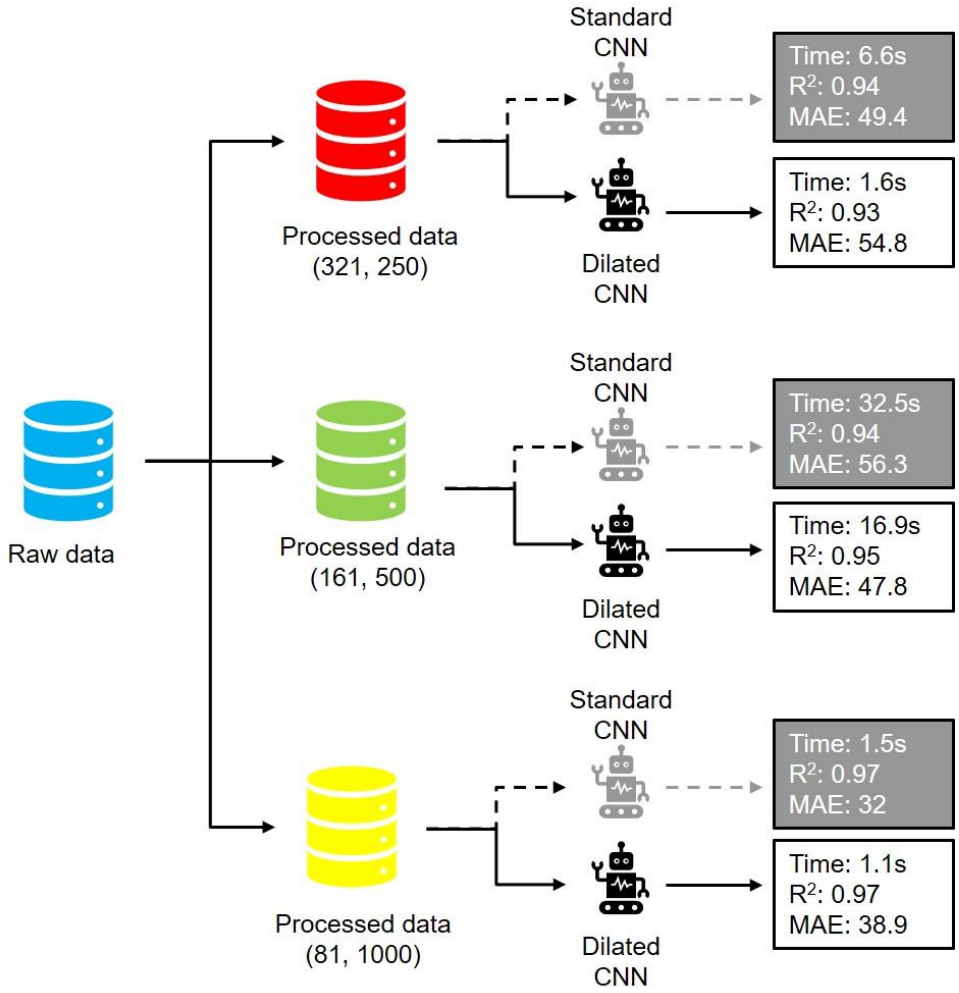
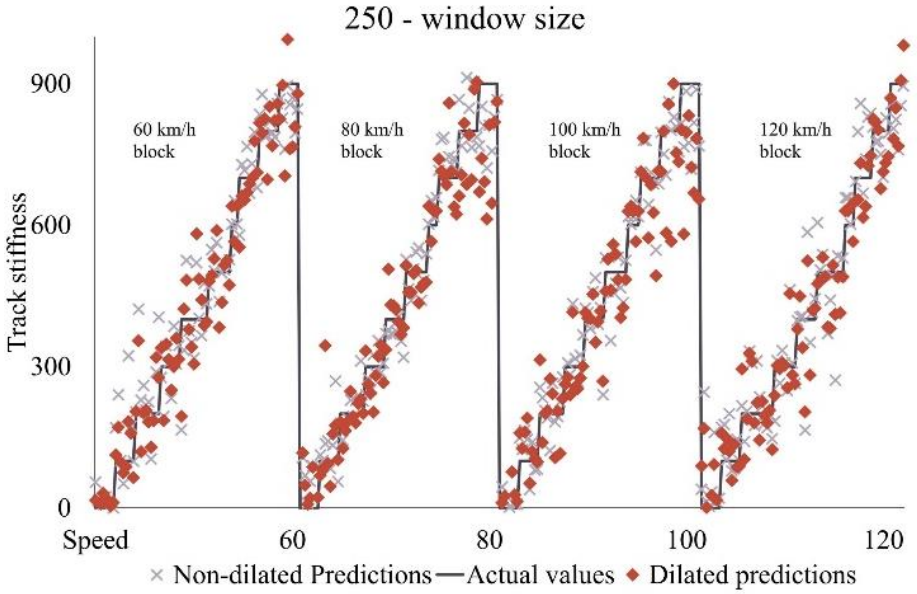


Figure 4-6 Model performances using the test set

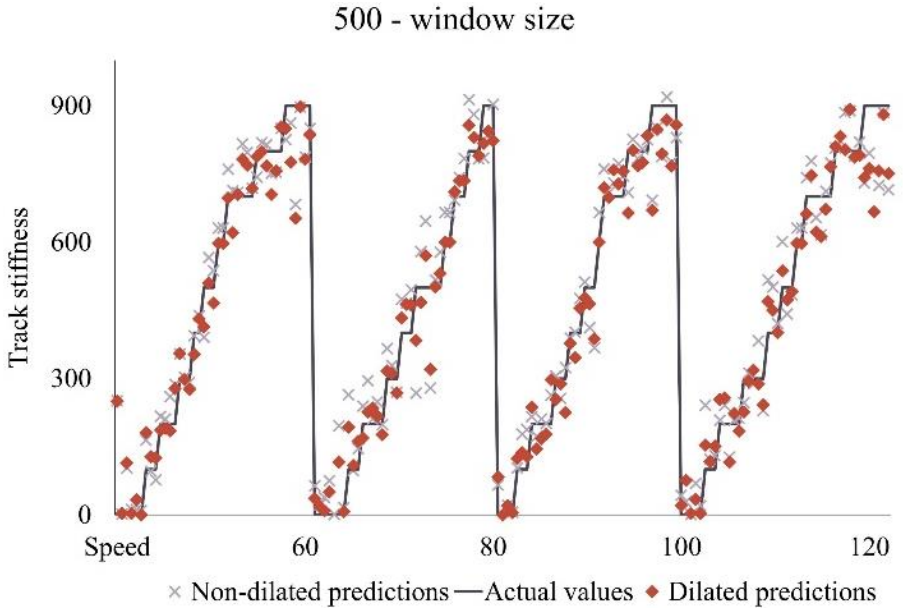
All models achieve over 90% R^2, with peak values obtained using a 1000-time step window size. The performance difference between standard CNN and dilated CNN is minor; however, the dilated CNN significantly reduces computation time by 75%, 45%, and 27% for window sizes 250, 500, and 1000, respectively. This reduction is particularly advantageous given railway networks' extensive and energy-intensive nature.

Figure 4-7 provides a comparison between the predicted values and actual values to see the finer resolution of the performance. Figure 4-7 highlights that all three fine-tuned

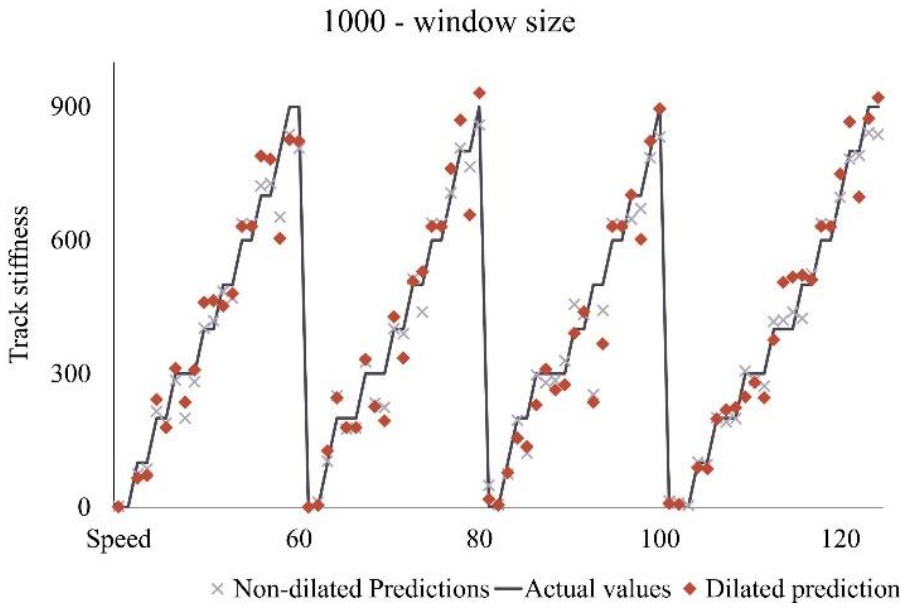
models yield promising results, with errors under 50 MN/m using a 1,000-time step window size.



(a)



(b)



(c)

Figure 4-7 Actual values vs predictions for three datasets (a) 250 (b) 500 (c) 1000

The models' performance remains consistent across different window sizes and is not significantly affected by track irregularities. There is no additional error from the three track irregularities introduced at stiffness levels of 100 MN/m, 600 MN/m, and 800 MN/m compared to other stiffness levels without irregularities.

To assess the impact of speed, Figure 4-7 uses speed as the X-axis, showing four blocks representing different speed settings. The scatter within these blocks indicates no significant performance gap between dilated and non-dilated CNN due to speed variations. The main error arises in the high track stiffness range of 700 MN/m to 900 MN/m, where the model tends to overpredict for 700 MN/m and underpredict for 800 MN/m and 900 MN/m. This issue may be due to the vibration pattern's insensitivity to changes in high track stiffness. A potential solution could be incorporating additional features like acoustics and angular velocity. High stiffness values are uncommon, as the optimal track stiffness is typically between 70 MN/m and 130 MN/m [131-133]. The

proposed framework allows easy retrofitting of service trains without disrupting daily operations. Additionally, the proposed framework supports continuous, fast, and cost-effective track stiffness measurement compared to traditional methods. For instance, a rolling stiffness measurement vehicle costs 7,000 Euro/day (excluding transport and locomotive costs), and a track loading vehicle costs 4,000 Euro/day (excluding data evaluation, transport, and locomotive costs).

4.4 Conclusions

This section proposes an ML model for real-time global track stiffness estimation. It has been confirmed that the model performs satisfactorily, the dilated technique reduces computational costs, and track irregularities have no significant impact on performance. Another important thing that can be confirmed is the correlation between the track characteristics and vibration data found by an ML model under different operational contexts. This finding provides a profound foundation for later studies that can extend the usage of smartphone-based vibration with ML in railway sectors.

One implication of this study is the potential for mounting wireless accelerometers on the axle boxes of service trains. To mitigate the risk of interference with essential signalling and control systems, it is crucial to manage wireless transmissions from accelerometers with precision. This requires rigorous frequency planning and the implementation of appropriate shielding measures to ensure adherence to electromagnetic compatibility standards, thus maintaining the integrity of onboard systems and communication networks. The captured vibrations can be sent to the train operation centre or stored for later analysis by the dilated CNN model. Continuous track stiffness monitoring allows for early maintenance and preventive measures.

Chapter 5

Train Motion Classification

5.1 Introduction to Train Motion Dynamics

Evaluating railway train motion is essential for assessing RC because it directly impacts passenger experience through several mechanisms. The stability and smoothness of the train's motion affect the overall RC, as sudden jolts or sways can significantly reduce comfort. Track irregularities, such as dips and joints, also lead to bumps that affect the RC. The design of seating and suspension systems plays a crucial role in mitigating the effects of these motions. Therefore, this section aims to evaluate and understand train motions.

Train motion encompasses the dynamic movements a train undergoes during operation: acceleration, deceleration, lateral movements, and vibrations caused by track geometry irregularities and operational conditions. These motions are influenced by factors like train speed, axle load, track alignment, and track condition [134]. Understanding train

motion is crucial for enhancing safety, RC, and predictive maintenance. Current methods for classifying train motion rely on traditional sensors—IMUs and GPS—combined with complex analytical models to understand train dynamics and interactions. Traditional approaches to identifying train dynamics are based on train motion models derived from Newton's second law. These models predict train motion by incorporating train and track features: tractive power and track geometry [135, 136]. However, traditional methods typically involve static data analysis that processes data after collection. This introduces a critical time lag between data collection and analysis, delaying identifying issues. For example, if a sensor detects an anomaly—excessive vibration or unusual acceleration—the delay in processing this data means the train might continue operating under potentially unsafe conditions until the issue is reviewed. This reactive approach reduces the system's ability to respond to real-time conditions, which is principle for maintaining safety and operational efficiency in dynamic railway environments [137]. Maintaining precise parameter settings for the mathematical models used in traditional methods is also challenging. Accurate and up-to-date parameters are necessary to simulate train dynamics accurately, but obtaining and maintaining these parameters is difficult due to varying operational conditions, wear of train components, and changes in track conditions. This requires constant fine-tuning of model parameters, without which the analysis can become inaccurate and unreliable [138].

The rapid advancement of smartphone technologies and ML algorithms presents new opportunities for train motion classification. Smartphones with advanced sensors—accelerometers, gyroscopes, and GPS—can acquire accurate and comprehensive motion and location data [139, 140]. This reduces the financial and logistical burdens of deploying and maintaining specialised hardware across a train. ML algorithms can

identify patterns in the data to classify train motion accurately. The real-time data processing capabilities of ML address the time lag inherent in traditional methods, enabling immediate detection of train motion. Moreover, ML enhances predictive maintenance by continuously monitoring motion data, reducing downtime and maintenance costs while improving the reliability and safety of train operations.

5.2 Methodologies

5.2.1 Defining Train Motions

In this section, the train motion is categorised into two groups: the train body's motion and the wheelset's motion. The motion of the train body is further divided into roll, yaw, pitch (front/back), and bounce (vertical), as well as the motion of the wheelset - hunting behaviours.

5.2.1.1 Motions of the Train Body

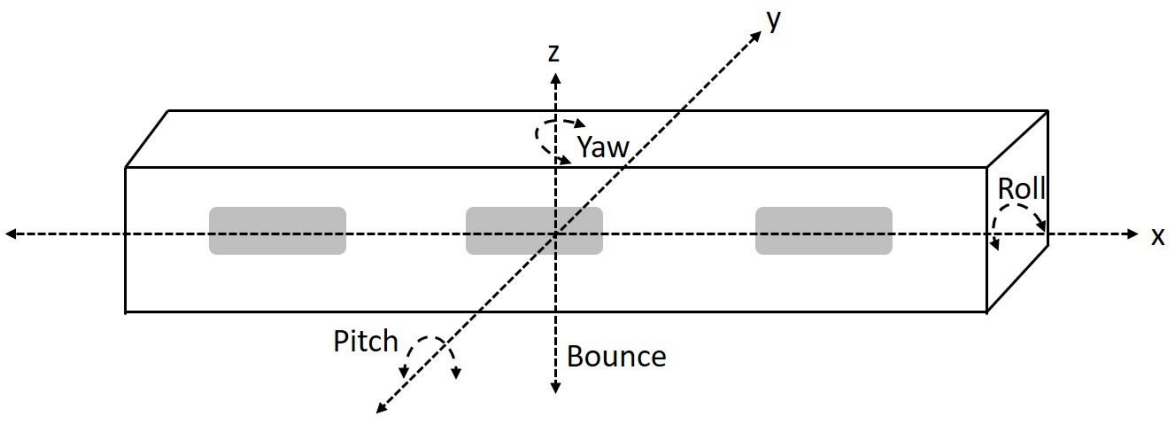


Figure 5-1 Train motions

Figure 5-1 schematically illustrates the different types of train body motion. Rolling refers to the rotational movement of the train body around the X-axis, typically occurring when

the train traverses curves with a tilting mechanism, where acceleration along the X- and Z-axes is predominant. Yawing describes the twisting movement of the train body, where the front and rear ends move laterally in opposite directions. Pitching involves rotational movement along the lateral axis, resulting in forward and backward tilting, often caused by uneven track surfaces. This motion can impact load distribution across the wheelsets and the train's traction and braking performance. Bouncing refers to the vertical motion of the train body, usually due to track irregularities such as dips or humps.

5.2.1.2 Hunting Behaviour

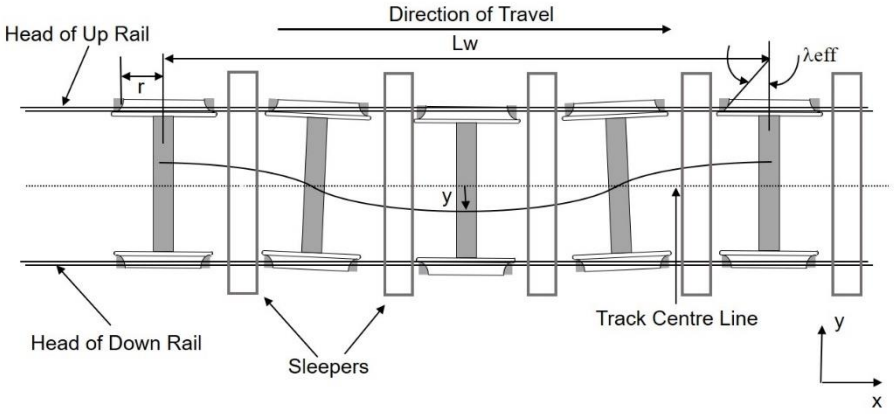


Figure 5-2 Hunting behaviours

Hunting behaviour in railway systems refers to the oscillatory motion of train wheelsets, where the wheels move back and forth across the rails laterally. This side-to-side movement typically occurs at high speeds and is driven by the dynamics of wheel-rail interaction. Figure 5-2 illustrates the hunting behaviour, characterised by lateral displacement *y* and significant lateral forces [141]. The hunting problem in high-speed railways was systematically analysed by T. Matsudaira [142] decades ago, demonstrating that hunting behaviour presents substantial challenges to the stability and speed optimisation of trains. Hunting frequencies range from 2 to 8 Hz [143], depending on the

characteristics of the suspension components. This frequency range is a differentiating feature for distinguishing hunting behaviour from other types of motion.

The wavelength L_w is determined by the wheel radius, the distance between the contact points on the rails, and the effective conicity of the wheels. The wavelength directly influences the stability and smoothness of train operations. A shorter wavelength can amplify hunting oscillations, leading to instability and increased wear on both rails and wheels. Conversely, a longer wavelength can enhance stability but might affect the RC [141]. Accurately estimating the wavelength effectively prevents dynamic issues in rail vehicles, ensuring safe and efficient operations [142]. The wavelength L_w is defined as:

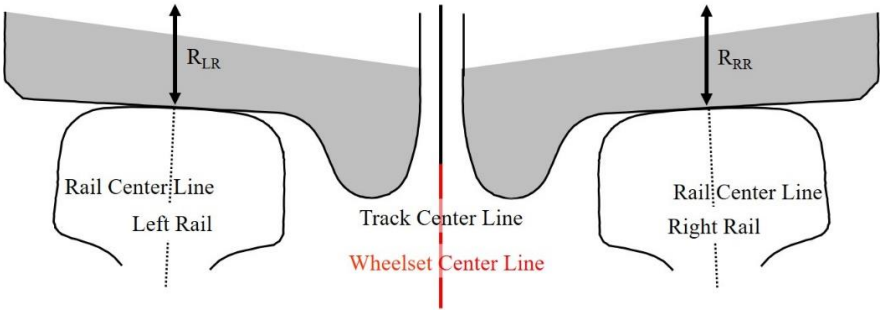
$$L_w = 2\pi\sqrt{\frac{rs}{2\lambda_{eff}}} \qquad \text{Equation 30}$$

Where r is the radius of the wheel, s is the distance between contact points on the rails (approximately the track gauge), λ_{eff} is the wheel's effective conicity. Based on this, the wavelength for bogies L_b is defined as:

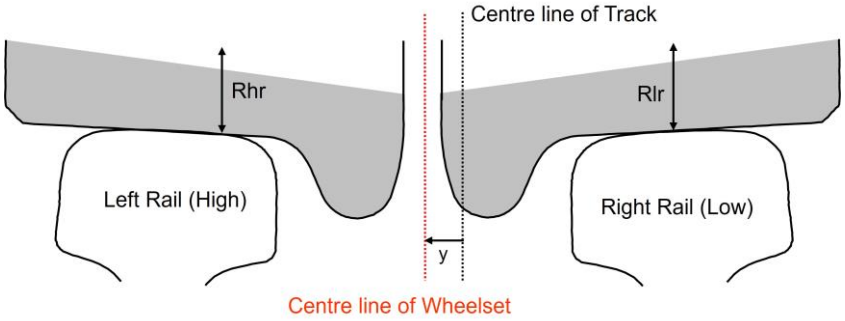
$$L_b = L_w\sqrt{1 + \frac{w}{s}} \qquad \text{Equation 31}$$

Where w is the bogie wheelbase.

Due to the hunting behaviour, the wheelset moves out of its equilibrium position (where the centre of the track and the centre line of the wheelset overlap, as shown in Figure 5-3(a), resulting in a difference in rolling radius (ΔR) as depicted in Figure 5-3(b).



(a)



(b)

Figure 5-3 Rolling radius differential (a) the equilibrium position (b) the position away from the equilibrium position

Controlling ΔR is crucial as it impacts the lateral force within the safe limit of 10 mm specified in EN14363 [144]. Excessive ΔR causes increased wear and tear on the wheel and rail, leading to higher maintenance costs and potential safety hazards such as derailments. Optimising ΔR can minimise wheel-rail contact stresses, thereby reducing rail fatigue and wheel flange wear [141]. Accurate control of ΔR enhances the RC by reducing lateral vibrations and noise [145]. Therefore, this study also estimates ΔR for hunting behaviours to monitor the range of ΔR since maintaining ΔR within safe limits can extend the lifespan of both wheels and rails. The ΔR is defined as:

$$\Delta R = 2\ \lambda_{eff}\ y \qquad \text{Equation 32}$$

5.2.2 Data Acquisition Methods

This section outlines the data collection process, including sensor deployment and data acquisition. Collecting data using passengers' smartphones raises concerns about handling sensitive information, such as location and motion data. Adhering to data privacy and security standards is necessary to protect this information. It is imperative to follow the General Data Protection Regulation (GDPR) in Europe, which provides guidelines for data protection and privacy. This includes obtaining informed consent from passengers, anonymising data, and implementing robust security measures [146].

Smartphone deployment strictly adheres to the ISO 2631 standard, which provides guidelines for evaluating human exposure to WBV. The reason for employing ISO 2361 is that this chapter aims to evaluate the motion at locations where ISO 2631 is considered crucial to RC. Smartphones are placed in multiple train carriages to ensure comprehensive data capture. Each smartphone is securely mounted at designated locations, such as seats and tables, with the X-axis of the phone aligned with the train's running direction.

This phase is critical for train motion classification, ensuring that the collected data reflects the dynamic behaviours experienced by a train during operation. Dynamic movements of a train include longitudinal, lateral, and vertical acceleration. Data acquisition is conducted between Selly Oak and Redditch stations in the UK, using different trains and carriages at various times (peak and off-peak) to account for different loading and crowding conditions.

Ensuring the accuracy and reliability of the collected data is a critical aspect of data acquisition. Sensor calibration is a prerequisite for high accuracy and reliability, avoiding

biases or drifts that might affect measurements. Studies have shown that sensor data accuracy can vary significantly between different phone models and manufacturers [147, 148], highlighting the need for phone calibration before data collection. Routine calibration ensures that the collected data accurately reflects operational conditions, improving data quality for ML models. Once deployed, the smartphones continuously record acceleration in three dimensions during train operation. The data is stored locally and transferred to computers for later processing.

5.2.3 Extraction of Train Motions

Data pre-processing begins by dividing the continuous data stream into manageable segments. This process is event-driven, with filter conditions set to label the acquired accelerations according to different types of motion, as shown in Figure 5-4. R.M.S values for all segmented windows are calculated, and conditions are established to filter out various types of motion. Each refined window is visualised to verify the integrity of the label. The conditions for identifying different types of train motion are fine-tuned to improve label quality. It is important to note that acceleration along the Y-axis is predominant in yawing and hunting. However, hunting behaviour is distinguished by higher frequency oscillations in the Y-axis.

The chunks, resulting from equal-sized intervals, facilitate the subsequent feature extraction process before addressing missing values. Missing values and incomplete data, often due to phone performance issues, can distort the analysis. Therefore, missing values must be imputed using statistical methods, or incomplete records should be removed if they constitute only a small portion of the entire dataset [149].

