MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Juiz de Fora
Faculdade de Engenharia
Programa de Pós-Graduação em Engenharia Civil

Renato da Silva Melo

Out-Of-Roundness Damage Wheel Identification in Railway Vehicles Using AutoEncoder Models

Juiz de Fora
2025

Renato da Silva Melo

Out-Of-Roundness Damage Wheel Identification in Railway Vehicles Using AutoEncoder Models

Dissertation presented to the Postgraduate Program in Civil Engineering at the Federal University of Juiz de Fora as part of the requirements for obtaining the title of Master's Degree in Civil Engineering.

Date of approval: 19/03/2025

Area of Concentration: Structures and Materials
Research Line: Structural Mechanics.
Advisor: Prof. Dr. Alexandre Abrahão Cury
Co-Advisor: Prof. Dr. Flávio de Souza Barbosa
Co-Advisor: Dr. Rafaelle Piazzaroli Finotti Amaral

Juiz de Fora
2025

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL (PEC)

Renato da Silva Melo

Título: *"Out-of-roundness damage wheel identification in railway vehicles using autoencoder models."*

Dissertação apresentada ao Programa de Pós - Graduação em Engenharia Civil da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Civil. Área de concentração: Estruturas e Materiais.

Aprovada em 19/03/2025.

BANCA EXAMINADORA

Prof. Dr. Alexandre Abrahao Cury - Orientador e Presidente da Banca
Universidade Federal de Juiz de Fora

Prof. Dr. Flávio de Souza Barbosa- Coorientador
Universidade Federal de Juiz de Fora

Profa. Dra. Rafaelle Piazzaroli Finotti Amaral - Coorientadora
Universidade Federal de Juiz de Fora

Profa. Dra. Paula de Oliveira Ribeiro - Membro titular interno
Universidade Federal de Juiz de Fora

Prof. Dr. Rafael Holdorf Lopez- Membro titular externo
Universidade Federal de Santa Catarina

Juiz de Fora, 20/03/2025.

Documento assinado eletronicamente por **Rafael Holdorf Lopez**, **Usuário Externo**, em 20/03/2025, às 15:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Alexandre Abrahao Cury**, **Professor(a)**, em 20/03/2025, às 15:14, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Paula de Oliveira Ribeiro**, **Professor(a)**, em 26/03/2025, às 14:58, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Flavio de Souza Barbosa**, **Professor(a)**, em 27/03/2025, às 17:27, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rafaelle Piazzaroli Finotti Amaral**, **Usuário Externo**, em 28/03/2025, às 20:50, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **2304400** e o código CRC **FB5DE66D**.

ACKNOWLEDGMENTS

I would first like to thank God and my parents, Zilda and Alecir, for all their support throughout this journey. You always believed in this dream and never spared any effort to make it come true. You will always be with me.

I would also like to express my deep gratitude to my girlfriend, Clara, whose support was essential for the completion of this work. Your love, patience, and encouragement were a constant source of motivation. In challenging moments, your presence brought comfort, continuously reminding me of the purpose of this project.

To my advisor, Alexandre Cury, who masterfully guided the entire supervision process, providing much more than technical-scientific direction. Your dedication, vast knowledge, and critical perspective were essential to the development of this work. My sincerest gratitude for all the trust, encouragement, and for believing in my potential.

To my co-advisors, Flávio Barbosa and Rafaelle Finotti, whose guidance and wisdom were fundamental to the success of this study. Each contribution, marked by excellence and precision, not only elevated the quality of this work but also expanded my academic and professional vision, challenging me to grow at every step of this journey.

To CNPq (grant 407256/2022-9) and Rumo Logística S.A. for funding this study, which was essential for its development and for allowing me to continue in the program.

I am grateful to the Federal University of Juiz de Fora for being part of making this dream come true, always upholding free, high-quality public education.

Finally, my gratitude also goes to everyone who, in some way, contributed to the completion of this work, especially the members of this examination committee.