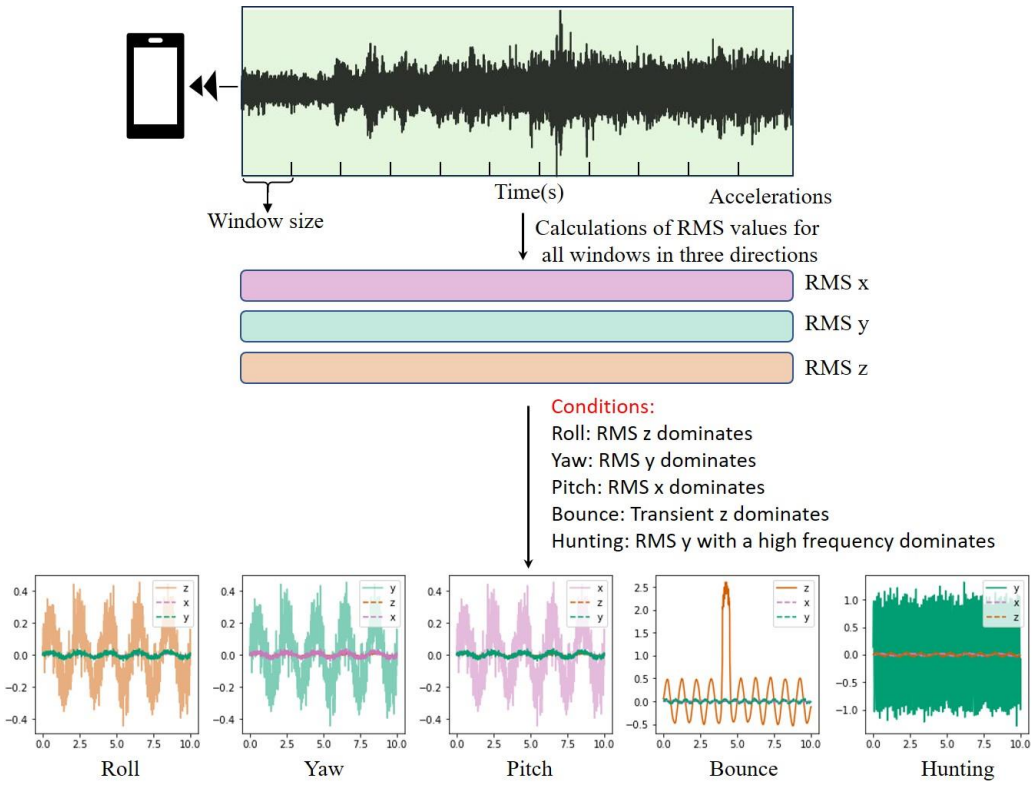


Figure 5-4 Motion labelling

5.2.4 Feature Extraction for Model Training

The primary aim of this subsection is to extract meaningful features for modelling. This study employs both time-domain and frequency-domain feature extraction techniques.

In the time domain, a set of statistical features that describe the fundamental characteristics of motion signals is considered. These features include the acceleration data's maximum, minimum, mean, standard deviation, R.M.S, skewness, and kurtosis. The maximum and minimum values provide insights into the extreme conditions experienced by the train, highlighting peaks and troughs in the motion data. The mean value represents the overall trend of motion over time, while the standard deviation measures the dispersion of the signal, indicating the level of fluctuation in the train's motion. R.M.S. unfolds an understanding of the overall energy. Skewness quantifies the

asymmetry of the signal distribution, and kurtosis measures the presence of outliers by indicating the extent of data residing in the tails. These features are defined as follows:

$$\max = \max(x_i) \quad \text{Equation 33}$$

$$\min = \min(x_i) \quad \text{Equation 34}$$

$$mean = \frac{\sum x_i}{n} \quad \text{Equation 35}$$

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{n}} \quad \text{Equation 36}$$

$$RMS = \sqrt{\frac{1}{n}\sum_{I} x_i^2} \quad \text{Equation 37}$$

$$\tilde{\mu}_3 = \frac{\sum_i^n (x_i - \bar{x})^3}{(n-1) * \sigma^3} \quad \text{Equation 38}$$

$$Kurt = \frac{\mu_4}{\sigma^4} \quad \text{Equation 39}$$

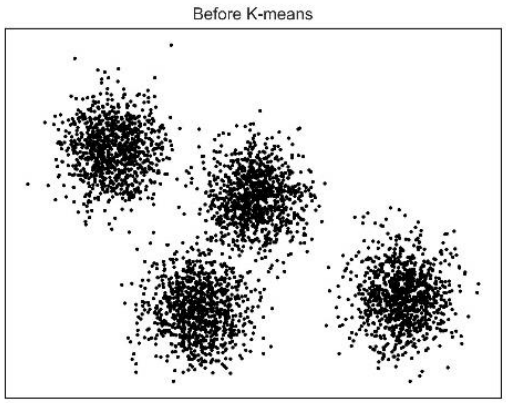
In the frequency domain, the analysis begins by applying the FFT to the time-domain signals to extract the frequency components of the motion data and reveal periodic patterns and dominant frequencies that are not easily found in the time domain. The frequency spectrum is then divided into several bands, and the energy—a measure of the signal's power within each band—is calculated as frequency-domain features. These features reveal how energy is distributed across different frequencies, which helps differentiate various types of motion.

By combining time-domain and frequency-domain features, a comprehensive representation of the different types of motion is obtained. This approach not only provides a holistic view but also includes features that can distinguish between different motions.

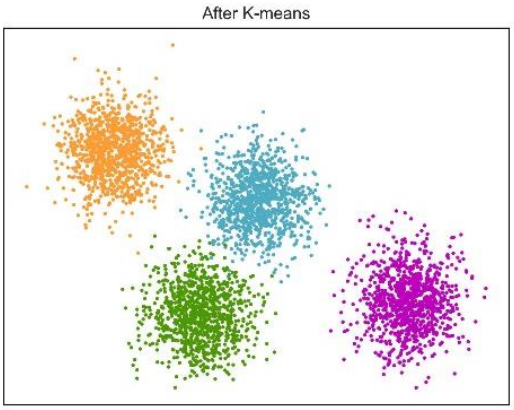
5.2.5 Model Development and Training

5.2.5.1 K-means

K-Means clustering is a powerful and widely used unsupervised learning algorithm. Its advantages include the ability to handle high-dimensional data and computational efficiency, making it suitable for real-time analysis [150]. By dividing data into *K* groups, K-Means can use predefined time and frequency domain features to categorise various motion patterns. This capability is used to distinguish normal operational states and abnormal states [151, 152]. Additionally, K-Means is highly compatible with large datasets due to its computational efficiency [151].



(a)



(b)

Figure 5-5 K-means clustering

Figure 5-5 showcases the function of K-means clustering. Figure 5-5(a) shows an unlabelled and ungrouped dataset, while Figure 5-5(b) demonstrates the application of K-means clustering to the same dataset. K-means also provide interpretable output by defining centroids for each cluster. Model interpretability is crucial for engineers, enabling them to understand the underlying factors contributing to the output [153]. Additionally, an interpretable model enhances transparency for stakeholders and builds trust in the system's capabilities [154, 155].

The technical function of K-means can be categorised into two phases: selecting K centroids and assigning each data sample to the nearest centroid [151, 156]. The detailed process is outlined step-by-step as follows:

1. K-means starts with defining K initial centroids. These centroids can be selected randomly or using more sophisticated methods, such as the K-means++ algorithm, to improve convergence.
2. Each data point is assigned to the closest centroid, typically measured using Euclidean distance, as defined in Equation 40.

$$d(x_i, c_j) = \sqrt{\sum_{m=1}^{D} (x_{im} - c_{jm)}{}^2} \qquad \text{Equation 40}$$

Where $d(x_i, c_j)$ represents the Euclidean distance calculated from a data point x_i and a centroid c_j; *D* is the dimensionality of the data.

3. After all samples are assigned to a specific group based on step 2, the centroids are recomputed by taking the mean of all samples in that cluster. If a cluster *Cj* contains n_j samples, the new centroid c_j is calculated as:

$$c_j = \frac{1}{n_j} \sum_{x_i \in C_j} x_i \qquad \text{Equation 41}$$

4. Steps 2 and 3 are repeated iteratively until the model converges, meaning the centroids no longer undergo significant changes. This process refines the clusters and minimises the distance (variance) between samples of the same cluster, which can be defined as:

$$V = \sum_{j=1}^{K} \sum_{x_i \in C_j} \left\| x_i - c_j \right\|^2 \qquad \text{Equation 42}$$

5. The algorithm converges when the assignment of samples to clusters remains unchanged between iterations or when the change in centroids falls within a predefined threshold.

5.2.5.2 Agglomerative Clustering

Agglomerative clustering, a type of hierarchical clustering technique, begins by treating each sample as an individual cluster. Samples are then progressively merged based on their similarity until all samples are combined into a single cluster or specific conditions are met [157]. Figure 5-6 illustrates this process with a dendrogram representing the hierarchy and distances. In the example shown, the algorithm first merges data points '15' and '16' into a cluster due to their minimal distance. Based on distance measurements, this merging process is repeated for other data points until all samples are joined into a single cluster, as indicated by the blue line in Figure 5-6.

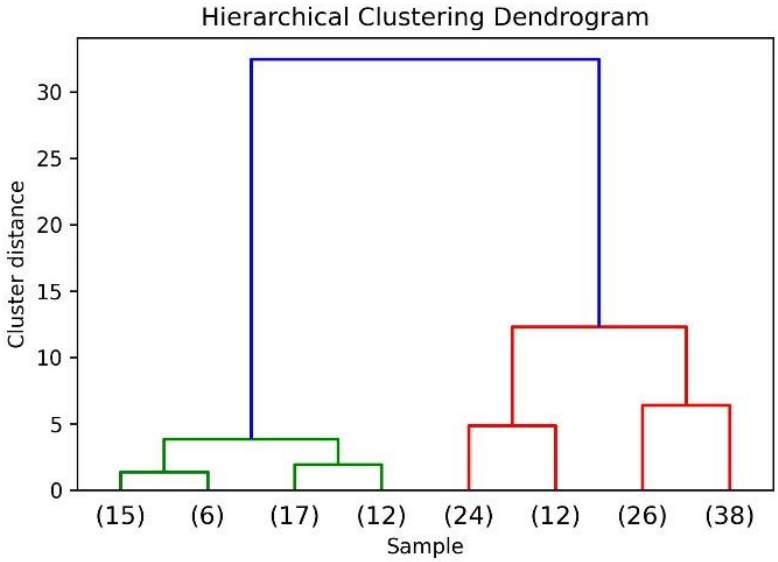


Figure 5-6 Agglomerative clustering

The hierarchical structure is easy to interpret and visualise using dendrograms, resulting in an explainable and interpretable model. The model's adaptability to different types of data and various distance measures makes it beneficial for analysing train motion data. However, this method can be computationally intensive, as it requires calculating distances between all pairs of clusters, leading to poor scalability. Despite these limitations, agglomerative clustering is introduced in this section as a benchmark model alongside K-means, thereby enhancing the robustness of the outcomes produced by K-means.

5.2.6 Results and Discussions

5.2.6.1 Dataset

This section outlines the correlations between types of train motion and the raw data presented in the subsection "Filtering Motion". After refining the motion data, extracted

features are obtained for subsequent use in ML models. Details are provided in the subsection “Extracted Features”.

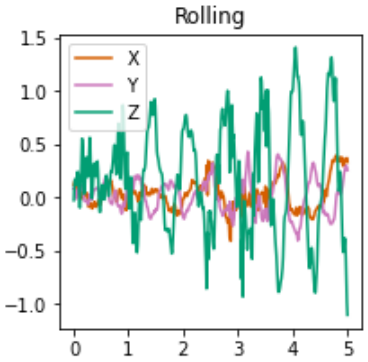
Data collection is conducted over seven journeys, with multiple phones placed in different carriages. Data segmentation is performed using 5-second windows, resulting in 7,633 samples, as this duration is sufficient to capture and settle the motion, as ISO 2631 specifies that the duration should be long enough to capture representative vibration data [158].

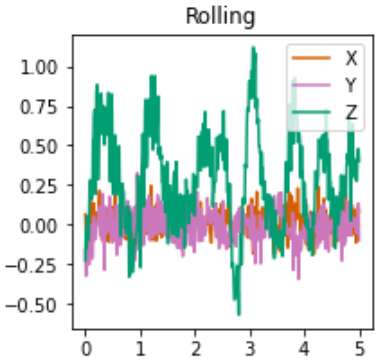
Filtering Motion

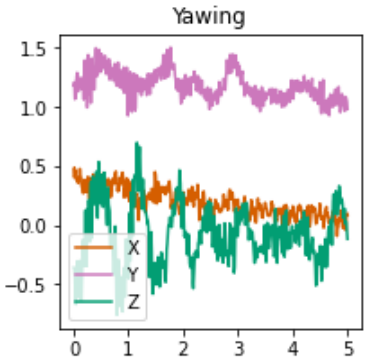
Table 10 Conditions for labelling different types of motion

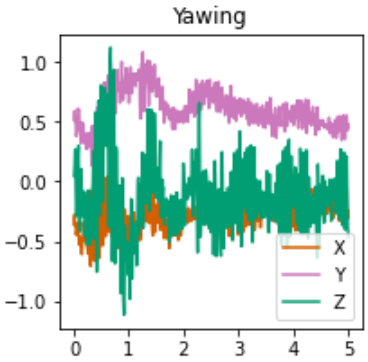
Motion	**R.M.S x**	**R.M.S y**	**R.M.S z**	**Max z**	**Frequency y**
Roll	< 0.1	< 0.1	> 0.4	-	-
Yaw	< 0.1	> 0.4	< 0.1	-	< 4 Hz
Pitch	> 0.4	< 0.1	< 0.1	-	-
Bounce	-	-	-	> 4	-
Hunting	-	> 0.4	-	-	4 Hz – 8 Hz

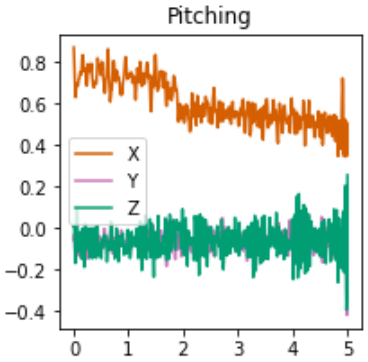
By applying the conditions outlined in Table 10, relevant windows are labelled according to the characteristics of the motion. Out of the 7,633 windows, there are 12 instances of rolling, 38 of yawing, 12 of pitching, 2 of bouncing, and 7 of hunting.

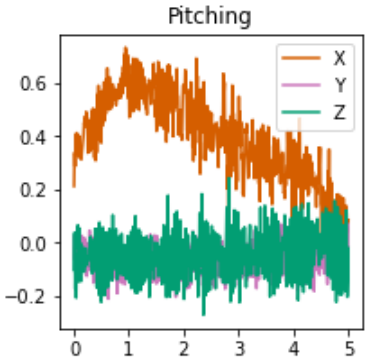

Rolling
X
Y
Z

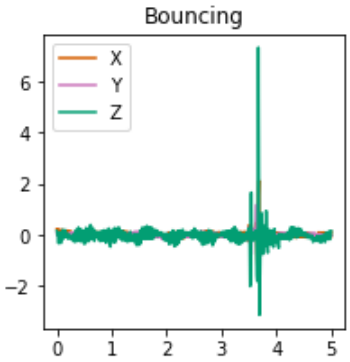

Rolling
X
Y
Z

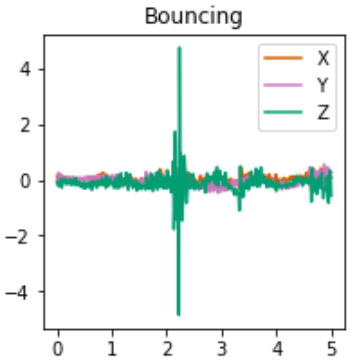

Yawing
X
Y
Z


Yawing
X
Y
Z


Pitching
X
Y
Z


Pitching
X
Y
Z


Bouncing
X
Y
Z


Bouncing
X
Y
Z

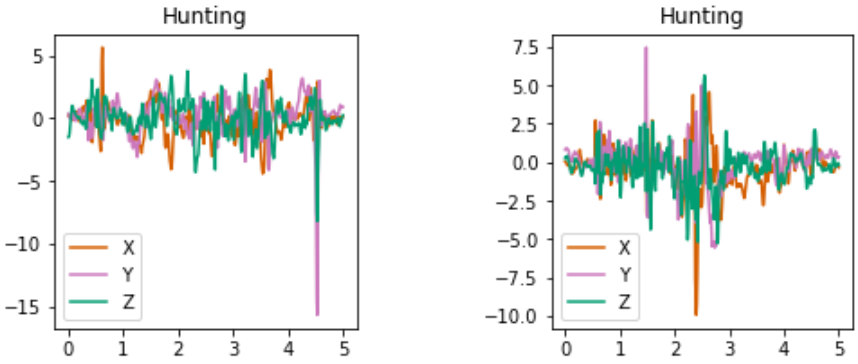


Figure 5-7 Train motions

Figure 5-7 depicts the connection between raw data and different types of motion. Rolling, yawing, and pitching are primarily influenced by one of the three axes. For instance, the Z-axis acceleration has the most substantial influence on rolling, although it has a transient effect. The rolling radius differential (RRD) between the wheels of a railway vehicle is a significant factor affecting hunting behaviour. RRD refers to the difference in the effective rolling radius of the wheels, which becomes notable when a train passes a curve, causing the outer wheel to travel a greater distance than the inner wheel. This difference is managed by the conicity of the wheels, allowing them to adjust their rolling radius dynamically.

Studies have shown that a sizeable rolling radius differential exacerbates hunting behaviours because the differential creates asymmetrical forces that push the vehicle laterally across the track, amplifying oscillations [159, 160]. Therefore, regular maintenance is required to minimise the rolling radius differential and reduce hunting tendencies [161].

Extracted Features

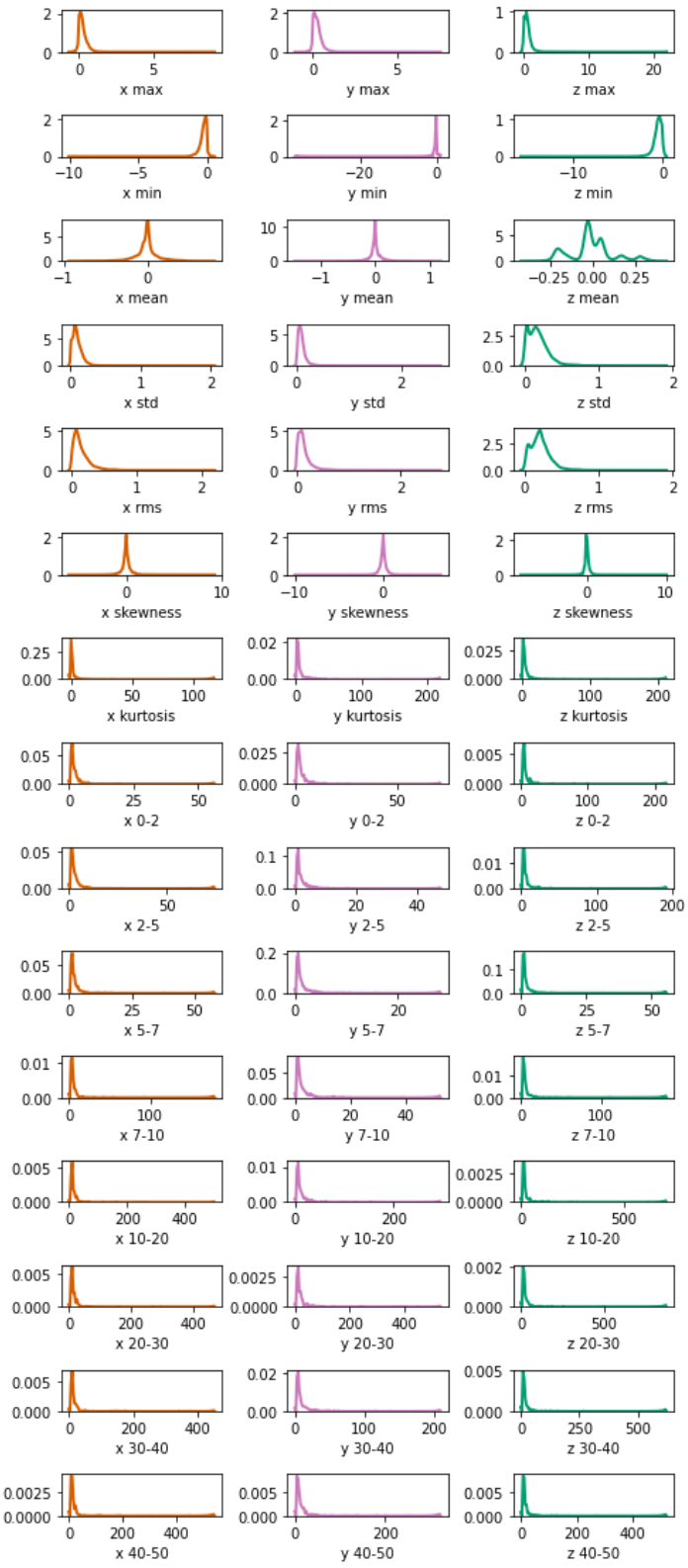


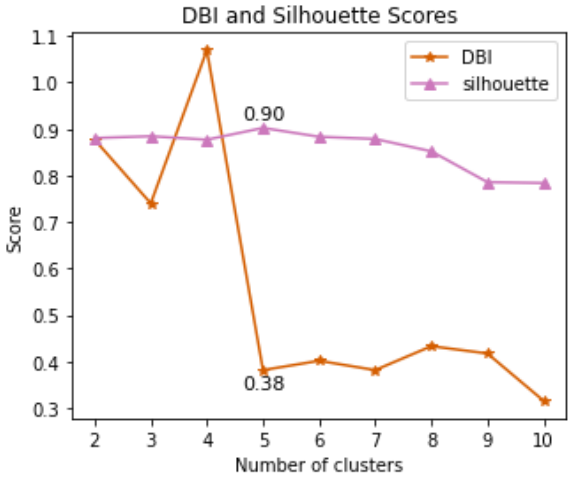
Figure 5-8 Distributions of x, y, and z features using kernel density estimation

Figure 5-8 presents the post-processed data using Kernel Density Estimation (KDE) to display the distributions of the x, y, and z features. KDE is a non-parametric method for estimating the probability density function of a dataset, providing a smooth estimate without assuming a specific underlying distribution [162]. This method effectively reveals the distributional characteristics of the data and identifies patterns, peaks, and anomalies that are not easily detected by parametric methods [163]. The X-axis in each plot represents the range of values, while the Y-axis denotes the density, illustrating the probability of different features occurring in the dataset.

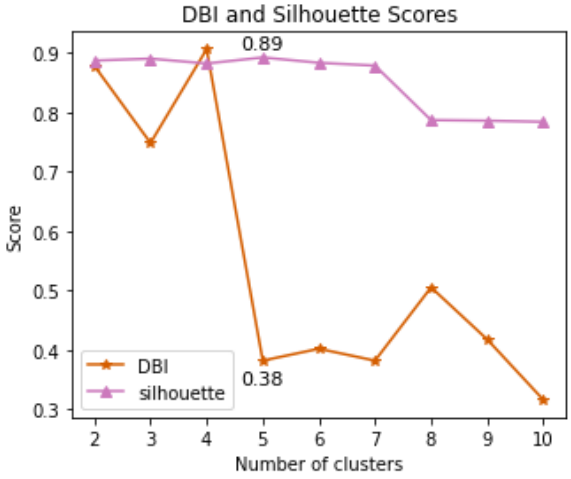
The first part of the figure shows the time-domain features. Among these, features in the x and y directions exhibit a more consistent pattern, with most values centred around 0. However, features in the z direction show more significant fluctuations. The latter part of the figure displays the frequency-domain features, with directions and band ranges indicated, such as 'x 2-5', which represents the energy in the x direction within the frequency range of 2 to 5 Hz. The energy concentrations are predominantly within the lower frequency range (0–10 Hz), indicating that most motion energy falls within this span.

Selecting Best K

This study uses varying the number of clusters K to select the optimal K. The best K is determined by the most significant silhouette coefficient and the lowest DBI.



(a)



(b)

Figure 5-9 DBI and silhouette for K-means (a) K-means (b) Agglomerative

Figure 5-9 shows the DBI and Silhouette scores for different numbers of clusters, evaluated by two unsupervised learning models. Both models indicate minimum DBI values at 5 clusters. Simultaneously, the highest silhouette coefficient is observed at 5 clusters for both models. The alignment of these metrics at 5 clusters in both models highlights its validity as the optimal choice for clustering. This agreement supports the model's robustness in accurately identifying and clustering the majority of the samples.

The Performance for K-mean Clustering

By setting K to 5, the K-means model is trained and yields the results in Figure 5-10.

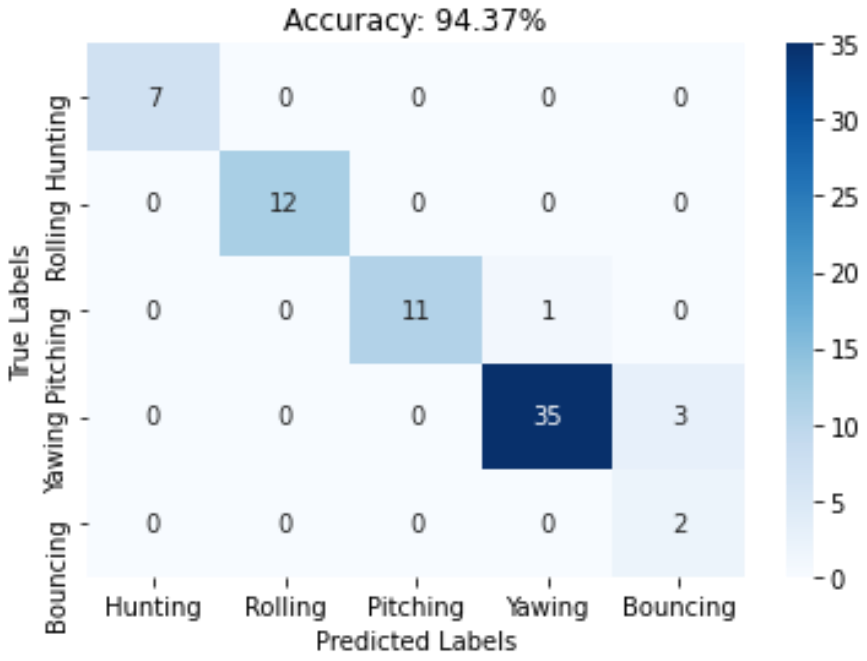


Figure 5-10 Confusion matrix for K-means clustering

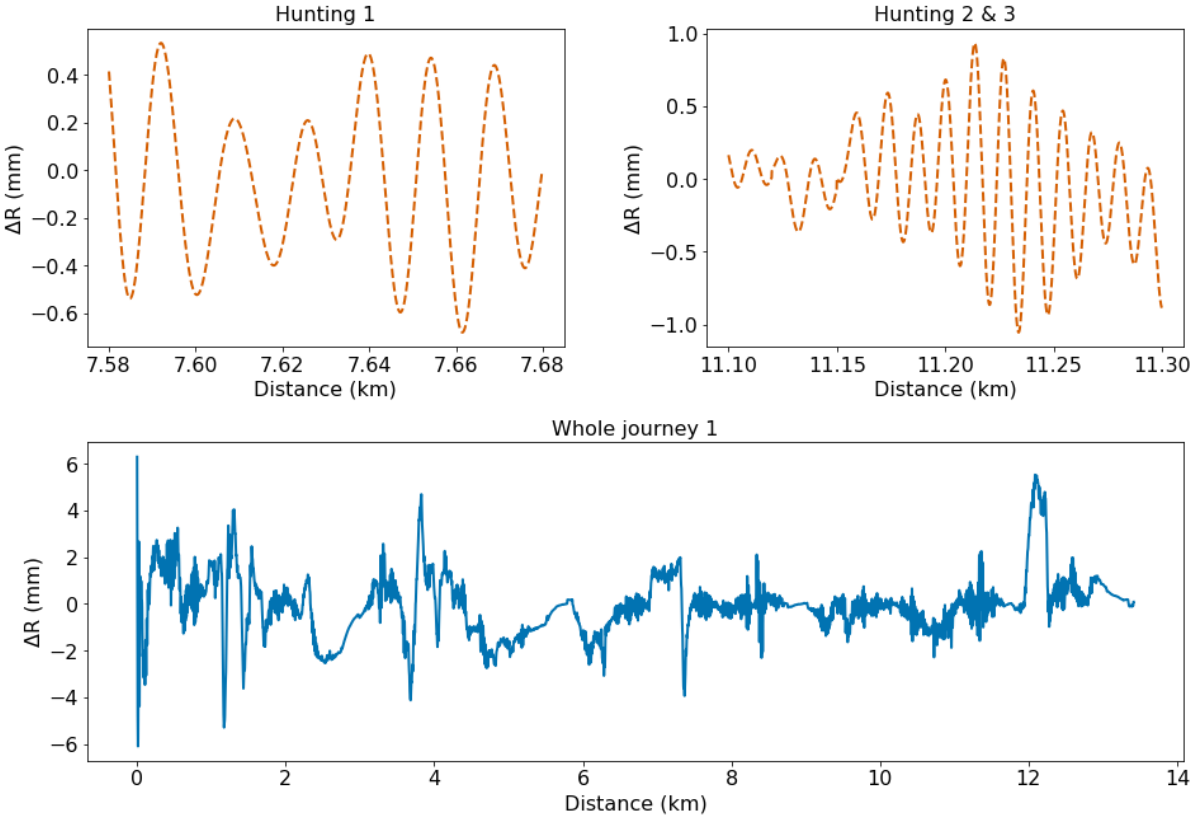
Figure 5-10 presents a confusion matrix illustrating the performance of K-means clustering, with an overall accuracy of 94.37%. The matrix shows clear diagonal dominance, indicating that most samples are correctly clustered. Specifically, 7 hunting, 12 rolling, and 2 bouncing motions are all correctly assigned.

Notably, there are misclassifications between pitching and yawing, as well as between yawing and bouncing. A detailed examination of these errors provides insights into potential limitations of the model and areas for improvement in the future direction. The misclassification indicates that the model cannot distinguish between certain types of train motion. This challenge may stem from compound motions that the clustering algorithm cannot detect. The underlying cause of these misclassifications is likely the

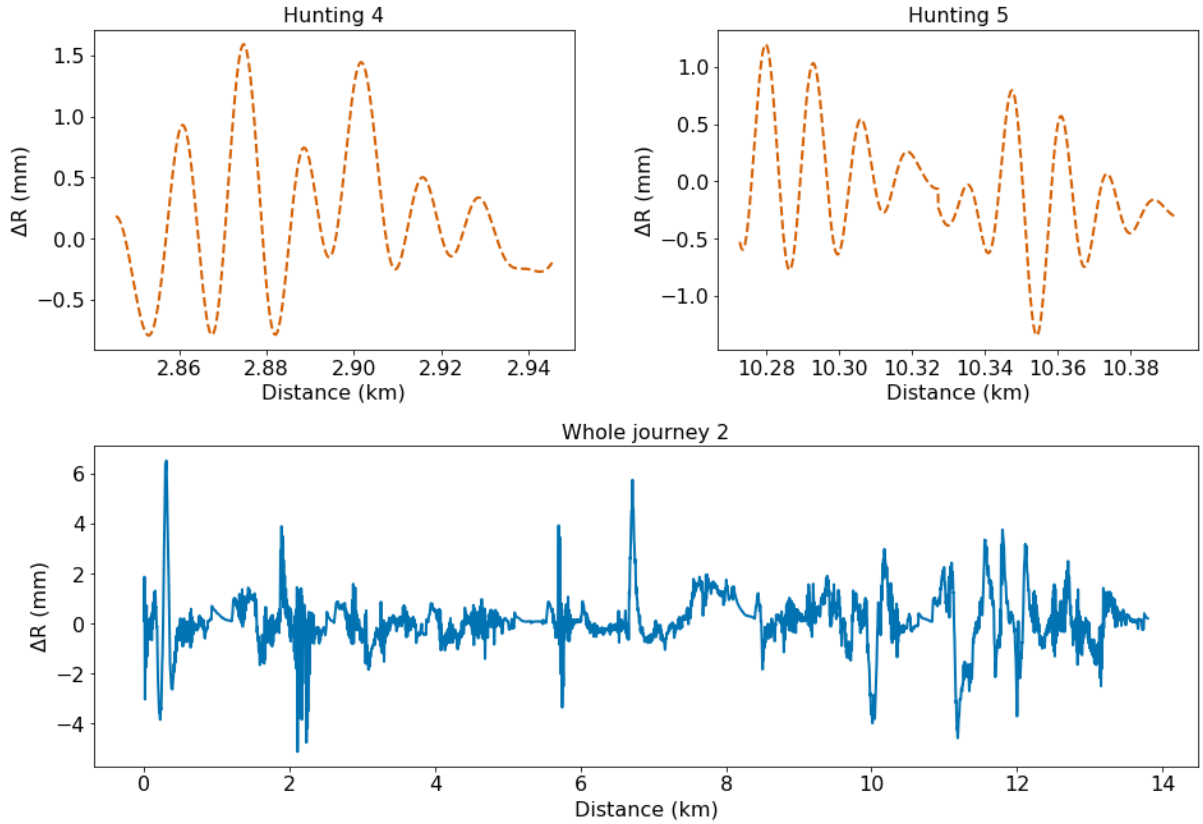
overlapping features in compound train motions. When multiple types of motion coincide, the signals they generate can share similar characteristics, making it difficult for the model to assign classes accurately. Additionally, the current set of features used for classification may not adequately capture the differences between the compound motion types. If the feature extraction process focuses on individual characteristics like frequency or amplitude without considering how these features interact in compound motions, misclassifications are likely to occur.

To address these issues, several improvements could be considered. One potential approach is to implement a multi-label classification system. This would allow the model to assign multiple labels to each sample, more effectively capturing the complexity of compound motions. In addition, enhancing the feature extraction process to represent interactions between different types of motion better could significantly improve the model's performance.

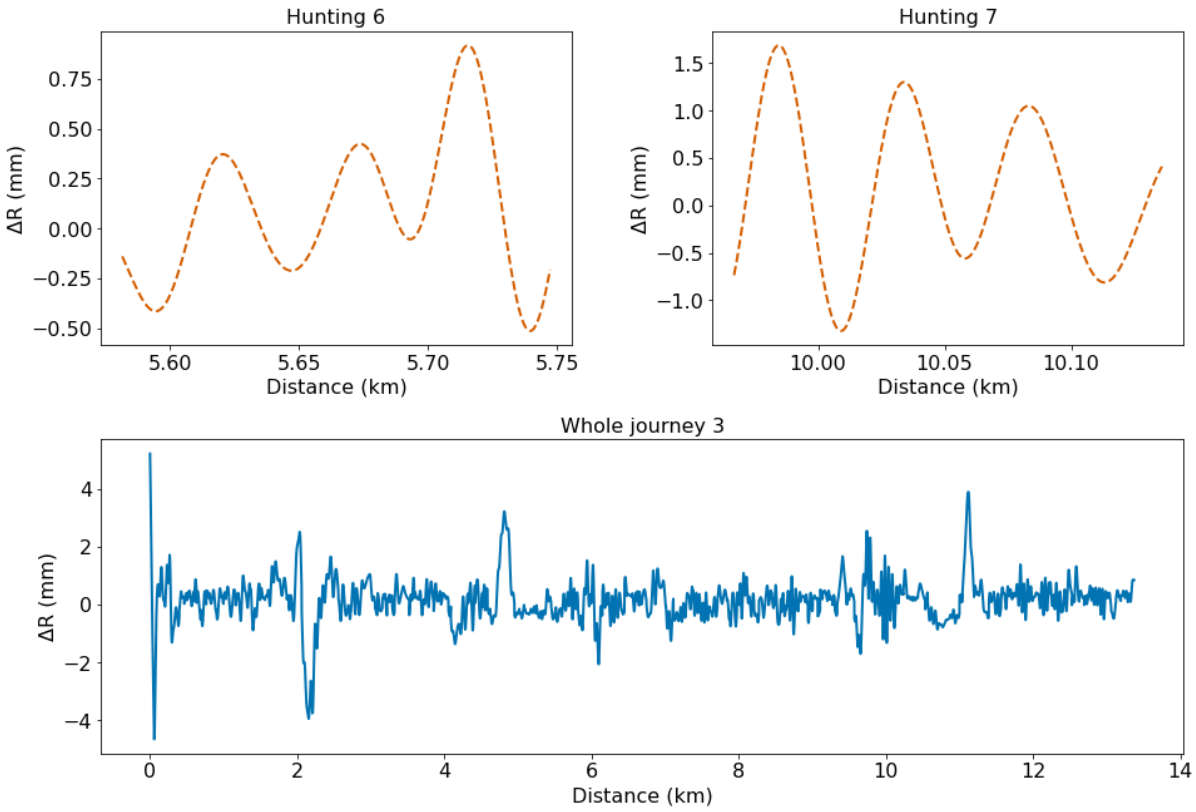
Data collection is conducted on a Class 323 train. Given the standard gauge of 1.435m [164], a bogie wheelbase of 2.5m, a wheel radius of 0.46m [165], and an effective conicity of 0.05 [166], L_w is calculated as 16.1 using Equation 30 and L_b as 26.7 using Equation 31. The rolling radius differentials for the 7 hunting behaviours are calculated using Equation 32 and shown in Figure 5-11.


Hunting 1
ΔR (mm)
Distance (km)
Hunting 2 & 3
ΔR (mm)
Distance (km)
Whole journey 1
ΔR (mm)
Distance (km)

(a)


Hunting 4
ΔR (mm)
Distance (km)
Hunting 5
ΔR (mm)
Distance (km)
Whole journey 2
ΔR (mm)
Distance (km)

(b)



(c)

Figure 5-11 Rolling radius differentials for hunting behaviours in three journeys

An increasing ΔR suggests growing lateral displacement, while a decreasing ΔR indicates a reduction in lateral movement. Notably, a sharp slope in ΔR implies rapid changes in wheel-rail contact conditions. Significant and rapid changes in ΔR can lead to increased lateral forces, potentially causing wheel flange contact with the rail and, in extreme cases, derailment. Stability is compromised when ΔR oscillates dramatically, which causes dropping RC and increased dynamic loads on the track structure.

Predicting ΔR allows for real-time control to improve train stability and safety. The suspension system can be adjusted to accommodate variations in ΔR and mitigate instability. Additionally, this prediction can signify the need for maintenance practices to reduce the risks associated with hunting behaviours.

5.3 Conclusions

Chapter 5 explores the train motions by integrating smartphone sensors and unsupervised learning techniques. The findings demonstrate that leveraging smartphone sensors for monitoring train dynamics enhances the precision of motion detection while substantially reducing the costs associated with traditional sensor systems. RC closely correlates with accurately classifying train body motions—such as roll, yaw, pitch, and bounce—and wheelset behaviours like hunting. The ML model proficiently differentiates these motions by analysing features in both the time and frequency domains, which provides the potential for the real-time identification and management of dynamic train behaviours. Enhanced monitoring of lateral forces, vertical accelerations, and other critical parameters significantly contributes to the stability and smoothness of train operations. This proactive approach to detecting and classifying train motions improves RC and supports predictive maintenance strategies. More importantly, it reduces risks and improves overall safety in railway systems.

Building upon the insights from Chapter 5, which explores the dynamics of train motions using advanced classification techniques, the understanding of different types of motion is now well-established. The accurate identification of motions such as rolling, yawing, and hunting plays a critical role in predicting and mitigating the factors that influence RC. With these motion profiles clearly defined, Chapter 6 quantifies RC at specific points within the train. Leveraging the methodologies developed earlier, Chapter 6 seeks to refine the real-time estimation of comfort levels, demonstrating the application of ML models and smartphone-based data collection to provide a localised understanding of

RC. This transition marks a step towards integrating these findings into practical, real-world applications to improve overall passenger satisfaction.

Chapter 6

Point-Specific Train Ride Comfort Quantification

6.1 Introduction to Localised Comfort Assessments

Chapter 4 establishes the correlation between vibration features and track characteristics. Chapter 5 utilises ML algorithms and smartphone data to estimate train movements relevant to RC. The findings indicate that integrating vibration data and smartphone technology benefits on-board train analysis. Consequently, there is increased confidence in the proposed methodology's capability for RC estimation. This chapter applies the methodology to estimate RC at a single location on a train. Concurrently, the chapter explores the methodology's potential to predict RC at specific points before advancing to crowd-sensing techniques for evaluating the overall vehicle's RC.

Most existing research recognises the critical role of vibrations in RC. The study asserts [167] that vibration-related RC for passenger trains and vibration-related safety concerns for freight trains should be addressed from the design stage. Other studies, such as [168], have aimed to enhance bogie suspension systems to increase RC. A study [169] employed a numerical model validated with field tests, concluding that rail irregularities with short wavelengths significantly impact the RC. Subsequently, [170] utilised a magneto-rheological (MR) damper to mitigate the adverse effects of large amplitude vibrations caused by rail irregularities. A full-scale examination of the MR damper at speeds ranging from 80 to 350 km/h demonstrated excellent comfort according to UIC 513 standards. A study [171] identified a strong relationship between the number and distance of train/bus stops and RC levels. The common motivation behind the studies in [168-171] is the importance of vibrations: improved bogie suspensions can better isolate vibrations from car bodies; irregularities cause dynamic train movements affecting vibrations; and more frequent stops lead to increased acceleration and deceleration, thereby impacting the RC. Combining the support from the well-established literature and this thesis on the application and feasibility of using vibrations, it is rational to adopt this approach for calculating RC.

6.2 Methodologies

6.2.1 Data Collection for Point-Specific Assessments

In this chapter, UIC 513 is used to calculate the RC, and the details of UIC 513 are provided in Chapter 3.3.4. However, Figure 6-1 outlines how UIC 513 is applied in this chapter to validate if the proposed method is functional on other standards apart from ISO 2631.

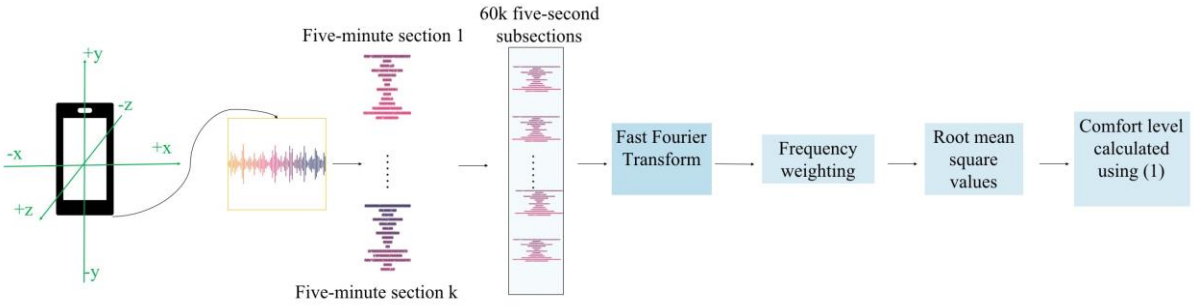


Figure 6-1 Flowchart to calculate the RC using UIC 513

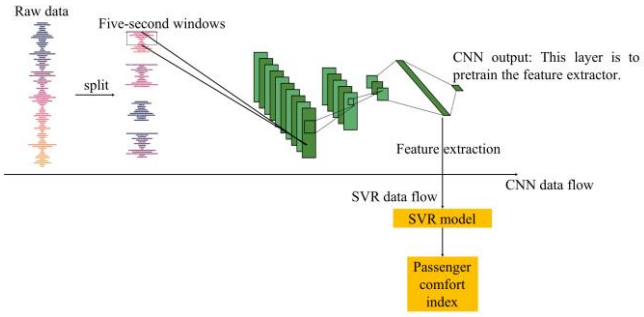
As illustrated in Figure 6-1, a smartphone is used to acquire vibration data in three directions. Data is collected using a smartphone during a round trip from University Station to Birmingham International Station. The phone is placed stably on the floor with no relative movement to the moving train to ensure that the acceleration data recorded by the phone was solely from the train.

According to UIC 513, RC is evaluated in five-minute sections. Each five-minute section is divided into 60 five-second sub-sections, which are then transformed into the frequency domain and weighted based on Figure 3-8. The R.M.S value of each sub-section is calculated. Finally, a 95% confidence probability is applied to compute the RC using Equation 10. The comfort index can be referenced in Table 7 for UIC 513. However, this study employs a slightly different approach to enable passengers to quantify their comfort levels using their smartphones. This method requires passengers to keep their phones stationary for five minutes, as specified in UIC 513, which is impractical. More importantly, the comfort level is evaluated on a more detailed scale of five seconds to estimate the RC in a finer resolution.

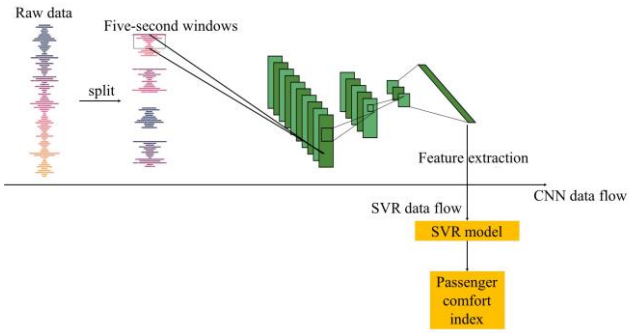
6.2.2 Developing an Ensemble Model for Comfort Quantification

CNN is initially designed for image classification. However, their application is extended in this chapter for two reasons: (1) CNN provide a more efficient computational approach than ANN, making them more suitable for implementation on mobile devices by reducing computational costs; (2) CNN functions as automatic feature extractors, thereby minimizing the need for handcrafted features which can introduce human bias if the dataset is not fully understood.

Extensive research has demonstrated the efficacy of CNN in various domains, including sentence modelling [172], heartbeat classification [173], and image classification tailored for mobile devices with significant computational cost reduction [174]. One substantial advantage of using CNN is that CNN can act as a feature extractor or dimension reduction to extract useful low-dimensional features. Figure 6-2(a) illustrates pre-trained CNN to enhance the extracted features' informativeness and representativeness. This mechanism is similar to an autoencoder, which compresses input data into a lower-dimensional representation and reconstructs it using convolutional layers for efficient feature extraction and data reconstruction. It is evident that the raw data in Figure 6-2(a) has a higher dimension compared to the extracted features, especially when the node count in the feature extraction layer is set low. If the CNN achieve high accuracy, it indicates that the extracted features are representative of the raw data.



(a)



(b)

Figure 6-2 Two scenarios to implement feature extractor (a) with the implementation of output for the CNN and (b) without the implementation of output for the CNN

Figure 6-2(b) presents a configuration without the pretraining mechanism for a benchmark to evaluate the effectiveness of the configuration shown in Figure 6-2(a). In both configurations, the final prediction uses a Support Vector Regressor (SVR) [175]. Comprehensive details are attached in section 4.2.2 so that the following details are included in the SVR.

SVR is an extension of the Support Vector Machine (SVM), a supervised learning method widely used for classification and regression tasks. Initially developed in the early 1990s, SVR has become a standard tool in ML, particularly for function estimation in high-dimensional spaces. Given a training dataset $\{(x_i, y_i)\}_{i=1}^{N} \in R^d \times R$, where $x_i \in R^d$ is the input vector with *d* representing the dimension of the input vector, and $y_i \in R$ is the label.

SVR aims to find a function *f(x)* that approximates the label values with an error less than ε and remains the function is as flat as possible, which means the error larger than ε is neglectable. The function *f(x)* is defined as:

$$f(x) = \langle w, x_i \rangle + b \qquad \text{Equation 43}$$

Where $w \in R^d$ and $b \in R$ are parameters to be learned, and $< w, x >$ is the dot production of *w* and *x*.

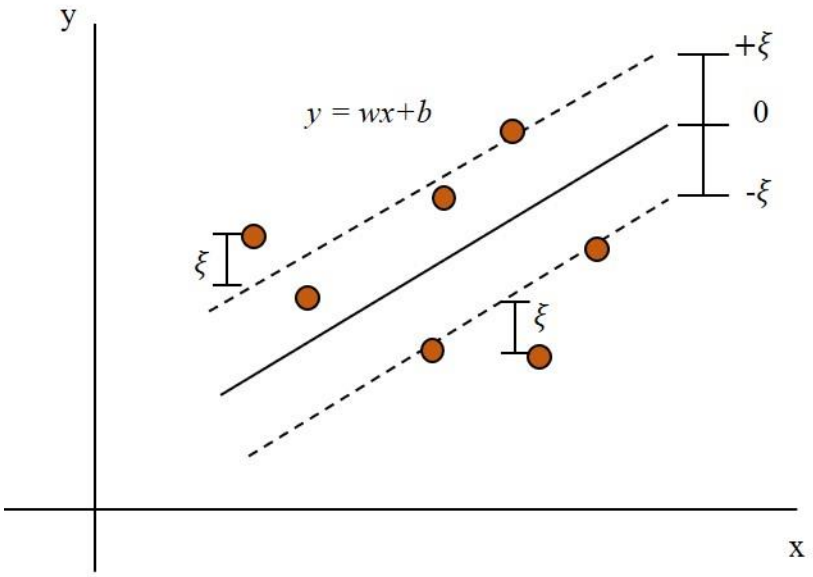


Figure 6-3 Regression problem using Support vector regression

The flatness of *f(x)* is ensured by minimising the norm of the weight vector *w*, leading to the following convex optimisation problem:

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \qquad \text{Equation 44}$$

Subject to the constraints:

$$y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \quad \text{Equation 45}$$

$$\langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \quad \text{Equation 46}$$

where ξ_i, ξ_i^* ≥0 are slack variables introduced to handle infeasible constraints and allow for some errors in the predictions.

6.3 Results and Discussions

In this section, we present the results produced by four models (CNN, SVR, pre-trained CNN + SVR, and non-pre-trained CNN + SVR) based on two types of datasets (time domain and frequency domain).