ABSTRACT

The present study investigates the use of unsupervised Machine Learning (ML) techniques applied to the Structural Health Monitoring (SHM) of railway wheels with different stages of polygonalization in the running band. The proposed methodology is based on a comparative study of different AutoEncoder (AE) architectures, including the Variational AutoEncoder (VAE), the Sparse AutoEncoder (SAE), and the Convolutional AutoEncoder (CAE). The models are trained using vertical acceleration data from a virtual track monitoring system, and their performance is evaluated in distinguishing between normal and abnormal structural conditions, as well as in the processing time of input data. The integration of AE models with Hotelling's T^2 control charts is explored as a strategy to enhance anomaly detection. The study highlights the potential of these approaches for predictive maintenance applications, contributing to greater efficiency and safety in railway operations.

Keywords: Structural Health Monitoring, Railways, Damage Detection, Out-Of-Roundness, Sparse AutoEncoder, Convolutional AutoEncoder, Variational AutoEncoder.

RESUMO

O presente estudo investiga o uso de técnicas de Aprendizado de Máquina não supervisionado (ML, do inglês "*Machine Learning*") aplicadas ao Monitoramento da Integridade Estrutural (SHM, do inglês "*Structural Health Monitoring*") de rodas ferroviárias com distintos estágios de poligonalização na banda de rodagem. A metodologia proposta baseia-se em um estudo comparativo entre diferentes arquiteturas de AutoEncoders (AE), incluindo o AutoEncoder Variacional (VAE, do inglês "*Variational AutoEncoder*"), o AutoEncoder Esparso (SAE, do inglês "*Sparse AutoEncoder*") e o AutoEncoder Convolucional (CAE, do inglês "*Convolutional AutoEncoder*"). Os modelos são treinados com dados de aceleração vertical provenientes de um sistema virtual de monitoramento em via, e seus desempenhos são avaliados na diferenciação entre condições estruturais normais e anormais, e no tempo de processamento dos dados de entrada. A integração dos modelos AE com gráficos de controle de Hotelling T^2 é explorada como uma estratégia para aprimorar a detecção de anomalias. O estudo destaca o potencial dessas abordagens para aplicações de manutenção preditiva, contribuindo para maior eficiência e segurança na operação ferroviária.

Palavras-chave: Monitoramento de Integridade Estrutural, Ferrovias, Detecção de danos, Ovalização, Autocodificador Esparso, Autocodificador Convolucional, Autocodificador Variacional.

LIST OF ABBREVIATIONS AND ACRONYMS

AE	AutoEncoder
AI	Artificial Intelligence
ANN	Artificial Neural Network
ARX	Exogenous Autoregressive Model
CAE	Convolutional AutoEncoder
CNN	Convolutional Neural Network
COMAC	Coordinate Modal Assurance Criterion
CPU	Central Processing Units
CWT	Continuous Wavelet Transform
DLM	Linear Dynamic Bayesian Model
FEM	Finite Element Method
GAN	Generative Adversarial Network
GPU	Graphical Processing Unit Central Processing Units (CPUs),
IoT	Internet of Things
IWS	Instrumented Wheelset
KL	Kullback-Leibler
MAC	Modal Assurance Criterion
MAE	Mean Absolute Error
ML	Machine Learning
MSE	Mean Squared Error
OOR	Out-Of-Roundness
ORSR	Original to Reconstructed Signal Ratio
PCA	Principal Component Analysis
PC-GAN	Physics-Constrained Generative Adversarial Network
PI-GAN	Physics-Informed Generative Adversarial Network
SAE	Sparse AutoEncoder
SDA	Symbolic Data Analysis
SHM	Structural Health Monitoring
TPE	Tree-structured Parzen Estimator
UCL	Upper Control Limit
VAE	Variational AutoEncoder
VBI	Vehicle-Bridge Interaction
WILD	Wheel Impact Load Detector
WSN	Wireless Sensor Network

SUMMARY

1 INTRODUCTION

1.1 Contextualization

Since its beginning, railway transportation has played a vital role in improving the living conditions of developing societies. This mode of transportation is characterized by its efficiency and versatility, serving both passenger and freight transport, covering different cargo volumes and distances [1,2]. The growing need for efficient transportation between countries has become a central challenge for developed nations, as modern and sustainable transportation systems are essential for structuring advanced economies [3].

In the context of an increasingly interconnected world, the continuous upgrading and expansion of railway networks are fundamental, given their relevance to the economic progress of these nations [1,3]. In the United States, the total length of the railway network exceeds 250.000 km [3], while in Brazil, this value is approximately 30.660 km, corresponding to 15.1% of the country's transportation modal share [4].

The railway mode offers numerous advantages compared to road and waterway transportation, such as high cargo capacity, lower cost per ton transported, safety, and the long service life and durability of the infrastructure. These factors have driven the expansion of railway systems globally over recent years. However, a significant portion of existing railway networks were designed for demands that are vastly different from current ones, highlighting the need for constant monitoring and modernization of the infrastructure [5]. Currently, railways face challenges associated with higher axle loads, increased speeds, and greater frequency of use [3]. In addition to operational demands, railway infrastructure is significantly impacted by climate change, facing risks such as erosion, landslides, and flooding, as well as other biophysical effects. These issues lead to service disruptions, increased operational costs, and the compromise of the safety and durability of the entire railway structure [4].

Although rails are one of the main components of railway systems, whose monitoring and interventions are of greater interest to concessionaires, the structural condition of the wheels plays a fundamental role in the safety of railway vehicle operations. Wheels deteriorate over time due to wear and fatigue, potentially developing various defects, such as Out-Of-Roundness (OOR), commonly represented by polygonal wheels and wheel flats, which substantially alter the wheel-rail contact characteristics [1,6]. These scenarios are illustrated in Figure 1, where (a) represents a wheel flat and (b) depicts a polygonal wheel. Wheel defects are identified as one of the main causes of railway accidents [7,8], and although polygonal wheels generate

fewer negative effects on the track compared to wheel flats [9], these defects can cause severe damage to both the track and the vehicle due to the excitation induced by such phenomena. These defects also increase noise emissions inside and outside the vehicle and considerably reduce railway safety and comfort levels.

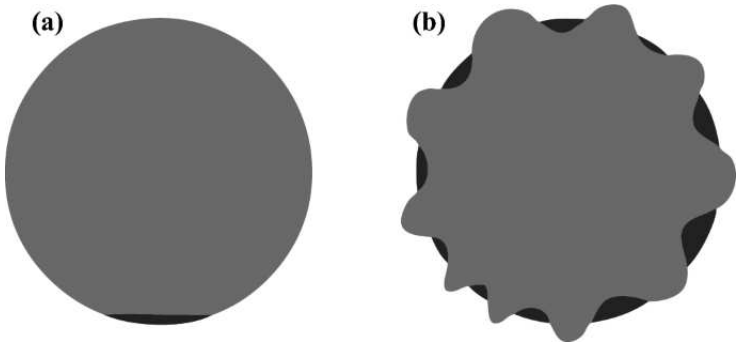

Figure 1. Common Out-Of-Roundness (OOR) defects in railway wheels: (a) wheel flat and (b) polygonal wheel.

The wheel-rail impact loads caused by wheel defects exhibit significantly higher values than those observed under normal operating conditions [10,11]. These high impacts substantially contribute to the accelerated wear of both vehicle components and railway infrastructure, inducing issues such as abnormal vibrations in axles, rail abrasion, and fractures in fasteners. To mitigate these effects, railway operation departments frequently implement preventive measures and carry out periodic renewals of defective elements. Among the preventive measures, the adoption of advanced anti-slip systems in most passenger trains stands out, as they help reduce the frequency of wheel-rail slips, one of the main causes of wheel flats [12,13]. However, in scenarios with high operating speeds and increased axle loads, the occurrence of wheel defects still persists. In contrast, freight trains, which often lack anti-slip technologies and are subjected to more severe traffic conditions, present even more critical wheel conditions, significantly compromising railway infrastructure [10].