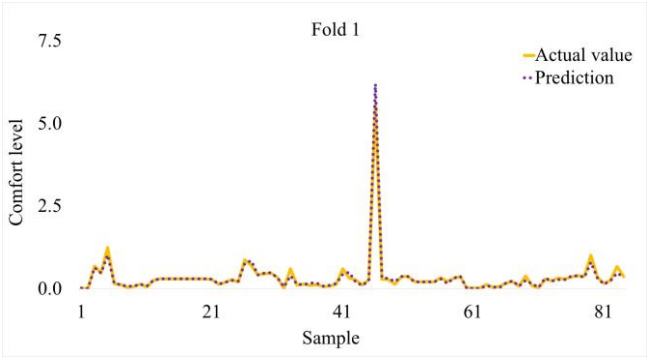
Table 11 Performance for the four models

	R^2		**MAE**		**RMSE**	
Data domain	**Time**	**Frequency**	**Time**	**Frequency**	**Time**	**Frequency**
CNN	0.706 ± 0.129	0.592 ± 0.161	0.07 ± 0.02	0.083 ± 0.031	0.177 ± 0.123	0.222 ± 0.185
SVR	0.365 ± 0.134	N/A	0.119± 0.04	N/A	0.268 ± 0.216	N/A
Non-pre-train CNN+SVR	0.51 ± 0.202	0.463 ± 00.14	0.099 ± 0.03	0.112 ± 0.036	0.250 ± 0.223	0.250 ± 0.226
Pre-train CNN+SVR	0.925 ± 0.043	0.988 ± 0.004	0.036 ± 0.007	0.02 ± 0.005	0.079 ± 0.047	0.028 ± 0.015

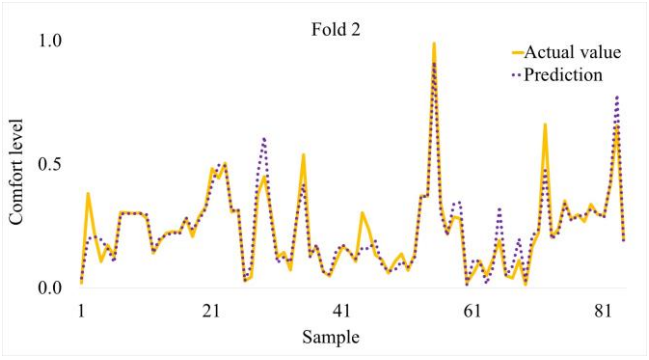
Table 11 shows the performance of the four models with 10-fold cross-validation using data from both the time and frequency domains. The results, calculated as the mean and standard deviation in ten folds (in the form of μ ± σ), indicate the overall performance comparison.

The proposed model demonstrates superior performance and robust stability. It achieves average MAE and RMSE values of 0.02 and 0.028, respectively, indicating minimal prediction errors. Detailed results at the sample level are provided in Figure 6-4. Most actual values fall within the 0 to 1 range, suggesting that the tested train line section maintains an excellent comfort level, as referenced in Table 7. Notably, this study employs a regression model to predict the exact comfort level value rather than the comfort level interval outlined in Table 7.

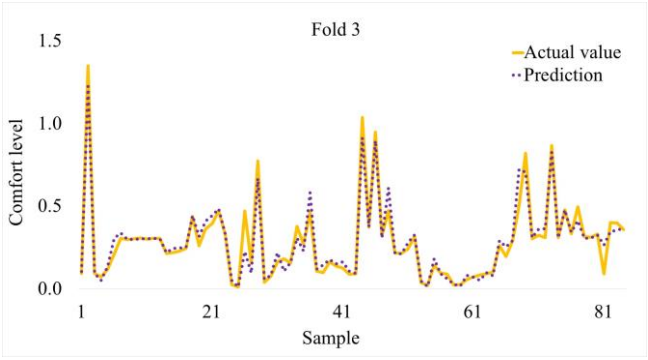
Analysing Figure 6-4(i)-fold 9 reveals that predictions at two peaks show a noticeable gap between the actual and predicted values; however, both the actual values and predictions remain within the same comfort interval. This indicates that the model's performance could be improved by transforming the dataset labels to reflect comfort level intervals.



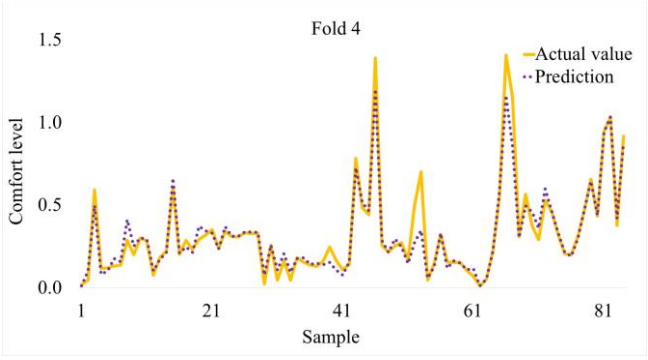
(a)



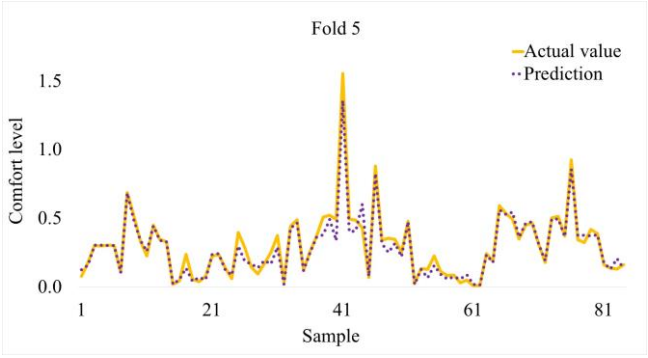
(b)


Fold 3
Actual value
Prediction
Comfort level
1.5
1.0
0.5
0.0
1
21
41
61
81
Sample

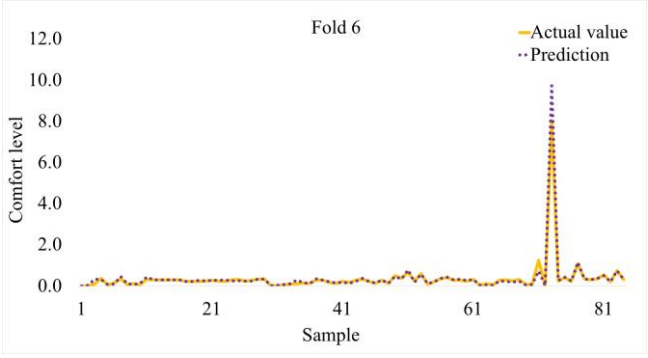
(c)


Fold 4
Actual value
Prediction
Comfort level
1.5
1.0
0.5
0.0
1
21
41
61
81
Sample

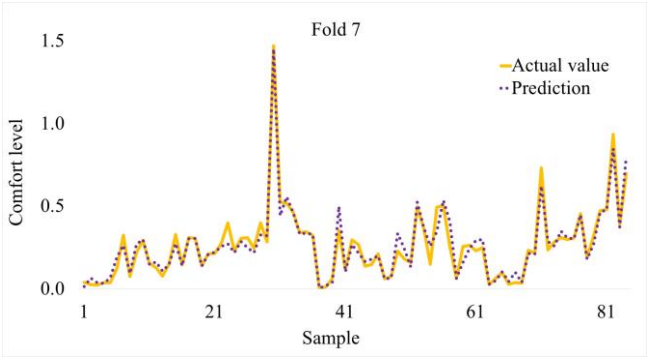
(d)


Fold 5
Actual value
Prediction
Comfort level
1.5
1.0
0.5
0.0
1
21
41
61
81
Sample

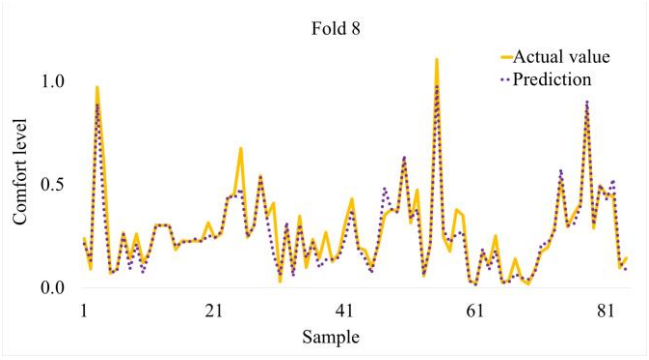
(e)


Fold 6
Actual value
Prediction
Comfort level
12.0
10.0
8.0
6.0
4.0
2.0
0.0
1
21
41
61
81
Sample

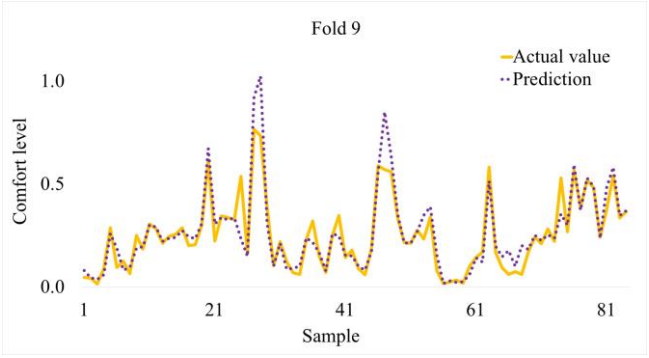
(f)



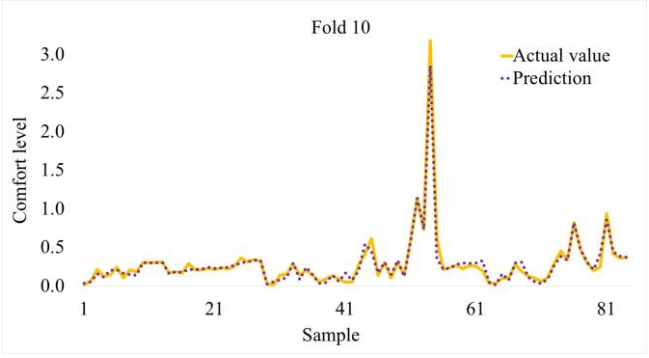
(g)



(h)



(i)



(j)

Figure 6-4 The actual values and predictions correspond to samples in ten folds. (a)-(j) are fold 1 to fold 10, respectively

The pre-trained CNN+SVR model consistently outperforms the alternatives, achieving R^2 values of 0.925 and 0.988 for the time and frequency domains, respectively. This superior performance can be attributed mainly to the pre-training process, which enables extracting more informative features. Conversely, the performance of the CNN model, with an R^2 of 0.706 in the time domain, demonstrates the model's competence in feature extraction. However, its predictive capacity is limited without a more specialised regression mechanism such as SVR. With an R^2 of 0.365, SVR alone struggles to generalise effectively, highlighting its dependence on robust feature extraction.

The remarkable improvement in performance when pre-trained CNN is coupled with SVR (shown in Figure 6-2(a)), as compared to the non-pre-trained variant (shown in Figure 6-2(b)), underscores the CNN's feature extraction capability. Pre-trained models are optimised to extract informative features, substantially enhancing their predictive capability.

It is worth discussing the performance across different data domains. Whilst most models exhibit more substantial results with time-domain data, the pre-trained CNN+SVR model performs well in both time and frequency domains, which shows its robustness. This contrast may suggest that frequency-domain data present greater complexity, potentially requiring more sophisticated pre-processing or feature extraction techniques to achieve high predictive accuracy.

Additionally, while the pre-trained CNN+SVR model demonstrates good performance overall, the analysis reveals specific errors, particularly in predicting peak values. These deviations, although not significant in altering the comfort interval classification, suggest areas for improvement in the model's accuracy. One potential approach is the

transformation of dataset labels to reflect comfort level intervals, which could allow for more accurate predictions in regions. This refinement would likely enhance the model's performance, particularly in complex regions of the dataset where predictive gaps occur.

Furthermore, the comparison of regression layers within the models provides insights into their varying effectiveness. The study emphasises the superior generalisation ability of SVR when used with a pre-trained feature extractor, as opposed to the dense layers traditionally employed in CNN. SVR is renowned for its robustness in handling non-linear data and its capacity for generalisation [175], which likely contributes to its enhanced performance. In contrast, while adequate for classification tasks, the dense layers within CNN may lack the specificity required for regression-based problems.

6.4 Conclusions

This chapter aims to measure train RC levels using a phone-based ML model. An ensemble model combining a CNN and an SVR is adopted. The pre-trained CNN extracts informative features, and the SVR uses these features to predict the RC. The ensemble model demonstrates superior performance compared to its individual components, with the pre-trained feature extractor significantly outperforming the non-pre-trained one.

These findings are particularly relevant to railway companies focused on enhancing the RC. This study offers passengers a simple method to quantify comfort levels and provides engineers with a straightforward way to calculate the comfort index, which is traditionally more complex. More importantly, this chapter demonstrates that ML models can effectively use data collected by a smartphone to estimate the RC. It highlights the potential of utilising multiple smartphones to measure the overall RC for an entire vehicle

rather than a single point. The next chapter expands on this by using crowd-sensing, where multiple smartphones collect data from different locations in the train. This gives a more complete picture of RC across the whole vehicle. It also introduces new methods, like GAT, to improve the accuracy of the predictions and make the comfort estimation process more practical for real-time use.

Chapter 7

Overall Train Ride Comfort Using Crowd-Sensing Data

7.1 Introduction to Crowd-Sensing for Comfort Assessment

Adhering fully to ISO 2631 is both challenging and uncommon among researchers due to its inherent complexity and the number of sensors required for comprehensive evaluation. The motion of trains is intricate and cannot always be accurately analysed as a rigid body, as specified by ISO 2631-4, which focuses on passenger and crew comfort in fixed-guideway transport systems. This standard requires three measuring points—both ends and the middle of the vehicle. When both standing and seated passengers are considered, the number of measurement locations increases from three to twelve, given that vibrations for a seated passenger originate from three sources (feet, seat surface, and seat back).

The considerable complexity and expense place researchers in a challenging position, requiring a balance between precision and budget constraints. In the field of data collection for RC analysis, methodologies vary significantly in their complexity and resource requirements. The extensive effort needed to install and manage long cables and multiple sensors reduces mobility and presents significant challenges for real-time evaluation. Consequently, many studies adopt a simplified approach, using fewer pieces of equipment to measure acceleration at limited locations. For example, [35] employed a single accelerometer for data collection.

In contrast, a more detailed study [76] employed three triaxial accelerometers on a dummy, six triaxial accelerometers on a train, and two seat cushion accelerometers, facilitating a more comprehensive acquisition of acceleration data. Procuring, maintaining, and calibrating such specialised sensors entail a substantial financial investment, particularly for large-scale studies. Moreover, the installation process for these advanced systems is time-consuming and requires high technical expertise. This complexity is further compounded by the need for extensive cabling, which impedes mobility and flexibility, especially when measurements are needed at multiple or varying locations. Additionally, managing and processing the extensive data generated by these numerous sensors is a significant challenge. The complexity of data management increases the risk of equipment malfunctions, potentially leading to compromised data integrity.

The proposed crowdsensing paradigm, which involves five smartphones, can address the issues above. These phones are spatially distributed throughout the vehicle to capture vibrations. The data collected can then be saved locally or transmitted wirelessly to a central hub for further processing. This method provides significant advantages in

terms of mobility and flexibility. Given the complex dynamics of vibration, this chapter introduces the first use of GAT in this domain. This thesis pioneers the integration of crowd-sourced data in assessing the RC, representing a novel endeavour in this field. Furthermore, this work innovatively employs GAT to model the correlation of the RC across different locations by utilising spatial distances. To demonstrate the superior effectiveness of GAT, their performance is compared against a range of traditional ML models.

7.2 Methodologies

7.2.1 Train Dynamic

This section elaborates on the theory of train motion, which provides a firm foundation for developing ML models in later stages. Chapter 5 strives to use data-driven methods to identify the train motion without diving into the principle of the train motion, while this subsection aims to study and understand the mathematical model of train dynamic to provide a theoretical foundation that the acceleration at different locations of a train is interacted and correlated. These dynamics are complex and vary across different points on the train due to differences in structural characteristics and external influences. A comprehensive understanding of these principles is vital for developing accurate and robust predictive models. It provides insights into feature selection, data pre-processing, and the interpretation of model outputs, allowing the models to account for spatial and temporal dependencies within the data. Additionally, integrating train motion theory with ML enhances model interpretability, providing a scientifically grounded framework that supports the validation of predictive capabilities.

The theory of train motion is fundamentally based on Newton's second law of motion [136], which is the foundation for formulating equations that describe the dynamics of a train moving along a track. According to this principle, the motion of a train can be represented by a differential equation that accounts for the various forces acting upon it. The equation is defined as follows:

$$mp\frac{d_v}{d_t} = f(v) - r(v) - g(s) \qquad \text{Equation 47}$$

where *m* is the train mass, p represents the rotating mass factor accounting for the inertia of the train's rotating parts, *v* denotes the train speed, *t* signifies time, *f(v)* corresponds to tractive and braking efforts, *r(v)* is the running resistance, and *g(s)* reflects the effects of track geometry on the train's dynamics. The running resistance *r(v)* is determined by the Davis equation [176] in Equation 48:

$$r(v) = r_0 + r_1 v + T_f r_2 v^2 \qquad \text{Equation 48}$$

Here, r_0, r_1, and r_2 are the parameters for motion resistance and T_f accounts for additional resistance when trains run through tunnels. The resistance due to track geometry *g(s)*, described in Equation 49, considers gradients and curves:

$$g(s) = mgsin\big(a(s)\big) + m\frac{k}{R(s)} \qquad \text{Equation 49}$$

In this equation, *g* denotes gravitational acceleration, *a(s)* represents the angle of the track at location *s*, and *R(s)* is the curve radius. Factors such as curve radius, superelevation, train speed and length, wheel condition and design, rail maintenance, and lubrication affect the curve resistance [161]. Although the precise effect of each factor

on motion resistance is not fully understood [177], the curve resistance is modelled as an equivalent gradient, producing similar motion resistance with a hyperbolic dependence on the curve radius *R(s)*. This means that as the radius of the curve decreases (tighter curves), the resistance increases significantly. The proportional constant *k* indicates that tighter curves generate higher resistances.

7.2.2 Data Collection from Multiple Devices

Data collection is conducted on a Class 323 train operating along the corridor from Selly Oak to Redditch. Five calibrated smartphones are installed at different points within a carriage to represent diverse passenger experiences. These devices are placed either on the floor or the seats, all oriented with the X-axis aligned in the direction of the train's movement. To minimise the impact of extraneous movements, all devices are securely mounted using blue tack. Figure 7-1 depicts the placement of one phone in the middle of the carriage and two phones at each end. Each smartphone records triaxial acceleration data at a maximum sampling rate of 100 Hz [136].

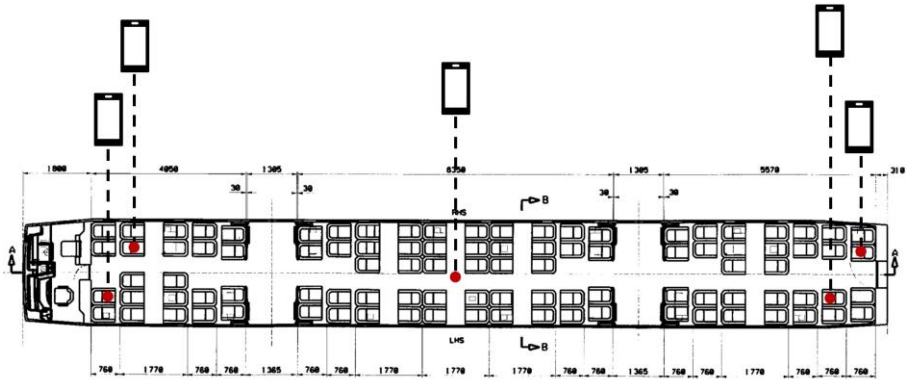


Figure 7-1 Phone distribution [178]

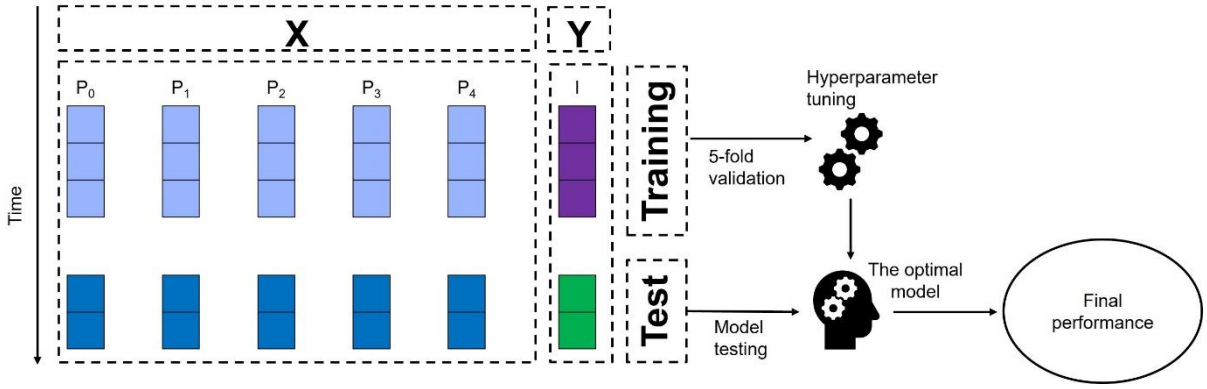
Data collection spanned four train journeys to capture variability in ride conditions, including occupancy, speed, and track quality. It is important to note that the recorded

data may exhibit variations due to the accelerometers' sensitivity to temperature and potential degradation over time, even though the phones remain in the same locations [176]. This study utilises a range of mobile phone models, including two iPhone X, one iPhone 6, and two iPhone 8. The diversity of phone models enhances the model's generalisability and ability to tolerate noise, and it is impractical to assume all passengers will be using identical phones.

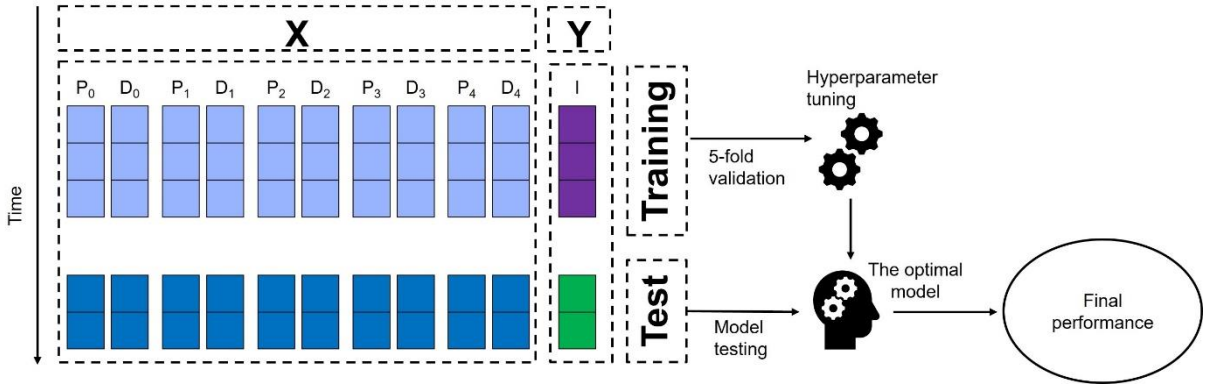
7.2.3 Data Preparation

Conducting quantitative experiments to gauge generalisation error and assess model performance is a standard practice in ML [177]. To ensure the robustness and reliability of the model, it is necessary to quantify its uncertainty. The KCV scheme in Figure 3-16 is an effective method for estimating model errors, with K commonly set to 5 or 10 to balance bias and computational cost [179]. The K value here is set to 5 to save computational resources. The advantage of this method is that all observations are used for training and validation, with each subset being used for validation exactly once. This technique reduces model variance, providing a comprehensive insight into the model's performance and avoiding biased results due to specific data splits [180].

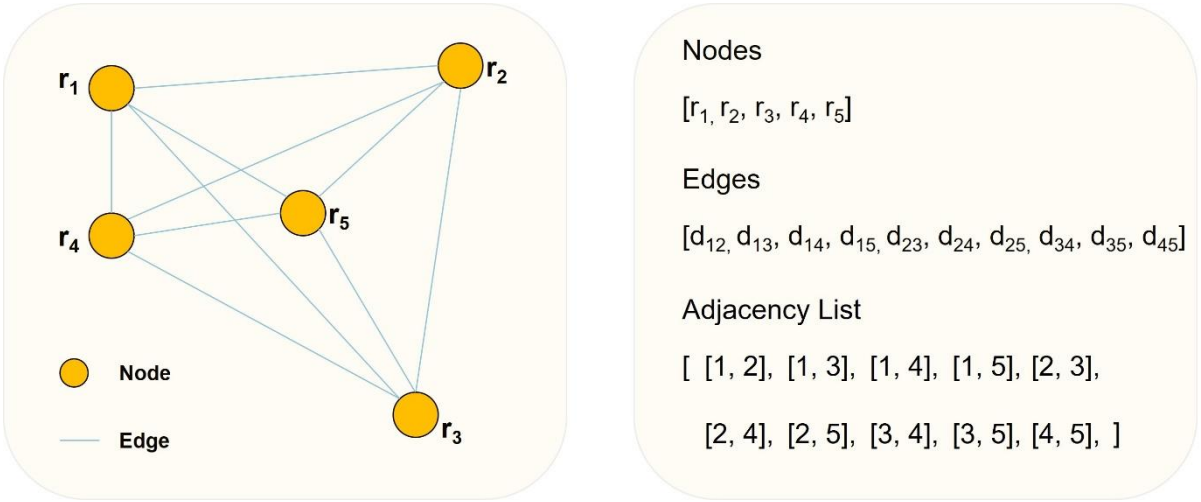
It is important to note that the dataset is thoughtfully partitioned to withhold the latter 10% of the data using a time-series split, as shown in the 'Test set' in Figure 7-2(a)(b). This reserved portion is crucial for the final testing of model performance, and its segregation prevents inadvertent information leakage during the model development process. Furthermore, the optimal model derived from the training set undergoes rigorous evaluation using the hold-out test set to enhance the reliability of the model's performance.