Given this context, the need to implement effective methodologies for monitoring railway infrastructures and vehicles arises. This monitoring aims to assess operational conditions and safety limits, as well as to identify faults or defects at early stages, seeking to minimize damage and its consequent impacts [14,15]. In this regard, Structural Health Monitoring (SHM) emerges as a promising technological alternative, enabling the automatic and continuous assessment of the structural integrity of various railway components, including wheels.

SHM is a field of study dedicated to the continuous evaluation of structural conditions, aiming to ensure performance, safety, and integrity. This methodology relies on a comprehensive set of technologies, equipment, and data analysis techniques, often integrated with predictive models, to effectively monitor and assess the state of structures [16]. The implementation of SHM systems enables more efficient management of repairs and maintenance, establishing itself as an essential tool in multiple application fields. Traditionally, SHM relied on visual inspections and non-destructive techniques, processes that, although effective, present limitations related to high costs and long durations. However, advancements in computational technologies and data acquisition systems have transformed this landscape, allowing SHM methodologies to process and analyze a significant volume of information more quickly and accurately [16]. Consequently, techniques such as Machine Learning (ML) have emerged as potential solutions to enhance the effectiveness of these systems.

Among the Machine Learning techniques applied to SHM, AutoEncoders (AEs) [17] stand out, demonstrating great potential in identifying anomalous patterns and detecting signs of structural degradation or damage. AEs are a type of unsupervised Artificial Neural Network (ANN) specialized in learning compact and efficient data representations by compressing information into a lower-dimensional space and subsequently reconstructing it. This dimensionality reduction capability, combined with appropriate evaluation metrics, enables the identification of structural changes that indicate the presence of damage. Thus, AEs constitute a robust tool for SHM applications, contributing to the optimization of monitoring processes and the maintenance of monitored infrastructures.

Although significant advances have been made in the use of AEs for SHM strategies, there is a notable gap in research exploring their application in monitoring railway wheels for defects associated with Out-of-Roundness. This study aims to address this gap by investigating the performance of Sparse AutoEncoder (SAE), Variational AutoEncoder (VAE), and Convolutional AutoEncoder (CAE) models in identifying and quantifying OOR wheels. To enhance the effectiveness of the analyses, the integration of AE models with Hotelling's T^2 Control Chart is proposed, a technique that allows for the quantification and monitoring of structural changes in railway wheels subjected to different levels of damage associated with OOR and defect evolution.

It is noteworthy that this study is part of the CNPq/MCT Project 407256/2022-9 in partnership with Rumo Logística S.A., which aims to develop SHM strategies for railways.

1.2 Objectives

The main objective of the study presented in this dissertation is to conduct a comparative analysis of three AutoEncoder models—Variational Autoencoder, Sparse Autoencoder, and Convolutional Autoencoder—to detect and quantify structural anomalies associated with the Out-of-Roundness in railway vehicle wheels. The expected results include the identification of a robust and effective methodology for assessing the structural integrity of the wheels, contributing to improved safety, optimization of maintenance and repair procedures, and cost reduction associated with these activities.

As specific objectives, complementary to the main objective, we have:

- Analysis of optimized SAE, CAE, and VAE models capable of learning high-quality latent representations from simulated railway vibration data.
- Estimation of the computational cost, in terms of processing time, to assess the feasibility of using the considered AutoEncoder methodologies.
- Conducting a comparative analysis in terms of input data reconstruction capability and computational cost in terms of processing time to evaluate the feasibility of using the Autoencoder methodologies considered in the study.

1.3 Study Structure

This study is organized into four chapters, which are briefly described below:

- **Chapter 1**: Explores the importance of the addressed topic, the motivation that drove the realization of this study, and details the objectives set to be achieved by the end of the research.
- **Chapter 2**: Presents a preliminary review of the literature, analyzing several relevant studies on SHM strategies. Then, the focus is directed to the railway context and to ORR-type wheels, highlighting their contributions to the deepening of the discussion on the topic and the identification of the existing gap that motivated the development of this study.
- **Chapter 3**: Presents the paper entitled *"Out-of-roundness damage wheel identification in railway vehicles using autoencoder models"*, accepted for publication in the journal *Applied Sciences*.
- **Chapter 4**: Presents the final considerations and proposes future studies that would contribute to the development of SHM strategies in the railway context.