(a)



(b)



(c)

Figure 7-2 Data flowing (a) data without distance features (b) data concatenated with distance features (c) graph data.

Figure 7-2(c) (left side) presents an undirected graph comprising five nodes labelled r_1 to r_5. These nodes are interconnected by edges that represent the distances between them. Figure 7-2(c) (right side) depicts the graph using arrays. The 'Nodes' array lists the ride index calculated from the corresponding phone, while the 'Edges' array specifies the distance between nodes. For example, d_{12} denotes the distance between node r_1 and r_2, and similarly for the other edges. The 'Adjacency List' defines the edges in a format compatible with computational processing, with each sub-list indicating which nodes are connected. For instance, '[1, 2]' signifies that node r_1 is connected to node r_2.

7.2.4 Model Development for Crowd-Sensed Data

7.2.4.1 Random Forest

Since its introduction by Breiman in 2001, the RF algorithm has achieved significant success in ML [181]. This ensemble model comprises multiple DT estimators (illustrated in Figure 7-3) and is utilised for both classification and regression tasks. To ensure diversity among the trees, each DT is trained on a sub-dataset generated through bootstrap sampling [182]. For classification problems, the final prediction is made by majority voting among the trees, while for regression problems, the results are averaged to produce the final prediction.

Using different data subsets for training each tree helps mitigate the risk of overfitting [181]. Additionally, overfitting can be further controlled by adjusting hyperparameters such as the maximum depth of the trees and the minimum number of samples required at leaf nodes. The ensemble nature of the RF also confers high robustness to noisy data, as it can smooth out the effects of outliers that do not conform to the central data patterns.

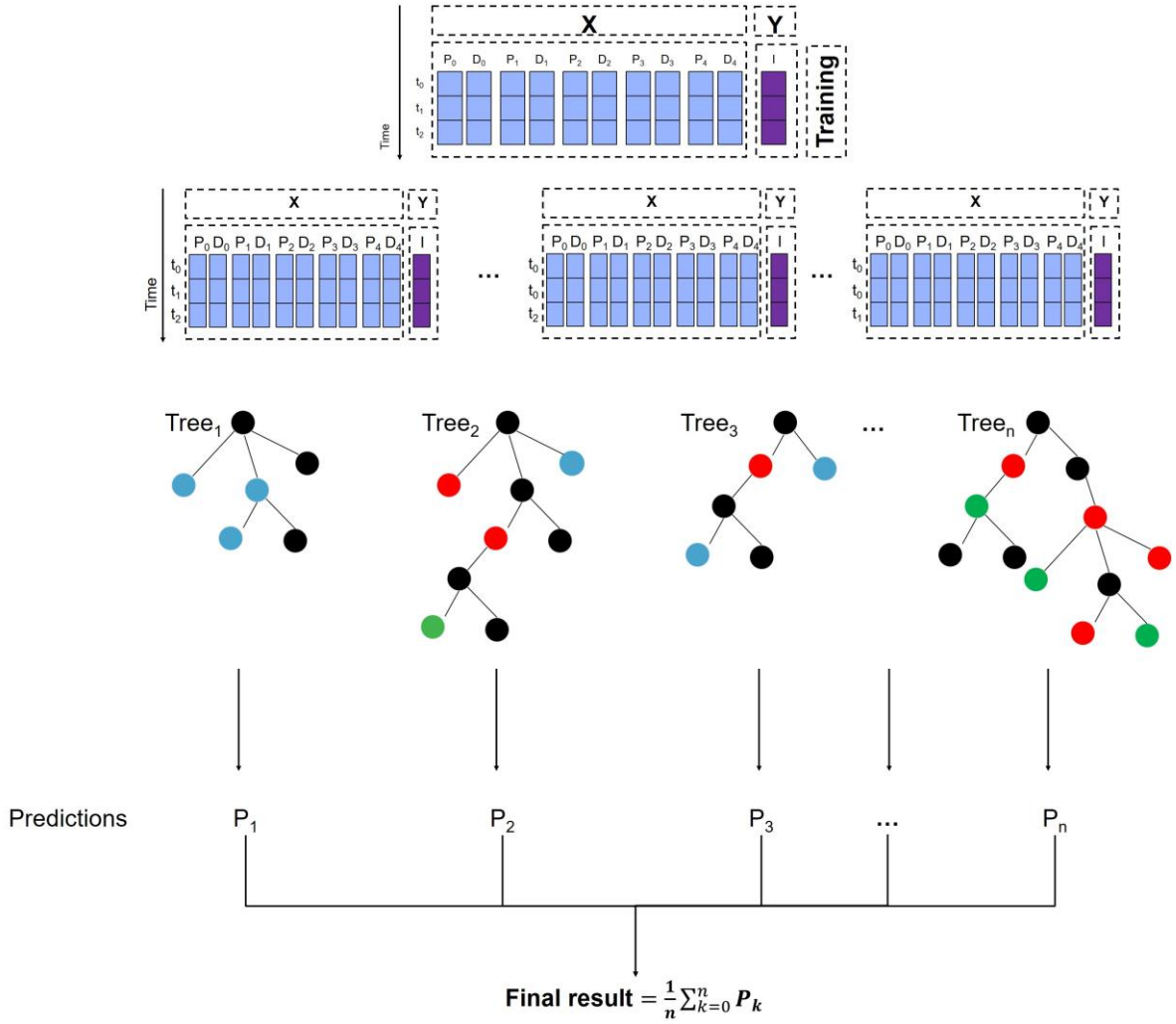


Figure 7-3 Random Forest

Another critical feature of RF is its implicit ability to perform feature selection. Since each tree is trained on randomly selected subsets of data and attributes, not all features are used at every split. This process identifies the most informative attributes, thereby reducing the influence of less essential features on the final predictions and minimising the need for extensive feature selection efforts. The RF algorithm is a powerful and versatile ML method well-suited for handling diverse data types and addressing various problems, including classification, regression, and feature selection. Its robustness and high performance make it popular for many practical applications.

7.2.4.2 Extreme Gradient Boosting

Tree boosting is a highly effective and widely utilised technique in ML. Chen and Guestrin [183] introduced an expandable end-to-end tree-boosting system called XGBoost, widely adopted by data scientists for achieving state-of-the-art results in numerous ML tasks. They developed an innovative algorithm that efficiently handles sparse data and introduced a weighted quantile sketch technique for approximate tree learning. Additionally, they explored cache access patterns, data compression strategies, and data sharding, demonstrating that XGBoost scales effectively to handle datasets with billions of examples while using significantly fewer computing resources than existing systems.

XGBoost shares a common feature with RF: it is an ensemble algorithm that combines predictions from multiple individual models, typically DTs, to deliver more accurate and robust final predictions. However, the methods by which these results are combined, and the construction of the trees differ between the two algorithms. Unlike RF, which averages the predictions from each tree, XGBoost builds DTs sequentially (as shown in Figure 7-4), with each new tree aiming to correct the errors made by the previous ones. Trees are created one at a time, with the residuals—the differences between the actual values and the predictions made by previous trees—being passed to the newly created tree.

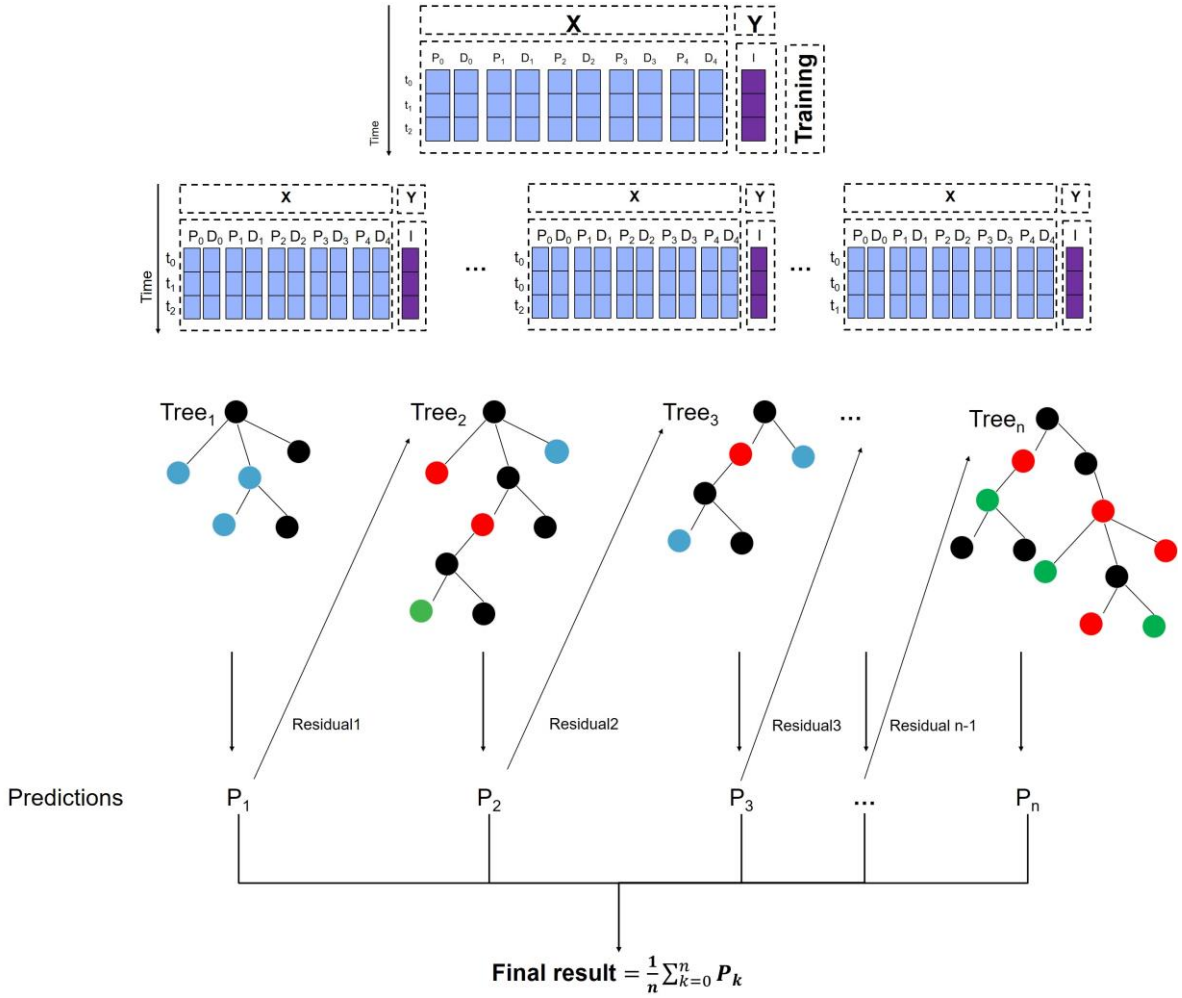


Figure 7-4 Extreme Gradient Boosting

7.2.4.3 K-nearest Neighbour

The K-Nearest Neighbours (KNN) algorithm is one of the most fundamental and straightforward methods in ML, often used as a preliminary approach when there is little or no prior information about the data [184]. The core of KNN involves calculating the Euclidean distance between samples using their features, as defined below:

$$d(x_i, x_n) = \sqrt{\sum_{\substack{i=0 \\ n=0}}^{\substack{i \\ n}} \left(x_{ip-} x_{np}\right)^2} \qquad \text{Equation 50}$$

Where x_i refers to a sample; x_{ip} represents the feature of x_i; *p* is the number of features; *i* and *n* are the number of samples.

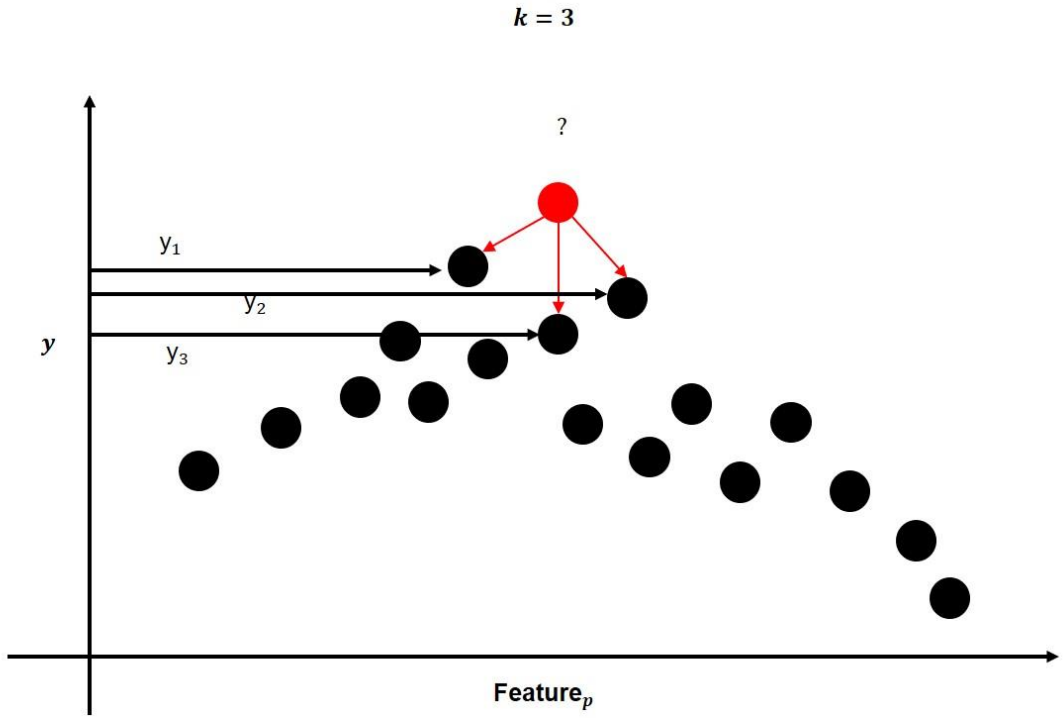


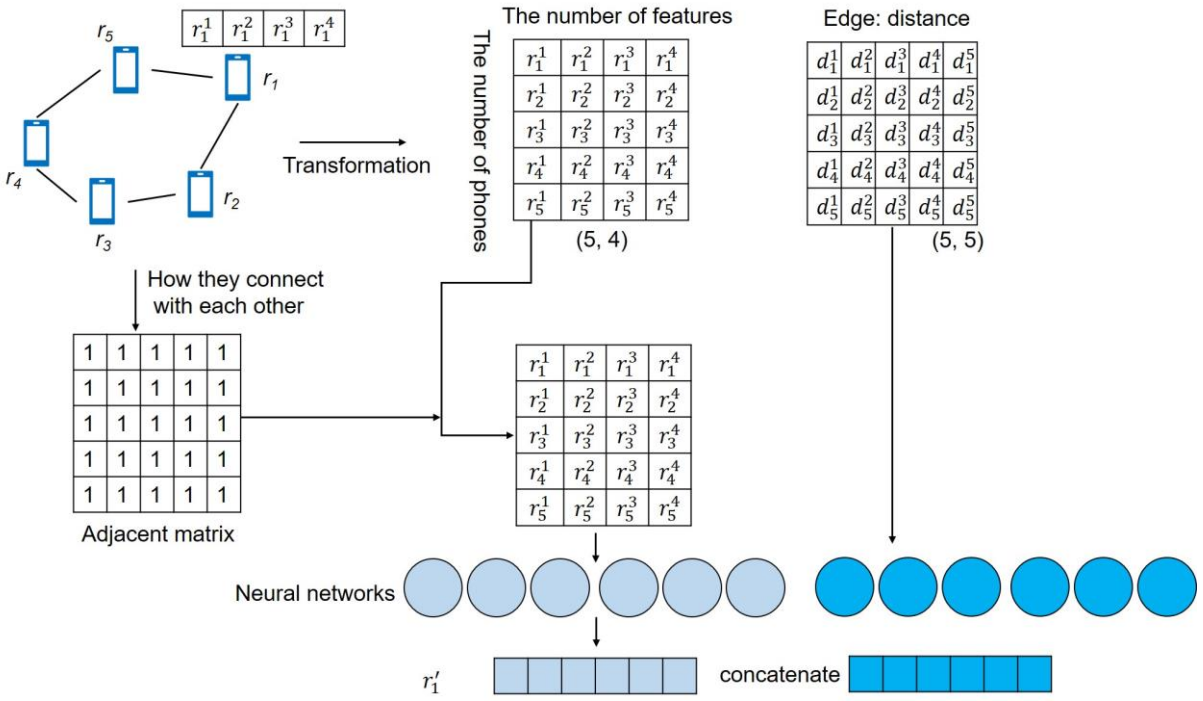
Figure 7-5 K-nearest neighbour

Figure 7-5 illustrates the KNN algorithm using three neighbours for prediction. The figure shows a sample to be predicted in red. In this case, the nearest three neighbours, y_1, y_2, and y_3, each have a known *y* value. The prediction for the unknown sample's *y* value is determined by averaging the *y* values of these three nearest neighbours (y_1, y_2, and y_3). This figure also emphasises the importance of selecting the hyperparameter "k" in KNN, as detailed in [185]. The choice of "k" directly impacts the model's performance. A small "k" value is more likely to lead to overfitting, where the model is overly complex and may not generalise well to new data. Conversely, an enormous "k" value tends to result in underfitting, where the model oversimplifies the data and fails to capture the underlying patterns.

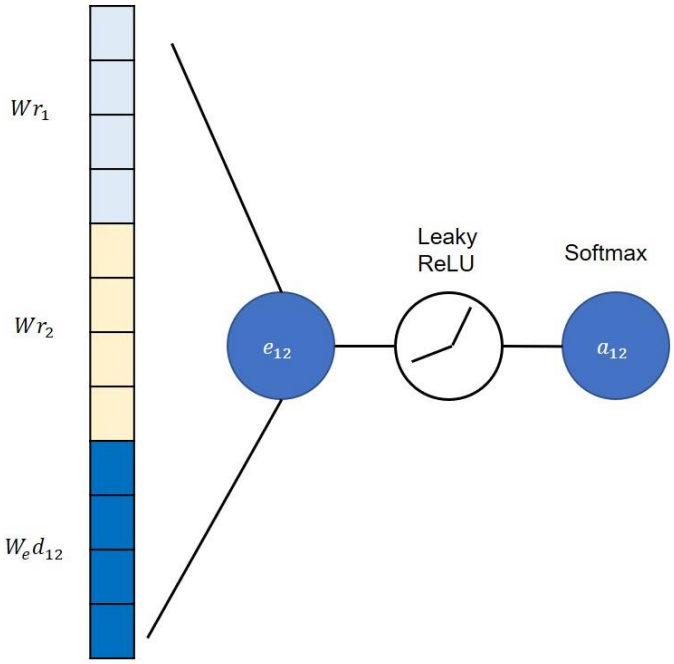
7.2.4.4 Graph Attention Network

GAT have gained increasing recognition for their proficiency in handling non-Euclidean data [186], presenting a significant advantage over traditional CNN that simplifies data into grid matrices. The design of GAT makes them particularly suitable for crowd-sensing applications. Onboard a train, for instance, vibrations measured at different points are interconnected, and the attention mechanism of GAT is ideal for evaluating the influence between each pair of measurements.

The superior capability of GAT to preserve data topology is seen in Figure 7-6, which depicts a single hidden layer GAT model proficient in mapping the complex topological interrelations among various smartphones. This model constructs a graph consisting of nodes, an adjacency list, and edges, processed through an input layer followed by a hidden layer that performs graph attention convolution. To validate the model's extensive predictive power, the nodes designated for prediction are varied individually, emphasising the model's robustness and potential applications in complex data structures. This implies that each smartphone is a predicting point in turn by the remaining four phones.



(a)



(b)

Figure 7-6 Graph attention network

In Figure 7-6(a), it is observed that the nodes (representing the comfort index calculated) and the edges (distances between each pair of phones) are transformed into vectors

before being fed into two separate neural networks with learnable weights. After passing through these neural networks, the attention mechanism is applied, as illustrated in Figure 7-6(b), using 'Phone 1' as an example. Notably, the distance between 'Phone 1' and 'Phone 2' is included in the calculation of the attention coefficient. This coefficient is subsequently processed by an activation function followed by the SoftMax function, which normalises the attention scores.

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{j \in N_i} \exp(e_{ij})} \quad \text{Equation 51}$$

$$e_{ij} = LeaklyReLU\left(\vec{a}(Wr_i || Wr_j || W_e d_{ij})\right) \quad \text{Equation 52}$$

$$r_i' = \sigma\left(\sum_{j \in N_i} a_{ij}(Wr_i + W_e d_{ij})\right) \quad \text{Equation 53}$$

Where:

6. a_{ij} is the normalized attention coefficient between r_i and its neighbour r_j.
7. e_{ij} is the attention coefficient before normalisation.
8. r_i and r_j are the comfort indices computed from 'Phone i' and 'Phone j'.
9. W is the learnable weight matrix for the node.
10. W_e is the learnable weight matrix for the edge.
11. d_{ij} is the distance between 'Phone i' and 'Phone j'.

The technical exposition culminates with a series of equations describing GAT's operational framework. These equations, from Equation 51 to Equation 53 [187, 188], break down the steps in calculating the attention coefficients and updating nodes via

graph convolutions. GAT extend traditional graph neural networks by incorporating an attention mechanism, which allows the network to assign varying importance to different neighbours of a node.

The attention coefficients a_{ij} are computed by normalising the raw attention scores e_{ij} using a SoftMax function [189], as shown in Equation 51. These raw attention scores are calculated via a LeakyReLU nonlinearity [190] applied to a shared attention mechanism $\vec{a}$, which combines the linear transformations of the feature vectors of nodes *i* and *j*, as well as the edge features d_{ij}, using weight matrices *W* and W_e, respectively, as depicted in Equation 52. The final node embeddings r_i' are then obtained by aggregating the transformed neighbour features weighted by their corresponding attention coefficients, followed by an activation function σ such as ReLU [191], as specified in Equation 53.

7.3 Results and Discussions

7.3.1 Data Collection Overview

Ensuring data integrity plays a vital role in ML, particularly in DL, where large datasets are crucial for effective training [192]. This study emphasises the importance of maintaining high data integrity and robustness to guarantee reliable findings.

Data collection occurs across various journeys under differing conditions to build a robust and generalisable model. Ensuring the diversity of the dataset is vital to reflect real-world scenarios accurately. Factors such as the direction of train travel, passenger load (crowdedness), operational speed, and different carriages are considered to enhance the model's robustness to cover a broad spectrum of train operation conditions. Varying distances between phones are included to simulate different passenger distributions,

demonstrating that the GATs can operate effectively even when phones are placed at variable distances.

Table 12 details the four datasets collected between Selly Oak and Redditch stations during peak and off-peak times to encompass different levels of crowdedness. The distances are represented by d_{ij}, indicating the distance between one phone and the others. Sample numbers vary due to differing train dwell times as the train waits for faster trains to pass. RC is calculated every 6 seconds, with a 3-second overlap. The overall journey dataset is a concatenation of the four individual datasets.

Table 12 Data collection

Dataset	**Direction**	**Crowdedness**	**Distance (d_{ij})**	**Window size (second)**	**Overlapped (second)**	**The number of samples**
1	Selly Oak to Redditch	Peak time	5, 5 5, 5 0.1, 10 10, 10 10, 0.1	6	3	327
2	Redditch to Selly Oak	Peak time	7, 7 7, 7 0.1, 14 14, 14 14, 0.1	6	3	402
3	Selly Oak to Redditch	Off-peak	10, 10 10, 10 0.1, 20 20, 20 20, 0.1	6	3	573

4	Redditch to Selly Oak	Off-peak	1, 1 1, 1 0.1, 2 2, 2 2, 0.1	6	3	522
Overall	Selly Oak to Redditch and return	Peak and off-peak time	The gathering of the above	6	3	1824

It is worth mentioning that the training set for the overall journey is composed of all the training sets from the four journeys, and similarly, the test set for the overall journey is composed of all test sets from these journeys to prevent information leakage. Both individual journey datasets and the combined dataset are used to assess model performance. The graph remains static during data collection for a single journey since the distances between phones are constant. However, passenger boarding and alighting at each station cause the graph to change, allowing evaluation of the model's performance in a dynamic context using the 'Overall' dataset.

Figure 7-7(a) illustrates the spatial arrangement of the smartphones. 'Phone 1' is centrally located on the floor, while 'Phone 2' and 'Phone 3' are positioned at one end, and 'Phone 4' and 'Phone 5' at the other, each maintaining a 10-meter distance from 'Phone 1'. Adjacent pairs are noted between 'Phone 2' and 'Phone 3', as well as 'Phone 4' and 'Phone 5', separated by a gap of 0.1 meters. Figure 7-7(b) presents a time-series analysis of tri-axial acceleration data collected from five smartphones positioned at various locations on a train. Each phone's dataset, labelled from 'Phone 1' to 'Phone 5', is graphically displayed across three sub-panels representing the X, Y, and Z axes. The

measurements, recorded in meters per second squared (m/s^2), are plotted over time in seconds. The tri-axial acceleration data show significant fluctuations along the Z-axis, more fluctuated than in the lateral (Y-axis) and longitudinal (X-axis) directions, which likely indicates a vertical response to track irregularities and bumps [193]. The X-axis data, aligned with the direction of train travel, primarily captures the train's acceleration and deceleration patterns [194]. In contrast, the Y-axis data suggests lateral sway movements, reflecting the train's side-to-side motion and stability. Additionally, Figure 7-7(b) includes a spectral analysis of the acceleration data across multiple journeys, highlighting that the most significant frequency components are below 20 Hz. Notably, the energy peaks are predominantly within the ranges of 0–2 Hz for the X-axis, 0–3 Hz for the Y-axis, and 0–5 Hz for the Z-axis, indicating that most train movements occur at these frequency bands.

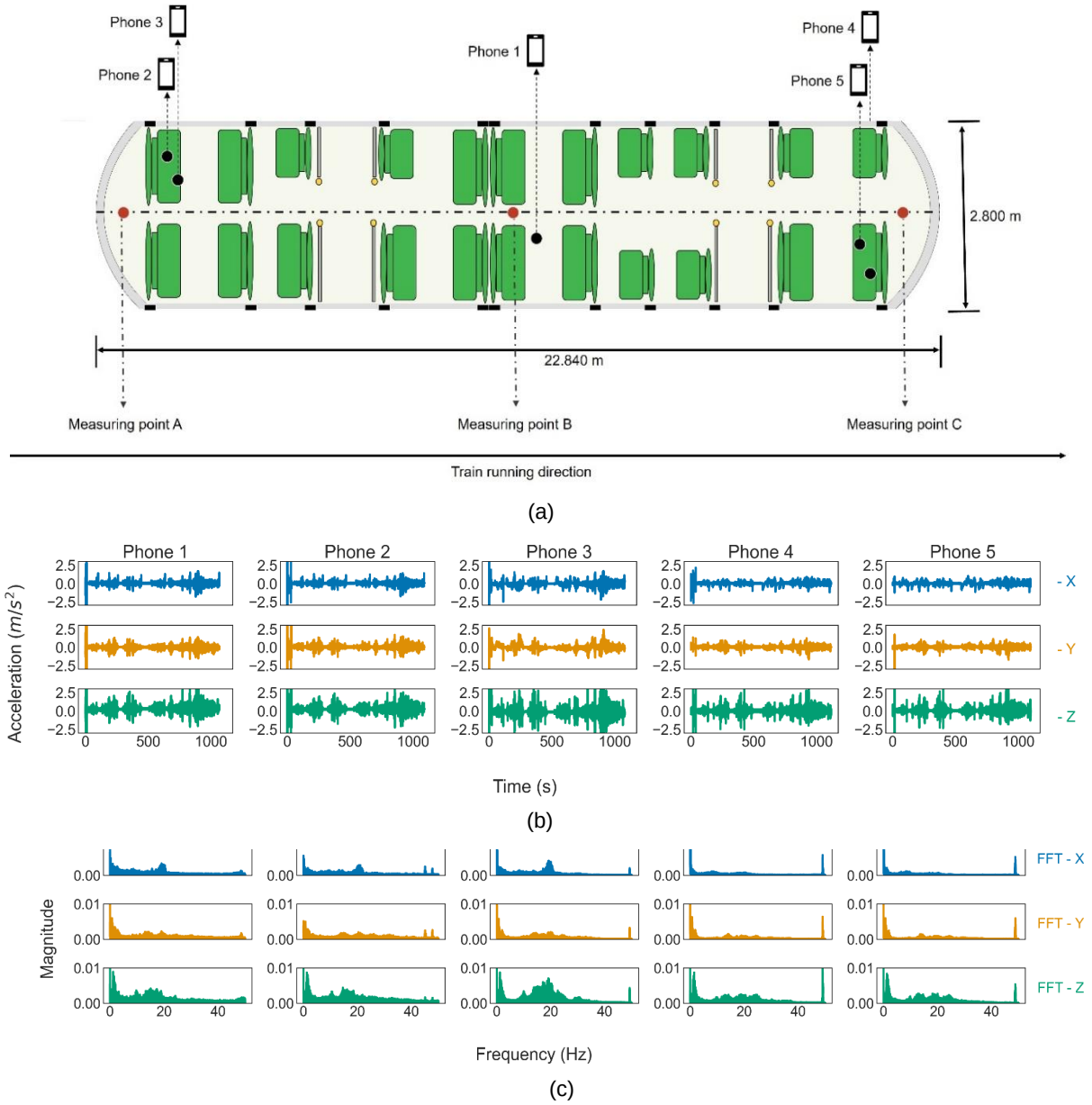


Figure 7-7 Data visualisation (a) the distribution of phones, (b) raw data from the phones in the time domain, (c) raw data from the phones in the frequency domain.

Figure 7-8 unfolds ride comfort data recorded by five smartphones over four journeys. The horizontal axis represents the number of samples, while the vertical axis indicates the comfort level in meters per second squared (m/s²). In 'Journey 1', 'Phone 3' and 'Phone 4' recorded shocks up to 2.5 m/s², reaching the 'very uncomfortable' threshold. Similarly, 'Phone 5' in 'Journey 2' also recorded a shock of up to 2.5 m/s². However, most

measurements from various locations remain below 1 m/s², which denotes 'a little uncomfortable'.

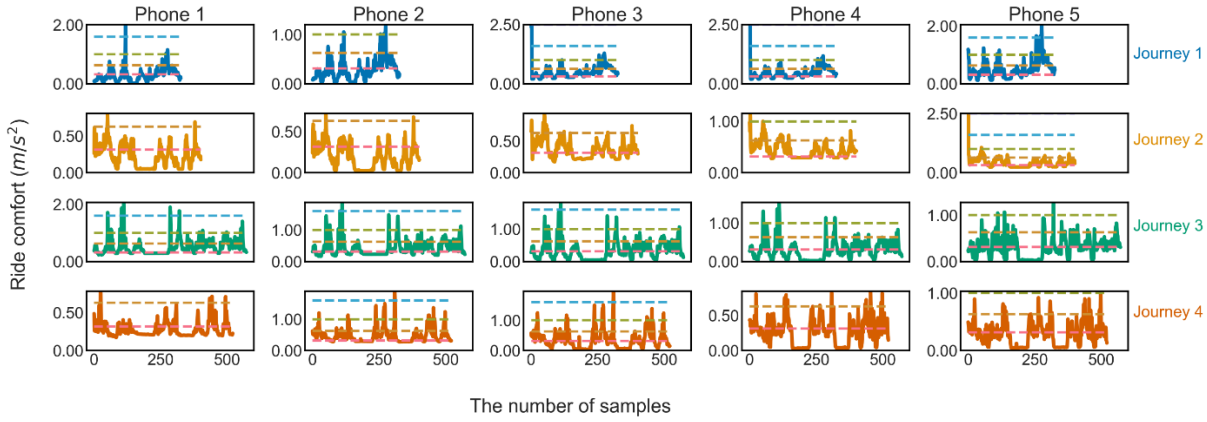


Figure 7-8 Ride comfort calculations. Dash lines divide the comfort level, as shown in Table 6. The dash line indicates values 0.315 m/s2, 0.63 m/s2, 1 m/s2, 1.6 m/s2, and 2.5 m/s2 from the bottom to the top.

The pattern shown in Figure 7-8 reveals clear differences in RC depending on where each coach is positioned within the train. At the front, as seen in Journey 4, passengers experience a smoother ride with fewer vibrations, likely because this section is less affected by the train's movements. In the middle, as represented by Journey 2, there is more noticeable shaking, with greater fluctuations in comfort levels due to the combined forces from the front and rear. Journey 3, from a coach slightly further back, shows moderate comfort with occasional peaks, suggesting it benefits somewhat from being buffered by other carriages. The rear, shown in Journey 1, has the highest and most frequent vibrations, consistent with the increased sway often felt in this section. These findings reflect typical train dynamics: the front tends to be more stable, while the back is prone to more movement, impacting overall comfort.

7.3.2 Evaluation of Model Performance

This section reveals the performance of the proposed system, focusing on both technical accuracy and ethical implications to ensure robustness and reliability. The chapter employs a diverse dataset, cross-validation techniques, and benchmark models to achieve this. The combination of these elements ensures the model's robustness. Rigorous data validation on GAT demonstrates their superior performance in handling complex data structures compared to benchmark models. Cross-validation techniques further confirm the GAT's generalisability and reliability across various scenarios, with data collected from different journeys featuring varied characteristics.

To enhance technical robustness, an extensive hyperparameter tuning process using Optuna is employed. This approach optimises the model by leveraging Optuna's innovative "defined-by-run" strategy, which allows for a dynamically configurable search space. Optuna integrates advanced searching and pruning algorithms to enhance optimisation cost-effectiveness and is highly compatible with diverse models and purposes [195]. In comparative evaluations [195], Optuna outperforms other hyperparameter tuning frameworks such as SMAC [196], Spearmint [197], Hyperopt [198], Autotune [199], Vizier [200], and Tune [201], particularly in ease of implementation, computational cost efficiency, and adaptability to various models.

Social robustness involves creating knowledge and solutions that are credible, relevant, and legitimately generated [202]. This is achieved by engaging diverse stakeholders throughout the development process, from data collection to model application. The proposed system leverages crowd-sourced data to offer a scalable, direct, and cost-effective way for passengers and train companies to estimate real-time ride comfort. The

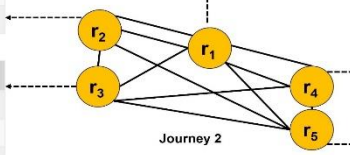
system ensures transparency in data collection by relying solely on smartphone accelerometers.

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_2	0.93 +- 0.1	0.91
ND	RF	r_2	0.97 +- 0.05	0.96
ND	XGB	r_2	0.96 +- 0.06	0.95
D	KNN	r_2	0.93 +- 0.1	0.91
D	RF	r_2	0.96 +- 0.06	0.94
D	XGB	r_2	0.96 +- 0.07	0.93
D	GAT	r_2	0.96 +- 0.07	0.94

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_1	0.94 +- 0.1	0.96
ND	RF	r_1	0.96 +- 0.07	0.98
ND	XGB	r_1	0.96 +- 0.07	0.98
D	KNN	r_1	0.94 +- 0.1	0.96
D	RF	r_1	0.96 +- 0.07	0.98
D	XGB	r_1	0.94 +- 0.1	0.96
D	GAT	r_1	0.93 +- 0.07	0.91

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_4	0.86 +- 0.14	0.87
ND	RF	r_4	0.86 +- 0.15	0.91
ND	XGB	r_4	0.84 +- 0.18	0.87
D	KNN	r_4	0.86 +- 0.17	0.93
D	RF	r_4	0.86 +- 0.15	0.91
D	XGB	r_4	0.86 +- 0.15	0.85
D	GAT	r_4	0.90 +- 0.05	0.91

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_3	0.88 +- 0.09	0.93
ND	RF	r_3	0.88 +- 0.07	0.91
ND	XGB	r_3	0.87 +- 0.1	0.93
D	KNN	r_3	0.88 +- 0.09	0.93
D	RF	r_3	0.88 +- 0.08	0.91
D	XGB	r_3	0.88 +- 0.08	0.92
D	GAT	r_3	0.88 +- 0.04	0.90



Journey 2

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_5	0.77 +- 0.38	-1.26
ND	RF	r_5	0.77 +- 0.38	-3.2
ND	XGB	r_5	0.75 +- 0.37	-0.48
D	KNN	r_5	0.77 +- 0.38	-2.07
D	RF	r_5	0.77 +- 0.38	-1.09
D	XGB	r_5	0.66 +- 0.38	-0.74
D	GAT	r_5	0.84 +- 0.03	0.84

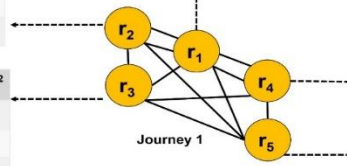
(a)

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_2	0.35 +- 0.5	0.69
ND	RF	r_2	0.57 +- 0.38	0.64
ND	XGB	r_2	0.53 +- 0.36	0.56
D	KNN	r_2	0.35 +- 0.5	0.69
D	RF	r_2	0.55 +- 0.36	0.64
D	XGB	r_2	0.56 +- 0.38	0.38
D	GAT	r_2	0.87 +- 0.21	0.72

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_1	0.44 +- 0.5	0.82
ND	RF	r_1	0.53 +- 0.29	0.8
ND	XGB	r_1	0.47 +- 0.32	0.66
D	KNN	r_1	0.44 +- 0.5	0.82
D	RF	r_1	0.58 +- 0.25	0.82
D	XGB	r_1	0.5 +- 0.31	0.67
D	GAT	r_1	0.87 +- 0.15	0.86

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_4	0.75 +- 0.31	0.86
ND	RF	r_4	0.69 +- 0.36	0.88
ND	XGB	r_4	0.75 +- 0.36	0.86
D	KNN	r_4	0.75 +- 0.31	0.86
D	RF	r_4	0.57 +- 0.35	0.95
D	XGB	r_4	0.45 +- 0.26	-0.23
D	GAT	r_4	0.91 +- 0.03	0.87

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_3	0.36 +- 0.26	0.77
ND	RF	r_3	0.29 +- 0.23	0.5
ND	XGB	r_3	0.23 +- 0.28	0.03
D	KNN	r_3	0.36 +- 0.26	0.77
D	RF	r_3	0.32 +- 0.24	0.32
D	XGB	r_3	0.24 +- 0.2	-0.15
D	GAT	r_3	0.83 +- 0.2	0.81



Journey 1

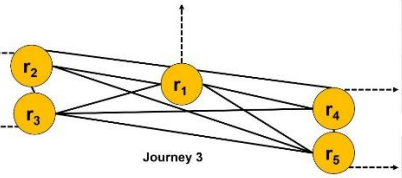
Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_5	0.71 +- 0.25	0.92
ND	RF	r_5	0.83 +- 0.14	0.89
ND	XGB	r_5	0.82 +- 0.1	0.92
D	KNN	r_5	0.71 +- 0.25	0.92
D	RF	r_5	0.83 +- 0.11	0.86
D	XGB	r_5	0.81 +- 0.13	0.9
D	GAT	r_5	0.90 +- 0.08	0.87

(b)

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_1	0.91 +- 0.02	0.86
ND	RF	r_1	0.92 +- 0.02	0.91
ND	XGB	r_1	0.91 +- 0.02	0.89
D	KNN	r_1	0.90+- 0.02	0.86
D	RF	r_1	0.91 +- 0.02	0.88
D	XGB	r_1	0.92 +- 0.02	0.91
D	GAT	r_1	0.93 +- 0.04	0.94

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_2	0.93 +- 0.01	0.90
ND	RF	r_2	0.94 +- 0.01	0.93
ND	XGB	r_2	0.93 +- 0.01	0.91
D	KNN	r_2	0.93 +- 0.01	0.94
D	RF	r_2	0.94 +- 0.01	0.93
D	XGB	r_2	0.93 +- 0.01	0.92
D	GAT	r_2	0.92 +- 0.02	0.94

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_3	0.89 +- 0.01	0.90
ND	RF	r_3	0.89 +- 0.01	0.94
ND	XGB	r_3	0.92 +- 0.03	0.96
D	KNN	r_3	0.94 +- 0.01	0.92
D	RF	r_3	0.91 +- 0.01	0.94
D	XGB	r_3	0.94 +- 0.05	0.91
D	GAT	r_3	0.94 +- 0.01	0.92



Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_4	0.97 +- 0.01	0.94
ND	RF	r_4	0.98 +- 0.01	0.98
ND	XGB	r_4	0.95 +- 0.02	0.95
D	KNN	r_4	0.97 +- 0.01	0.94
D	RF	r_4	0.98 +- 0.01	0.98
D	XGB	r_4	0.93 +- 0.04	0.95
D	GAT	r_4	0.96 +- 0.02	0.95

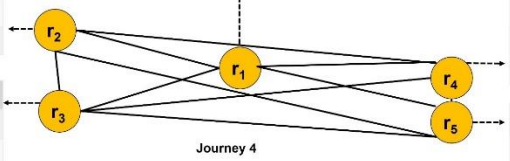
Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_5	0.53 +- 0.28	0.33
ND	RF	r_5	0.55 +- 0.22	0.36
ND	XGB	r_5	0.54 +- 0.21	0.46
D	KNN	r_5	0.53 +- 0.28	0.33
D	RF	r_5	0.54 +- 0.26	0.37
D	XGB	r_5	0.54 +- 0.21	0.36
D	GAT	r_5	0.81 +- 0.03	0.84

(c)

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_1	0.61 +- 0.3	0.77
ND	RF	r_1	0.58 +- 0.38	0.78
ND	XGB	r_1	0.57 +- 0.14	0.65
D	KNN	r_1	0.61 +- 0.3	0.77
D	RF	r_1	0.58 +- 0.4	0.78
D	XGB	r_1	0.6 +- 0.2	0.7
D	GAT	r_1	0.86 +- 0.15	0.81

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_2	0.92 +- 0.01	0.94
ND	RF	r_2	0.93 +- 0.02	0.94
ND	XGB	r_2	0.91 +- 0.03	0.93
D	KNN	r_2	0.92 +- 0.01	0.94
D	RF	r_2	0.92 +- 0.02	0.94
D	XGB	r_2	0.93 +- 0.01	0.94
D	GAT	r_2	0.91 +- 0.02	0.93

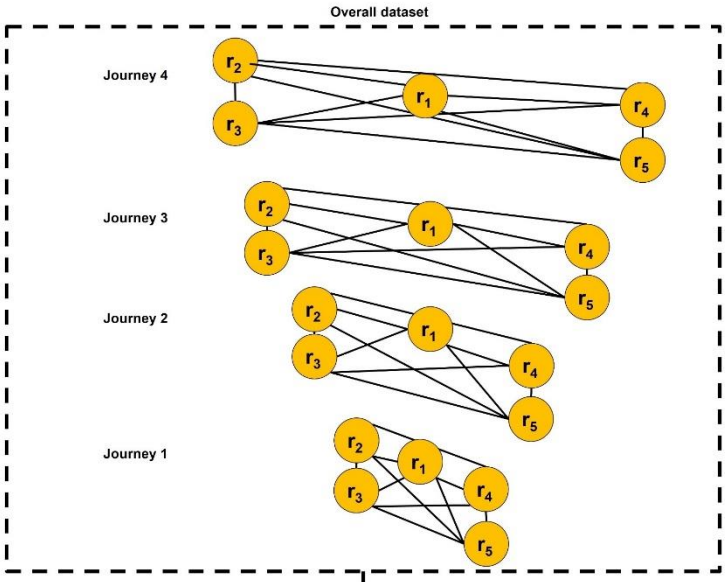
Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_3	0.92 +- 0.02	0.91
ND	RF	r_3	0.93 +- 0.01	0.94
ND	XGB	r_3	0.90 +- 0.02	0.92
D	KNN	r_3	0.92 +- 0.02	0.94
D	RF	r_3	0.92 +- 0.01	0.94
D	XGB	r_3	0.92 +- 0.02	0.94
D	GAT	r_3	0.92 +- 0.03	0.93



Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_4	0.92 +- 0.02	0.92
ND	RF	r_4	0.95 +- 0.02	0.93
ND	XGB	r_4	0.94 +- 0.03	0.91
D	KNN	r_4	0.92 +- 0.02	0.92
D	RF	r_4	0.95 +- 0.02	0.93
D	XGB	r_4	0.95 +- 0.03	0.93
D	GAT	r_4	0.94	0.87

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_5	0.92 +- 0.03	0.94
ND	RF	r_5	0.95 +- 0.02	0.93
ND	XGB	r_5	0.95 +- 0.02	0.93
D	KNN	r_5	0.92 +- 0.03	0.94
D	RF	r_5	0.95 +- 0.02	0.93
D	XGB	r_5	0.95 +- 0.01	0.92
D	GAT	r_5	0.93 +- 0.04	0.87

(d)



Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_2	0.62 +- 0.44	0.78
ND	RF	r_2	0.65 +- 0.4	0.87
ND	XGB	r_2	0.4 +- 0.49	0.67
D	KNN	r_2	0.77 +- 0.29	0.98
D	RF	r_2	0.83 +- 0.2	0.98
D	XGB	r_2	0.81 +- 0.21	0.96
D	GAT	r_2	0.89 +- 0.18	0.95

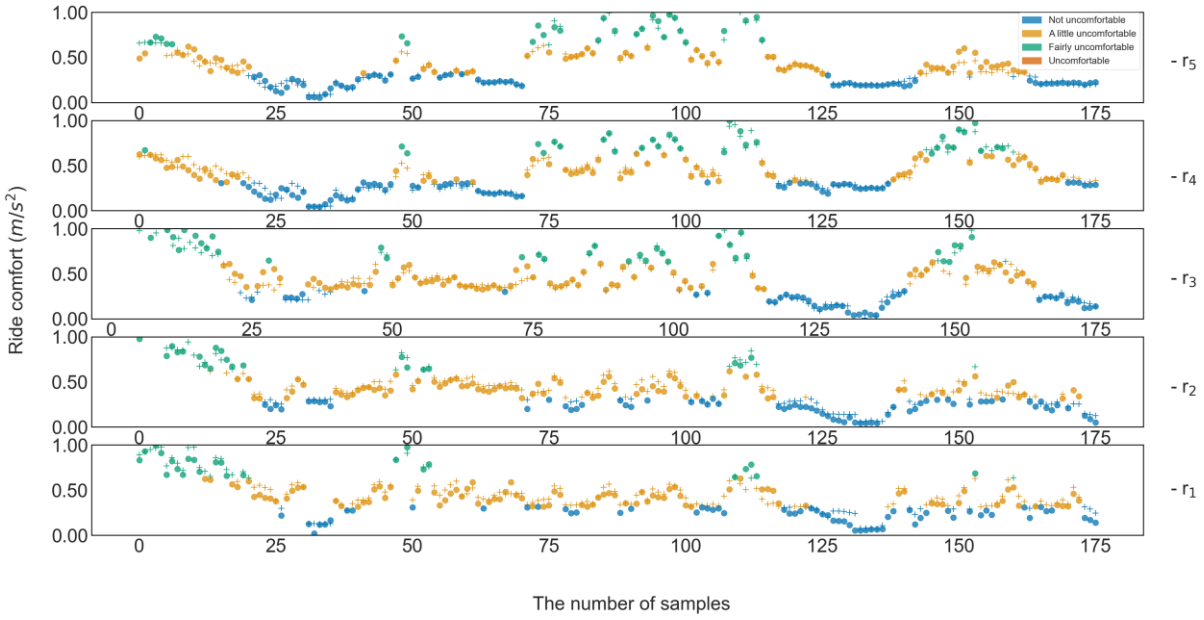
Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_1	-0.94 +- 3.46	-0.18
ND	RF	r_1	-0.24 +- 0.81	0.36
ND	XGB	r_1	-0.91 +- 3.07	-0.48
D	KNN	r_1	0.79 +- 0.23	0.7
D	RF	r_1	0.84 +- 0.14	0.76
D	XGB	r_1	0.65 +- 0.41	0.76
D	GAT	r_1	0.86 +- 0.14	0.84

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_4	0.75 +- 0.29	0.95
ND	RF	r_4	0.71 +- 0.25	0.95
ND	XGB	r_4	0.66 +- 0.24	0.8
D	KNN	r_4	0.68 +- 0.33	0.93
D	RF	r_4	0.75 +- 0.27	0.95
D	XGB	r_4	0.76 +- 0.27	0.91
D	GAT	r_4	0.86 +- 0.17	0.93

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_3	0.76 +- 0.33	0.98
ND	RF	r_3	0.75 +- 0.37	0.98
ND	XGB	r_3	0.62 +- 0.31	0.91
D	KNN	r_3	0.62 +- 0.46	0.98
D	RF	r_3	0.75 +- 0.37	0.98
D	XGB	r_3	0.72 +- 0.37	0.97
D	GAT	r_3	0.87 +- 0.24	0.92

Dataset	Model	Node	Training R^2	Test R^2
ND	KNN	r_5	0.63 +- 0.21	0.91
ND	RF	r_5	0.57 +- 0.18	0.90
ND	XGB	r_5	0.47 +- 0.33	0.92
D	KNN	r_5	0.54 +- 0.24	0.91
D	RF	r_5	0.61 +- 0.18	0.92
D	XGB	r_5	0.64 +- 0.2	0.95
D	GAT	r_5	0.88 +- 0.17	0.93

(e)



(f)

Figure 7-9 Model performance for (a) journey 1, (b) journey 2, (c) journey 3, (d) journey 4, (e) overall journey, (f) actual values vs predicted values

Figure 7-9 compares the performance of GAT with other models, such as KNN, RF, and XGBoost, across all journeys. Figure 7-9 (a)–(d) displays the performance of these models with fixed graphs. Comparing the performance of KNN, RF, and XGBoost using datasets – 'non-distance' (ND) and 'distance' (D) reveals that these models are not sensitive to the distance feature, as their performances with both datasets largely align. Nearly all predictions across different journeys and nodes achieve an R^2 value above 90%, indicating that the RC measurements from different locations on the train are highly correlated. This supports the hypothesis that the comfort level recorded by any four out of five phones can estimate the comfort of the remaining phone.

GAT is designed to understand the structural information inherent in the data, providing additional context that traditional ML models may miss. This is particularly evident in the superior performance of GAT at node 'r_5' of 'Journey 3' and node 'r_1' of 'Journey 4', where GAT significantly outperform KNN, RF, and XGBoost. In dynamic graphs (Figure 7-9(e)),

traditional ML models using the 'ND' dataset see a performance drop from 0.90 R^2 to about 0.70 R^2. At node 'r_1', traditional models even yield negative R^2 values, indicating predictions worse than a baseline using the average value of the test set's labels. Including the distance feature corrects this, improving R^2 to approximately 0.84 ± 0.14 on the training set and 0.76 on the test set.

When comparing GAT's performance on static and dynamic graphs, GAT demonstrates adaptability. Unlike traditional models that may struggle with changing graph structures over time, GAT can adjust the weights of adjacent nodes based on distance, showing superior results across all nodes on the dynamic dataset. This flexibility allows GAT to provide robust and accurate ride comfort assessments, adapting to real-time changes in distance.

GAT exhibits high accuracy and adaptability by integrating structural information into the graph for RC estimation. The distance feature benefits both GAT and traditional models when handling dynamic graphs. GAT outperforms other models by effectively extracting the topology of the data, making them well-suited for networked data typical in engineering contexts. Figure 7-9(f) confirms the robustness of GAT, with predicted values closely aligning with actual values at all locations. GAT's proficiency in handling acceleration data suggests their potential for other applications, such as real-time track condition monitoring. Additionally, the combination of crowd-sensing and GAT can be scaled to monitor entire trains across different lines and networks.

From an operational perspective, the cost of implementation and maintenance is essential. The proposed method is cost-effective, relying only on smartphones to sense train acceleration, with significant computational costs occurring during model training.

This method complements full-scale measurements specified by ISO 2631 and is easy to implement without requiring prior expert knowledge, ensuring satisfying performance at all nodes and providing a practical solution for measuring RC across entire trains or specific seats.

7.3.3 Robustness of AI in Crowd-Sensing

Trustworthy AI solutions must be robust, safe, and reliable when operating under real-world conditions involving passengers' private data. To achieve this, our novel approach is designed to function effectively in less-than-ideal circumstances while also being aware of when these conditions deteriorate. Our AI system is built to facilitate technical audits by user stakeholders (e.g., specialist engineers) and to explain how it reaches conclusions meaningfully for the user while also indicating when operational limits are reached.

In this study, we assess the vulnerabilities of our AI solution through extensive testing using perturbed data in compliance with the Joint Research Centre guidelines for the Robustness and interpretability of AI [155]. It is important to note that corrupted data will be identified and separated for evaluation by rail experts (i.e., human-centric AI), as shown in Figure 7-10. This methodology will ensure the trustworthiness of our AI/ML solution in real-world applications.

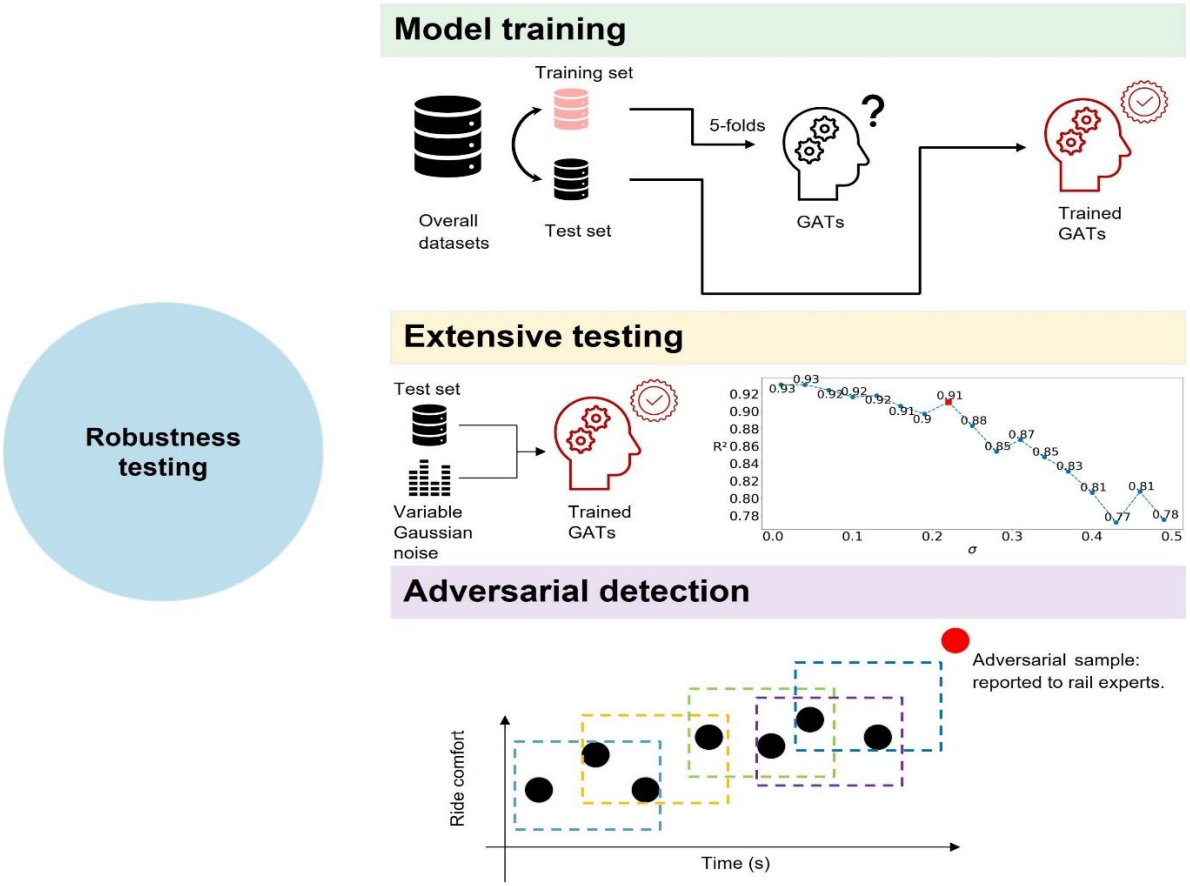


Figure 7-10 AI robustness assurances

The trained GATs are tested against perturbed data by introducing variable Gaussian noise to the test set. The initial plateau where R^2 values hover around 0.93 to 0.92 demonstrates the model's robustness to Gaussian noise with σ between 0 and 0.23. However, the model's performance declines dramatically when σ exceeds 0.23.

Adversarial detection is addressed by introducing an adaptive window to estimate the average value of the last three values (this parameter is adjustable). When the latest value exceeds a predefined threshold, the sample is classified as a potential adversarial sample and is sent to rail experts for further assessment.

7.4 Conclusions

Chapter 7 systematically examines the effectiveness of GAT in enhancing the accuracy of RC predictions by utilising crowd-sourced smartphone sensor data. To the best of our knowledge, leveraging crowd-sourced data to assess RC is unprecedented. The integration of GAT introduces a novel dimension in understanding the spatial relationships and structural dependencies inherent in the collected data, representing a significant advancement over conventional ML models.

The investigation is designed, utilising data from various journeys to ensure the robustness and generalisability of the findings. The results indicate that the GAT model consistently outperforms traditional models such as KNN, RF, and XGBoost, especially in scenarios where the data structure is dynamically changing. One of the most compelling insights is the importance of the distance feature in the models. The inclusion of spatial distances between smartphones is crucial in dynamic data contexts, where the GAT's attention mechanism effectively adjusts to the changing graph structures, demonstrating the model's adaptability and scalability.

Practically, this study underscores the potential of using widely available smartphones as reliable data collection tools. The application of GAT, combined with the ease of crowd-sensing, presents a cost-effective and scalable solution for real-time estimating and predicting the maintenance needs of railway systems. However, a primary concern with crowdsensing is the privacy of passengers' smartphone data.

While the results are promising, future research should aim to scale these findings by integrating them with other train management systems and conducting real-world trials to validate the effectiveness of GATs further using additional smartphones.

Chapter 8

Conclusions and Recommendations

8.1 Summary of Key Findings

This thesis aims to improve the assessment of RC in railway systems by combining ML algorithms with smartphone data. Traditional methods, relying on questionnaires or specialised sensors, lacked scalability and precision. In contrast, this study leverages smartphones, already equipped with advanced sensors, to gather real-time data on train dynamics. This chapter summarises the key research findings from quantifying track stiffness, classifying train motions, and assessing RC at specific points to an overall comfort evaluation using crowd-sensing techniques.

This thesis first examines the effectiveness of the vibration feature in predicting track stiffness. The proposed solution utilises a dilated CNN to analyse axle-box vibrations collected through the D-track simulation software. This method enables faster and more scalable track stiffness estimation. The dilated CNN model demonstrated excellent

performance, with over 90% R^2 across varying conditions of track stiffness and train speeds. One of the critical findings is that track irregularities (such as dipped joints and welds) have minimal impact on model performance, reinforcing the approach's robustness. This study suggests that ML-based track stiffness assessment could be integrated into daily operations without disrupting service. At the same time, the vibration feature is confirmed to be informative in railway systems.

Following the track stiffness assessment, the next focus is classifying train motions directly affecting RC. The research identifies motions, including roll, yaw, pitch, bounce, and hunting, all contributing to the RC. Smartphones placed at different points in the train provide real-time data for this classification.

With unsupervised learning, the ML models accurately identified motion types by analysing features from both time-domain and frequency-domain data. This classification enables better insights into the factors contributing to ride discomfort, such as yawing motions and hunting behaviour in the wheelsets. These results underscore the potential of smartphone-based sensors to provide a more precise and cost-effective alternative to traditional methods, improving both passenger safety and comfort.

Building on the successful classification of train motions, the study subsequently quantifies the RC at individual points within the train. The ensemble ML models, trained using smartphone data, can predict RC with high accuracy, which allows for identifying areas within the train that are more prone to discomfort and estimating the RC for each passenger. This is a significant advancement, as traditional methods often offer only a general assessment of RC instead of evaluating the impact of vibration on the passenger.

The final step in this research is to scale up the RC assessment to an overall train-level evaluation using crowd-sensing techniques. The study can construct a comprehensive, real-time comfort map across different seating locations by gathering data from multiple smartphones across the train. The GAT model utilised shows excellent robustness, effectively integrating crowd-sourced data to provide a reliable comfort assessment. This approach, which reduces reliance on expensive hardware and specialised sensors, demonstrates the potential for scalable and efficient comfort assessments across entire train systems. The ability to perform real-time, large-scale assessments opens new possibilities for railway operators to optimise service and maintenance.

The research findings clearly demonstrate the effectiveness of ML models and smartphone data in revolutionising RC assessments in railway systems. The integration of these technologies provides a scalable, cost-effective, and highly accurate method for real-time evaluation. The successful application of crowd-sensing further highlights the potential for broader implementation across railway networks, enhancing both passenger satisfaction and operational efficiency.

8.2 Recommendations for Future Research

Recommendations for the future direction are listed as follows:

1. Additional factors can be incorporated for a more comprehensive comfort assessment if the human response for the factors is well-quantified. For instance, high noise levels can increase discomfort, especially during long journeys, even if vibration levels are within acceptable limits.

2. The potential of the crowd-sensing model across other transport modes can be estimated to create adaptable, universal comfort standards. The crowd-sensing approach, which aggregates data from multiple smartphone sensors, can be extended to other transport modes such as buses and trams.
3. Adaptive AI models that respond to real-time disruptions and variability can be considered to improve robustness in dynamic environments.
4. Research on personalised comfort assessments based on individual preferences and physiological responses can be integrated into passenger Apps.
5. It is recommended that solid encryption and anonymisation techniques be implemented to safeguard passenger data while ensuring compliance with regulations.
6. Further investigation into the long-term effects of continuous monitoring on operational efficiency, maintenance, and customer satisfaction is recommended.
7. Working closely with railway operators, engineers, and industry stakeholders is required to improve the developed approaches to meet real-world demands and regulatory standards.
8. Further research might explore environmental impacts concerning the adoption of ML for railway operations.
9. Further experiment is required to examine the model's complexity and computational cost. As models grow in sophistication, with additional features or higher data resolutions, the computational cost and complexity increase. This can make it more challenging to implement the model on widely accessible devices, such as smartphones. Conducting further experiments can help identify the most efficient algorithms and optimisations, balancing accuracy with computational efficiency.

10. Evaluate the holistic benefits and disadvantages (economic, social, and environmental) of using the proposed methods in the railway industry.
11. Study the need for workforce reskilling and upskilling to effectively implement the proposed approaches, focusing on training for railway maintenance teams.
12. Integrating GPS into the proposed system would enhance performance by providing precise location data for each measurement. This spatial context allows for identifying specific track sections with recurring comfort issues, enabling targeted maintenance. GPS also aids in synchronising data from multiple devices, improving the accuracy of the crowd-sensing model.

References

1. da Silva, P.F.P. and J. Mendes, *Passengers comfort perception and demands on railway vehicles: A review.* KnE Engineering, 2020: p. 257–270-257–270.
2. Beirão, G. and J.S. Cabral, *Understanding attitudes towards public transport and private car: A qualitative study.* Transport policy, 2007. **14**(6): p. 478-489.
3. Givoni, M., C. Brand, and P. Watkiss, *Are railways climate friendly?* Built Environment, 2009. **35**(1): p. 70-86.
4. Mouwen, A., *Drivers of customer satisfaction with public transport services.* Transportation Research Part A: Policy and Practice, 2015. **78**: p. 1-20.
5. Soehodho, S., *Public transportation development and traffic accident prevention in Indonesia.* IATSS research, 2017. **40**(2): p. 76-80.
6. Borhan, M.N., et al., *Why public bus is a less attractive mode of transport: A case study of Putrajaya, Malaysia.* Periodica Polytechnica Transportation Engineering, 2019. **47**(1): p. 82-90.
7. Asplund, M., et al., *Introduction to the Collaboration Project: A Systematic Approach to Improve Passenger Ride Comfort.* Trafikverket TRV, 2023. **113404**: p. 178508100-001.
8. Oboknb, D. and M. Clarke, *The development of questionnaire surveys for the investigation of passenger comfort.* Ergonomics, 1973. **16**(6): p. 855-869.
9. Kim, Y., et al., *Correlation of ride comfort evaluation methods for railway vehicles.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2003. **217**(2): p. 73-88.
10. Lee, Y., et al., *Ride comfort analysis with physiological parameters for an e-health train.* Telemedicine and e-Health, 2009. **15**(10): p. 1010-1021.
11. Tang, Z. and J. Xu. *Research on the physiological and psychological factors' evaluation index framework of train ride comfort*. in *IEEE Conference Anthology*. 2013. IEEE.
12. da Fonseca-Soares, D., et al., *Greenhouse gas emissions in railways: Systematic review of research progress.* Buildings, 2024. **14**(2): p. 539.
13. Griffin, M.J., *Handbook of human vibration*. 2012: Academic press.
14. Stevens, S.S., *On the psychophysical law.* Psychological review, 1957. **64**(3): p. 153.
15. Howarth, H. *A study of the growth of human discomfort with increasing magnitude of mechanical shocks*. in *JOURNAL OF SOUND AND VIBRATION*. 1988. ACADEMIC PRESS LTD 24-28 OVAL RD, LONDON, ENGLAND NW1 7DX.
16. Rubin, C., et al., *Low mechanical signals strengthen long bones.* Nature, 2001. **412**(6847): p. 603-604.
17. Fuermaier, A.B., et al., *Good vibrations–effects of whole body vibration on attention in healthy individuals and individuals with ADHD.* PLoS One, 2014. **9**(2): p. e90747.

18. Coelho-Oliveira, A.C., et al., *Acute whole-body vibration exercise promotes favorable handgrip neuromuscular modifications in rheumatoid arthritis: A cross-over randomized clinical.* BioMed Research International, 2021. **2021**.
19. Marayong, P., et al. *Vibrotactile device for rehabilitative training of persons with lower-limb amputation*. in *2014 IEEE Healthcare Innovation Conference (HIC)*. 2014. IEEE.
20. Hornik, B., et al., *The Effect of Mechanically-Generated Vibrations on the Efficacy of Hemodialysis; Assessment of Patients' Safety: Preliminary Reports.* International Journal of Environmental Research and Public Health, 2019. **16**(4): p. 594.
21. Niehoff, A., et al., *Effect of whole-body vibration and insulin-like growth factor-I on muscle paralysis-induced bone degeneration after botulinum toxin injection in mice.* Calcified tissue international, 2014. **94**: p. 373-383.
22. de Normalisation, C.E., *Mechanical vibration—guide to the health effects of vibration on the human body.* CR report, 1996. **12349**.
23. Bovenzi, M., *Health risks from occupational exposures to mechanical vibration.* La Medicina del lavoro, 2006. **97**(3): p. 535-541.
24. MIYAZAKI, Y., *Adverse effects of whole-body vibration on gastric motility.* The Kurume medical journal, 2000. **47**(1): p. 79-86.
25. Nakashima, A. and B. Cheung, *The effects of vibration frequencies on physical, perceptual and cognitive performance.* Toronto, Ontario, CA: Defence Research and Development Canada, 2006.
26. Paul de Vos, S., *Railway induced vibration*, L.G. Nick Craven, Editor. 2017, International Union of Railways (UIC): Paris.
27. Ahn, Y., J. Rhie, and M.-G. Kim, *The relevant factors of work-related fatigue for occupational vibration-exposed employees.* Annals of Occupational and Environmental Medicine, 2022. **34**.
28. An, N. and S. SI, *Mechanical Vibration and Shock-Evaluation of Human Exposure to Whole-Body Vibration-Part 1: General Requirements.* 1997.
29. AN, N. and S. SI, *Mechanical vibration and shock–evaluation of human exposure to whole body vibration–part 4: Guidelines for the evaluation of the effects of vibration and rotational motion on passenger and crew comfort in fixed-guideway transport systems.* Mechanical Vibration and Shock–Evaluation of Human Exposure to Whole Body Vibration–Part 4: Guidelines for the Evaluation of the Effects of Vibration and Rotational Motion on Passenger and Crew Comfort in Fixed-Guideway Transport Systems, 2007.
30. Leaflet, U., *513R, Guidelines for evaluating passenger comfort in relation to vibration in railway vehicle.* International Union of Railways, 1994.
31. Sperling, E. and C. Betzhold, *Beitrag zur beurteilung des fahrkomforts in schienenfahrzeugen.* Glasers Annalen, 1956. **80**(10): p. 314-317.
32. CEN, E., *12299: Railway applications–Ride comfort for passengers–Measurement and evaluation.* British Standards Institution: London, UK, 2009.
33. Graa, M., et al., *Development of a reduced dynamic model for comfort evaluation of rail vehicle systems.* Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics, 2016. **230**(4): p. 489-504.

34. Karakasis, K., D. Skarlatos, and T. Zakinthinos, *A factorial analysis for the determination of an optimal train speed with a desired ride comfort.* Applied Acoustics, 2005. **66**(10): p. 1121-1134.
35. Zoccali, P., G. Loprencipe, and R.C. Lupascu, *Acceleration measurements inside vehicles: Passengers' comfort mapping on railways.* Measurement, 2018. **129**: p. 489-498.
36. Junior, J.U.P., J.N. Chaves, and L.A. de Souza Santos, *Developing a Ride Comfort Monitoring System from Scratch: An Experience in a Suburban Railway.* Sustainable Rail Transport: Proceedings of RailNewcastle 2017, 2018: p. 19.
37. Kardas-Cinal, E., *Ride comfort for various passenger positions in a railway vehicle-simulation study.* Archives of Transport, 2010.
38. Bakinowski, Ł. and B. Firlik, *Influence of the type of place occupied by a tram passenger on the ride comfort.* Archives of Transport, 2022. **62**.
39. Pradhan, S. and A.K. Samantaray, *Integrated modeling and simulation of vehicle and human multi-body dynamics for comfort assessment in railway vehicles.* Journal of Mechanical Science and Technology, 2018. **32**: p. 109-119.
40. Liu, C., et al., *Effect of train speed and track geometry on the ride comfort in high-speed railways based on ISO 2631-1.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2020. **234**(7): p. 765-778.
41. Dumitriu, M. and D.I. Stănică, *Study on the evaluation methods of the vertical ride comfort of railway vehicle—mean comfort method and sperling's method.* Applied Sciences, 2021. **11**(9): p. 3953.
42. Zboiński, K., P. Woźnica, and Y. Bolzhelarskyi, *Modelling of the shape of railway transition curves from the point of view of passenger comfort.* Archives of Transport, 2021. **60**(4).
43. Montenegro, P., et al., *Impact of the train-track-bridge system characteristics in the runnability of high-speed trains against crosswinds-Part II: Riding comfort.* Journal of Wind Engineering and Industrial Aerodynamics, 2022. **224**: p. 104987.
44. LEICHT, L., L. HECKELE, and P. GRATZFELD, *INFLUENCE OF MODIFIED POWERTRAIN MASSES ON THE RIDE COMFORT OF A LIGHT RAIL VEHICLE.* Computers in Railways XVIII: Railway Engineering Design and Operation, 2022. **213**: p. 77.
45. Isacchi, G., M. Corsi, and F. Ripamonti, *Ride comfort assessment of high-speed rail vehicles: influence of yaw dampers installation angle*, in *Vibroengineering Procedia*. 2023, ITA. p. 98-104.
46. Novillo, G., et al., *Analysis of Tram Comfort Using The UNE EN 12299: 2010 Standard and Sperling Method (WZ).* Computers in Railways XVIII: Railway Engineering Design and Operation, 2022. **213**: p. 3.
47. Dumitriu, M. and M.C. Leu, *Correlation between Ride Comfort Index and Sperling's Index for Evaluation Ride Comfort in Railway Vehicles.* Applied Mechanics and Materials, 2018. **880**: p. 201-206.
48. Deng, C., et al., *Analysis of the consistency of the Sperling index for rail vehicles based on different algorithms.* Vehicle System Dynamics, 2021. **59**(2): p. 313-330.
49. Kumar, V., V. Rastogi, and P.M. Pathak, *Simulation for whole-body vibration to assess ride comfort of a low–medium speed railway vehicle.* Simulation, 2017. **93**(3): p. 225-236.

50. Lee, J.H., B.S. Jin, and Y. Ji, *Development of a Structural Equation Model for ride comfort of the Korean high-speed railway.* International Journal of Industrial Ergonomics, 2009. **39**(1): p. 7-14.
51. Nicol, J., et al., *Comfort studies of rail passengers.* Occupational and Environmental Medicine, 1973. **30**(4): p. 325-334.
52. Sameni, M.K., M.S.G. Maghami, and H.S. Naeini, *Assessing Passenger Comfort in Compartments of Long Distance Conventional Trains: A Case Study.* International Journal of Occupational Hygiene, 2022. **14**(2): p. 157-163.
53. bin Mohd Masirin, M.I., et al., *A field study on urban rail transits in city of Kuala Lumpur: passenger views on train noise and vibration.* Applied Mechanics and Materials, 2015. **773**: p. 839-844.
54. Khan, M.S., *Evaluation of acoustical comfort in passenger trains.* Acta Acustica united with Acustica, 2002. **88**(2): p. 270-277.
55. Sanok, S., et al., *Passenger comfort on high-speed trains: effect of tunnel noise on the subjective assessment of pressure variations.* Ergonomics, 2015. **58**(6): p. 1022-1031.
56. Qiu, Y. and M.J. Griffin, *Modelling the fore-and-aft apparent mass of the human body and the transmissibility of seat backrests.* Vehicle system dynamics, 2011. **49**(5): p. 703-722.
57. Liang, C.-C. and C.-F. Chiang, *A study on biodynamic models of seated human subjects exposed to vertical vibration.* International Journal of Industrial Ergonomics, 2006. **36**(10): p. 869-890.
58. Qassem, W., *Model prediction of vibration effects on human subject seated on various cushions.* Medical engineering & physics, 1996. **18**(5): p. 350-358.
59. Peng, Y., et al., *Passenger overall comfort in high-speed railway environments based on EEG: assessment and degradation mechanism.* Building and Environment, 2022. **210**: p. 108711.
60. Geselowitz, D.B. and O. Schmitt, *Electrocardiography.* Biological Engineering, 1969: p. 333-390.
61. Huang, S., et al., *Detection of mental fatigue state with wearable ECG devices.* International journal of medical informatics, 2018. **119**: p. 39-46.
62. Wang, Z., R. Matsuhashi, and H. Onodera, *Towards wearable thermal comfort assessment framework by analysis of heart rate variability.* Building and Environment, 2022. **223**: p. 109504.
63. Al-Nashash, H., et al., *ECG response of the human body subjected to vibrations.* Journal of medical engineering & technology, 1996. **20**(1): p. 2-10.
64. Sattari, S., et al., *Evaluation of Sperling's Index in Passenger and Freight Trains Under Different Speeds and Track Irregularities.* International Journal of Advanced Design & Manufacturing Technology, 2022. **15**(4).
65. Liu, Y., et al., *Dynamic Train Vertical Sperling Index Evaluation Model Considering Wheel-Rail Contact Loss.* Journal of Shanghai Jiaotong University (Science), 2022: p. 1-13.
66. Rezvani, M.A., R. Arcos, and V. Bokaeian, *The effect of onboard passengers' seating arrangement on the vertical ride comfort of a high-speed railway vehicle.* Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2022. **236**(15): p. 8221-8230.

67. Sharma, R.C., *Evaluation of Passenger Ride Comfort of Indian Rail and Road Vehicles with ISO 2631-1 Standards: Part 1-Mathematical Modeling.* International Journal of Vehicle Structures & Systems (IJVSS), 2016. **8**(1).
68. Liu, D., et al., *Momentary discomfort of high-speed trains passing through complex terrain sections under strong wind conditions.* Vehicle system dynamics, 2020. **58**(9): p. 1428-1450.
69. Sun, J., et al., *An investigation into evaluation methods for ride comfort of railway vehicles in the case of carbody hunting instability.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2021. **235**(5): p. 586-597.
70. Nastac, S. and M. Picu, *Evaluating methods of whole-body-vibration exposure in trains.* Ann."Dunarea De Jos" Univ. Galati, Fasc. XIV Mech. Eng, 2010. **2**: p. 55-60.
71. Peng, Y., et al., *Assessment of passenger long-term vibration discomfort: a field study in high-speed train environments.* Ergonomics, 2022. **65**(4): p. 659-671.
72. Fedatto Neto, M. and H.M. Gomes, *Analysis of vibration levels in users of urban trains.* Estudos tecnológicos em Engenharia [recurso eletrônico]. São Leopoldo, RS. Vol. 12, n. 2 (jul./dez. 2018), p. 24-44, 2018.
73. Meng, Q., P. Lu, and S. Zhu, *A smartphone-enabled IoT system for vibration and noise monitoring of rail transit.* IEEE Internet of Things Journal, 2023. **10**(10): p. 8907-8917.
74. Zhang, Y., et al., *Elevator ride comfort monitoring and evaluation using smartphones.* Mechanical Systems and Signal Processing, 2018. **105**: p. 377-390.
75. Kim, Y.-G., et al., *AN EXPERIMENTAL STUDY ON THE RIDE COMFORT OF THE KOREAN HIGH-SPEED TRAIN.* Experimental techniques, 2009. **33**(6).
76. Yu, Z., et al., *Momentary ride comfort evaluation of high-speed trains based on feature selection and gated recurrent unit network.* Journal of Sound and Vibration, 2023. **558**: p. 117769.
77. Zhou, W., et al., *Enhancing thermal comfort prediction in high-speed trains through machine learning and physiological signals integration.* Journal of Thermal Biology, 2024. **121**: p. 103828.
78. Yuxue, B., et al., *Sitting comfort analysis and prediction for high-speed rail passengers based on statistical analysis and machine learning.* Building and Environment, 2022. **225**: p. 109589.
79. Kyung, G. and M.A. Nussbaum, *Driver sitting comfort and discomfort (part II): Relationships with and prediction from interface pressure.* International Journal of Industrial Ergonomics, 2008. **38**(5-6): p. 526-538.
80. Azzoug, A. and S. Kaewunruen, *RideComfort: a development of crowdsourcing smartphones in measuring train ride quality.* Frontiers in Built Environment, 2017. **3**: p. 3.
81. Tien Do, N., et al. *Evaluating Passenger Railway Ride Quality Over Long Distances Using Smartphones*. in *ASME/IEEE Joint Rail Conference*. 2020. American Society of Mechanical Engineers.
82. Rodriguez, A., et al., *Smartphones and tablets applications in railways, ride comfort and track quality. Transition zones analysis.* Measurement, 2021. **182**: p. 109644.

83. Amador-Jimenez, L., A. Mohammadi, and F. Nasiri. *Level of comfort and safety in railway transit*. in *2017 4th International Conference on Transportation Information and Safety (ICTIS)*. 2017. IEEE.
84. George, T.K., et al. *Sensing discomfort of standing passengers in public rail transportation systems using a smart phone*. in *2013 10th IEEE international conference on control and automation (ICCA)*. 2013. IEEE.
85. Barabino, B., et al., *Standing passenger comfort: A new scale for evaluating the real-time driving style of bus transit services.* IEEE Transactions on Intelligent Transportation Systems, 2019. **20**(12): p. 4665-4678.
86. Cai, Z., G. Raymond, and R. Bathurst, *Natural vibration analysis of rail track as a system of elastically coupled beam structures on Winkler foundation.* Computers & structures, 1994. **53**(6): p. 1427-1436.
87. Steffens, D.M., *Identification and development of a model of railway track dynamic behaviour*. 2005, Queensland University of Technology.
88. Sandhu, R.S. *On Five Definitions of Data Integrity*. in *DBSec*. 1993. Citeseer.
89. Byun, J.-W., Y. Sohn, and E. Bertino. *Systematic control and management of data integrity*. in *Proceedings of the eleventh ACM symposium on Access control models and technologies*. 2006.
90. Hawkins, D.M., *Identification of outliers*. Vol. 11. 1980: Springer.
91. Boukerche, A., L. Zheng, and O. Alfandi, *Outlier detection: Methods, models, and classification.* ACM Computing Surveys (CSUR), 2020. **53**(3): p. 1-37.
92. Smiti, A., *A critical overview of outlier detection methods.* Computer Science Review, 2020. **38**: p. 100306.
93. Pariyal, P.S., et al., *Comparison based analysis of different FFT architectures.* International journal of image, graphics and signal processing, 2016. **8**(6): p. 41.
94. Bulgakov, V. and A. Segal, *Dimensionality Reduction in Sentence Transformer Vector Databases with Fast Fourier Transform.* arXiv preprint arXiv:2404.06278, 2024.
95. Hall, D.L. and J. Llinas, *An introduction to multisensor data fusion.* Proceedings of the IEEE, 1997. **85**(1): p. 6-23.
96. Sudha, D. and M. Ramakrishna. *Comparative study of features fusion techniques*. in *2017 international conference on recent advances in electronics and communication technology (ICRAECT)*. 2017. IEEE.
97. Engelbrecht, A.P., *Computational intelligence: an introduction*. 2007: John Wiley & Sons.
98. Shaheen, M.Y., *Applications of Artificial Intelligence (AI) in healthcare: A review.* ScienceOpen Preprints, 2021.
99. Cao, L., *AI in finance: A review.* Available at SSRN 3647625, 2020.
100. Wang, K. and Y. Wang. *How AI affects the future predictive maintenance: a primer of deep learning*. in *Advanced Manufacturing and Automation VII 7*. 2018. Springer.
101. Murphy, R.R., *Introduction to AI robotics*. 2019: MIT press.
102. Cunningham, P., M. Cord, and S.J. Delany, *Supervised learning*, in *Machine learning techniques for multimedia: case studies on organization and retrieval*. 2008, Springer. p. 21-49.
103. Hastie, T., et al., *Unsupervised learning.* The elements of statistical learning: Data mining, inference, and prediction, 2009: p. 485-585.

104. Wiering, M.A. and M. Van Otterlo, *Reinforcement learning.* Adaptation, learning, and optimization, 2012. **12**(3): p. 729.
105. Agatonovic-Kustrin, S. and R. Beresford, *Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research.* Journal of pharmaceutical and biomedical analysis, 2000. **22**(5): p. 717-727.
106. Zupan, J., *Introduction to artificial neural network (ANN) methods: what they are and how to use them.* Acta Chimica Slovenica, 1994. **41**(3): p. 327.
107. Chicco, D., M.J. Warrens, and G. Jurman, *The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation.* PeerJ Computer Science, 2021. **7**: p. e623.
108. Shahapure, K.R. and C. Nicholas. *Cluster quality analysis using silhouette score*. in *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*. 2020. IEEE.
109. Davies, D.L. and D.W. Bouldin, *A cluster separation measure.* IEEE transactions on pattern analysis and machine intelligence, 1979(2): p. 224-227.
110. Bengio, Y. and Y. Grandvalet, *No unbiased estimator of the variance of k-fold cross-validation.* Advances in Neural Information Processing Systems, 2003. **16**.
111. Pita, A.L., P.F. Teixeira, and F. Robuste, *High speed and track deterioration: the role of vertical stiffness of the track.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2004. **218**(1): p. 31-40.
112. Ebersohn, W., *Substructure influence on track maintenance requirements*. 1995: University of Massachusetts Amherst.
113. Ebersöhn, W. and E.T. Selig, *Track modulus measurements on a heavy haul line.* Transportation Research Record, 1994. **1470**: p. 73.
114. Berggren, E., Å. Jahlénius, and B.-E. Bengtsson, *Continuous track stiffness measurement: An effective method to investigate the structural conditions of the track.* 2002.
115. Wang, P., et al., *Overview and outlook on railway track stiffness measurement.* Journal of Modern Transportation, 2016. **24**: p. 89-102.
116. Kerr, A.D., *On the determination of the rail support modulus k.* International Journal of Solids and Structures, 2000. **37**(32): p. 4335-4351.
117. Zhang, G., *Research on right level of track structure stiffness and track-part stiffness.* China Railw Sci, 2002. **23**(1): p. 51-57.
118. Berggren, E., *Railway track stiffness: dynamic measurements and evaluation for efficient maintenance*. 2009, KTH.
119. Kaewunruen, S. and A. Remennikov, *Application of experimental modal testing for estimating dynamic properties of structural components.* 2005.
120. Kaewunruen, S. and A. Remennikov, *Monitoring structural degradation of rail pads in laboratory using impact excitation technique.* 2005.
121. Hosseingholian, M., M. Froumentin, and A. Robinet. *Dynamic track modulus from measurement of track acceleration by portancemetre*. in *9th world congress on railway research*. 2011.
122. Burrow, M.P., A.H. Chan, and A. Shein, *Deflectometer-based analysis of ballasted railway tracks.* Proceedings of the Institution of Civil Engineers-Geotechnical Engineering, 2007. **160**(3): p. 169-177.

123. Wangqing, W., et al. *Development of inspection car for measuring railway track elasticity*. in *Proceedings from 6th international heavy haul conference, Cape Town*. 1997.
124. Quirke, P., et al., *Drive-by detection of railway track stiffness variation using in-service vehicles.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2017. **231**(4): p. 498-514.
125. Mehrali, M., M. Esmaeili, and S. Mohammadzadeh, *Application of data mining techniques for the investigation of track geometry and stiffness variation.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2020. **234**(5): p. 439-453.
126. Li, D., C. Ngamkhanong, and S. Kaewunruen. *Time-dependent topology of railway prestressed concrete sleepers*. in *IOP Conference Series: Materials Science and Engineering*. 2017. IOP Publishing.
127. Chen, L.-C., et al., *Semantic image segmentation with deep convolutional nets and fully connected crfs.* arXiv preprint arXiv:1412.7062, 2014.
128. Meng, H., et al., *Speech emotion recognition from 3D log-mel spectrograms with deep learning network.* IEEE access, 2019. **7**: p. 125868-125881.
129. Lei, X., H. Pan, and X. Huang, *A dilated CNN model for image classification.* IEEE Access, 2019. **7**: p. 124087-124095.
130. Jain, A., K. Nandakumar, and A. Ross, *Score normalization in multimodal biometric systems.* Pattern recognition, 2005. **38**(12): p. 2270-2285.
131. Hosseingholian, M., M. Froumentin, and D. Levacher, *Continuous method to measure track stiffness a new tool for inspection of rail infrastructure.* World Applied Sciences Journal, 2009. **6**(5): p. 579-589.
132. Puzavac, L. and Z. Popović, *Vertical track geometry deterioration modeling.* Izgradnja, 2010. **64**(1-2): p. 7-20.
133. Sussman, T., W. Ebersöhn, and E. Selig, *Fundamental nonlinear track load-deflection behavior for condition evaluation.* Transportation Research Record, 2001. **1742**(1): p. 61-67.
134. Esveld, C. and C. Esveld, *Modern railway track*. Vol. 385. 2001: MRT-productions Zaltbommel.
135. Jaekel, B. and T. Albrecht, *Comparative analysis of algorithms and models for train running simulation.* Journal of Rail Transport Planning & Management, 2014. **4**(1-2): p. 14-27.
136. Hansen, I. and J. Pachl, *Railway timetabling and operations: Analysis, modelling, optimisation, simulation, performance, evaluation*. 2014: Eurail press.
137. Cunillera, A., et al., *Real-time train motion parameter estimation using an unscented Kalman filter.* Transportation Research Part C: Emerging Technologies, 2022. **143**: p. 103794.
138. De Martinis, V. and F. Corman. *Online microscopic calibration of train motion models: towards the era of customized control solutions*. in *RailNorrköping 2019. 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), Norrköping, Sweden, June 17th–20th, 2019*. 2019. Linköping University Electronic Press.

139. Alavi, A.H. and W.G. Buttlar, *An overview of smartphone technology for citizen-centered, real-time and scalable civil infrastructure monitoring.* Future Generation Computer Systems, 2019. **93**: p. 651-672.
140. Chen, C.-Y., et al., *A review of ubiquitous mobile sensing based on smartphones.* International Journal of Automation and Smart Technology, 2014. **4**(1): p. 13-19.
141. Wickens, A., *Fundamentals of rail vehicle dynamics*. 2003: CRC Press.
142. Matsudaira, T. *Paper 7: hunting problem of high-speed railway vehicles with special reference to bogie design for the new tokaido line*. in *Proceedings of the Institution of Mechanical Engineers, Conference Proceedings*. 1965. SAGE Publications Sage UK: London, England.
143. Huang, C., J. Zeng, and S. Liang, *Carbody hunting investigation of a high speed passenger car.* Journal of mechanical science and technology, 2013. **27**: p. 2283-2292.
144. EN, *EN 14363*, in *Railway applications - Testing and Simulation for the acceptance of running characteristics of railway vehicles -Running Behaviour and stationary tests*. 2016: CEN, Brussels.
145. Knothe, K. and S. Stichel, *Rail vehicle dynamics*. 2017: Springer.
146. Goddard, M., *The EU General Data Protection Regulation (GDPR): European regulation that has a global impact.* International Journal of Market Research, 2017. **59**(6): p. 703-705.
147. Kuhlmann, T., P. Garaizar, and U.-D. Reips, *Smartphone sensor accuracy varies from device to device in mobile research: The case of spatial orientation.* Behavior research methods, 2021. **53**: p. 22-33.
148. Grouios, G., et al., *Accelerometers in our pocket: Does smartphone accelerometer technology provide accurate data?* Sensors, 2022. **23**(1): p. 192.
149. Rahm, E. and H.H. Do, *Data cleaning: Problems and current approaches.* IEEE Data Eng. Bull., 2000. **23**(4): p. 3-13.
150. Hu, H., et al., *An effective and adaptable K-means algorithm for big data cluster analysis.* Pattern Recognition, 2023. **139**: p. 109404.
151. Na, S., L. Xumin, and G. Yong. *Research on k-means clustering algorithm: An improved k-means clustering algorithm*. in *2010 Third International Symposium on intelligent information technology and security informatics*. 2010. Ieee.
152. Burkardt, J., *K-means clustering.* Virginia Tech, Advanced Research Computing, Interdisciplinary Center for Applied Mathematics, 2009.
153. Doshi-Velez, F. and B. Kim, *Towards a rigorous science of interpretable machine learning.* arXiv preprint arXiv:1702.08608, 2017.
154. Rudin, C., *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead.* Nature machine intelligence, 2019. **1**(5): p. 206-215.
155. Hamon, R., H. Junklewitz, and I. Sanchez, *Robustness and explainability of artificial intelligence.* Publications Office of the European Union, 2020. **207**: p. 2020.
156. Tyagi, K., et al., *Unsupervised learning*, in *Artificial intelligence and machine learning for edge computing*. 2022, Elsevier. p. 33-52.
157. Müllner, D., *Modern hierarchical, agglomerative clustering algorithms.* arXiv preprint arXiv:1109.2378, 2011.

158. Standardization, I.O.f., *ISO 2631-1:1997*, in *Mechanical vibration and shock—evaluation of human exposure*

to whole body vibration—part 1: general requirements. 1997.
159. Mohan, A., *Nonlinear investigation of the use of controllable primary suspensions to improve hunting in railway vehicles*. 2003, Virginia Tech.
160. Dongfang, S., *Running dynamic properties of small-radius railway wheels*. 2021: Technische Universitaet Berlin (Germany).
161. Iwnicki, S., *Handbook of railway vehicle dynamics*. 2006: CRC press.
162. Terrell, G.R. and D.W. Scott, *Variable kernel density estimation.* The Annals of Statistics, 1992: p. 1236-1265.
163. Chen, Y.-C., *A tutorial on kernel density estimation and recent advances.* Biostatistics & Epidemiology, 2017. **1**(1): p. 161-187.
164. Konowrocki, R. and A. Chojnacki, *Analysis of rail vehicles' operational reliability in the aspect of safety against derailment based on various methods of determining the assessment criterion.* Eksploatacja i Niezawodność, 2020. **22**(1): p. 73-85.
165. Croft, B., C. Jones, and D. Thompson, *Modelling the effect of rail dampers on wheel–rail interaction forces and rail roughness growth rates.* Journal of Sound and Vibration, 2009. **323**(1-2): p. 17-32.
166. Cheng, Y.-C. and P.-H. Wu, *Wheel nominal rolling radius difference on hunting stability of railway vehicle system under a speed-dependent nonlinear creep model.* Journal of Vibration and Control, 2018. **24**(14): p. 3107-3123.
167. Yoo, W.-S., et al., *Development and application of new evaluation system for ride comfort and vibration on railway vehicles.* Journal of Mechanical Science and Technology, 2005. **19**: p. 1469-1477.
168. Lee, C.-M., et al., *Ride comfort of a high-speed train through the structural upgrade of a bogie suspension.* Journal of Sound and Vibration, 2016. **361**: p. 99-107.
169. Sadeghi, J., S. Rabiee, and A. Khajehdezfuly, *Effect of rail irregularities on ride comfort of train moving over ballast-less tracks.* International Journal of Structural Stability and Dynamics, 2019. **19**(06): p. 1950060.
170. Alehashem, S.S., Y.-Q. Ni, and X. Liu, *A full-scale experimental investigation on ride comfort and rolling motion of high-speed train equipped with MR dampers.* IEEE access, 2021. **9**: p. 118113-118123.
171. Amador-Jimenez, L. and A. Christopher. *A comfort index for public transportation: Case study of Montreal*. in *2016 IEEE international conference on intelligent transportation engineering (ICITE)*. 2016. IEEE.
172. Kalchbrenner, N., E. Grefenstette, and P. Blunsom, *A convolutional neural network for modelling sentences.* arXiv preprint arXiv:1404.2188, 2014.
173. Acharya, U.R., et al., *A deep convolutional neural network model to classify heartbeats.* Computers in biology and medicine, 2017. **89**: p. 389-396.
174. Zhang, X., et al. *Shufflenet: An extremely efficient convolutional neural network for mobile devices*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
175. Awad, M., et al., *Support vector regression.* Efficient learning machines: Theories, concepts, and applications for engineers and system designers, 2015: p. 67-80.

176. Davis, W.J., *The tractive resistance of electric locomotives and cars*. 1926: General Electric.
177. Wu, Q., et al., *Implications of lateral coupler forces for rail vehicle curving resistance.* Journal of Computational and Nonlinear Dynamics, 2021. **16**(3): p. 031002.
178. Webber, B., *Class 323 electric multiple units.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 1999. **213**(1): p. 49-62.
179. Kwon, S., et al. *Validation of heart rate extraction through an iPhone accelerometer*. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011. IEEE.
180. Kos, A., S. Tomažič, and A. Umek, *Evaluation of smartphone inertial sensor performance for cross-platform mobile applications.* Sensors, 2016. **16**(4): p. 477.
181. Breiman, L., *Random forests.* Machine learning, 2001. **45**: p. 5-32.
182. Hesterberg, T., *Bootstrap.* Wiley Interdisciplinary Reviews: Computational Statistics, 2011. **3**(6): p. 497-526.
183. Chen, T. and C. Guestrin. *Xgboost: A scalable tree boosting system*. in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016.
184. Peterson, L.E., *K-nearest neighbor.* Scholarpedia, 2009. **4**(2): p. 1883.
185. Zuo, W., D. Zhang, and K. Wang, *On kernel difference-weighted k-nearest neighbor classification.* Pattern Analysis and Applications, 2008. **11**: p. 247-257.
186. Zhao, L., et al., *T-gcn: A temporal graph convolutional network for traffic prediction.* IEEE transactions on intelligent transportation systems, 2019. **21**(9): p. 3848-3858.
187. Velickovic, P., et al., *Graph attention networks.* stat, 2017. **1050**(20): p. 10-48550.
188. Zhou, J., et al., *Graph neural networks: A review of methods and applications.* AI open, 2020. **1**: p. 57-81.
189. Franke, M. and J. Degen, *The softmax function: Properties, motivation, and interpretation.* 2023.
190. Mastromichalakis, S., *ALReLU: A different approach on Leaky ReLU activation function to improve Neural Networks Performance.* arXiv preprint arXiv:2012.07564, 2020.
191. Ramachandran, P., B. Zoph, and Q.V. Le, *Searching for activation functions.* arXiv preprint arXiv:1710.05941, 2017.
192. Whang, S.E., et al., *Data collection and quality challenges in deep learning: A data-centric ai perspective.* The VLDB Journal, 2023. **32**(4): p. 791-813.
193. Remennikov, A.M. and S. Kaewunruen, *A review of loading conditions for railway track structures due to train and track vertical interaction.* Structural Control and Health Monitoring: The Official Journal of the International Association for Structural Control and Monitoring and of the European Association for the Control of Structures, 2008. **15**(2): p. 207-234.
194. Wu, Q., M. Spiryagin, and C. Cole, *Longitudinal train dynamics: an overview.* Vehicle System Dynamics, 2016. **54**(12): p. 1688-1714.
195. Akiba, T., et al. *Optuna: A next-generation hyperparameter optimization framework*. in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 2019.

196. Hutter, F., H.H. Hoos, and K. Leyton-Brown. *Sequential model-based optimization for general algorithm configuration*. in *Learning and Intelligent Optimization: 5th International Conference, LION 5, Rome, Italy, January 17-21, 2011. Selected Papers 5*. 2011. Springer.
197. Snoek, J., H. Larochelle, and R.P. Adams, *Practical bayesian optimization of machine learning algorithms.* Advances in neural information processing systems, 2012. **25**.
198. Bergstra, J., et al., *Hyperopt: a python library for model selection and hyperparameter optimization.* Computational Science & Discovery, 2015. **8**(1): p. 014008.
199. Koch, P., et al. *Autotune: A derivative-free optimization framework for hyperparameter tuning*. in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 2018.
200. Golovin, D., et al. *Google vizier: A service for black-box optimization*. in *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. 2017.
201. Liaw, R., et al., *Tune: A research platform for distributed model selection and training.* arXiv preprint arXiv:1807.05118, 2018.
202. Seijger, C., et al., *Socially robust knowledge in coastal projects.* Environmental Science & Policy, 2016. **55**: p. 393-407.

Appendices

Published Works

First author:

1. Huang, Junhui & Kaewunruen, Sakdirat. (2022). Train-ride quality evaluation of the Elizabeth Line using machine learning. Frontiers in Built Environment. 8. 10.3389/fbuil.2022.1034433.
2. Huang, Junhui & Yin, Xiaojie & Kaewunruen, Sakdirat. (2022). Quantification of Dynamic Track Stiffness Using Machine Learning. IEEE Access. 10. 1-1. 10.1109/ACCESS.2022.3191278.
3. Huang, Junhui & Kaewunruen, Sakdirat. (2021). PCEvaluation with Machine Learning. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3139465.
4. Huang, Junhui & Kaewunruen, Sakdirat. (2023). Forecasting Energy Consumption of a Public Building Using Transformer and Support Vector Regression. Energies. 16. 966. 10.3390/en16020966.
5. Huang, Junhui & Algahtani, Mohammed & Kaewunruen, Sakdirat. (2022). Energy Forecasting in a Public Building: A Benchmarking Analysis on Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost) Networks. Applied Sciences. 12. 9788. 10.3390/app12199788.

6. Huang, Junhui & Kaewunruen, Sakdirat & Ning, Jingzhiyuan. (2022). AI-Based Quantification of Fitness Activities Using Smartphones. Sustainability. 14. 690. 10.3390/su14020690.

Co-author:

1. Kaewunruen, Sakdirat & Adesope, Abdullah & Huang, Junhui & You, Ruilin & Li, Dan. (2024). AI-based technology to prognose and diagnose complex crack characteristics of railway concrete sleepers. Discover Applied Sciences. 6. 10.1007/s42452-024-05880-8.
2. Fu, Hao & Huang, Junhui & Kaewunruen, Sakdirat. (2023). Mechanical performance of Bio-Inspired Gyroid and Primitive concrete structures under combined compression and torsion Loads: A discrete element method study. Engineering Structures. 291. 116429. 10.1016/j.engstruct.2023.116429.
3. Kaewunruen, Sakdirat & Sresakoolchai, Jessada & Huang, Junhui & Zhu, Yingyu & Ngamkhanong, Chayut & Remennikov, Alex. (2022). MLBased Design of Railway Prestressed Concrete Sleepers. Applied Sciences. 12. 10311. 10.3390/app122010311.
4. Kaewunruen, Sakdirat & Sresakoolchai, Jessada & Huang, Junhui & Harada, Satoru & Wisetjindawat, Wisinee. (2021). Human Activity Vibrations. Data. 6. 104. 10.3390/data6100104