1.4 References

[1] GUEDES, A.; SILVA, R.; RIBEIRO, D.; VALE, C.; MOSLEH, A.; MONTENEGRO, P.; MEIXEDO, A. Detection of wheel polygonization based on wayside monitoring and artificial intelligence. **Sensors**. v. 23, n. 4, 2188, 2023.

[2] LAGNEBACK, R. **Evaluation of wayside condition monitoring technologies for condition-based maintenance of railway vehicles.** 138 p. Thesis (Ph.D.) - Luleå University of Technology, Sweden, 2007.

[3] MALEKJAFARIAN, A.; OBRIEN, E.J.; QUIRKE, P.; CANTERO, D.; GOLPAYEGANI, F. Railway track loss-of-stiffness detection using bogie filtered displacement data measured on a passing train. **Infrastructures**. v. 6, n. 6, 93, 2021.

[4] BRASIL. Ministério da Ciência, Tecnologia e Inovação. **Infraestrutura ferroviária**. 2024. Disponível em: https://adaptabrasil.mcti.gov.br/. Acesso em: 27 jan. 2025.

[5] GUNN, D.A.; CHAMBERS, J.E.; DASHWOOD, B.E.; LACINSKA, A.; DIJKSTRA, T.; UHLEMANN, S.; SWIFT, R.; KIRKHAM, M.; MILODOWSKI, A.; WRAGG, J.; DONOHUE, S. Deterioration model and condition monitoring of aged railway embankment using non-invasive geophysics. **Constr. Build. Mater.** v. 170, p. 668 – 678, 2018.

[6] NIELSEN, J.C.O.; JOHANSSON, A. Out-of-round railway wheels-a literature survey. **Proc. of the Inst. of Mechanical Engineers, Part F: J. of Rail and Rapid Transit**. v. 214, n. 2, p. 79 – 91, 2000.

[7] CHONG, S.Y.; LEE, J.R.; SHIN, H.J. A review of health and operation monitoring technologies for trains. **Smart Struct. Syst.** v. 6, n. 9, p. 1079 – 1105, 2010.

[8] VALE, C. Wheel flats in the dynamic behavior of ballasted and slab railway tracks. **Applied Sciences.** v. 11, n. 15, 7127, 2021.

[9] WU, X.; CHI, M.; WU, P. Influence of polygonal wear of railway wheels on the wheel set axle stress. **Veh. Syst. Dyn.** v. 53, n. 11, p. 1535 – 1554, 2015.

[10] ZHOU, C.; GAO, L.; XIAO, H.; HOU, B. Railway wheel flat recognition and precise positioning method based on multisensor arrays. **Applied Sciences.** v. 10, n. 4, 1297, 2020.

[11] BIAN, J.; GU, Y.; MURRAY, M.H. A dynamic wheel–rail impact analysis of railway track under wheel flat by finite element analysis. **Veh. Syst. Dyn.** v. 51, n. 6, p. 784 – 797, 2013.

[12] ZUO, J.; WU, M. Research on anti-sliding control of railway brake system based on adhesion-creep theory. **In Proceedings of the IEEE International Conference on Mechatronics & Automation**, Xi'an, China, p. 1690-1694, 2010.

[13] LUO, R. Anti-sliding control simulation of railway vehicle braking. **Chin. J. Mech. Eng.** v. 44, n. 3, p. 29-34, 2008.

[14] ALVES, V.; MEIXEDO, A.; RIBEIRO, D.; CALÇADA, R.; CURY, A. Evaluation of the performance of different damage indicators in railway bridges. **Procedia Eng.** v. 114, p. 746-753, 2015.

[15] VALE, C.; BONIFÁCIO, C.; SEABRA, J.; CALÇADA R.; MAZZINO, N.; ELISA, M.; TERRIBILE, S.; ANGUITA, D.; FUMEO, E.; SABORIDO, C.; VANHONACKER, T.; DONDER, E.; LAEREMANS, M.; VERMEULEN, F.; GRIMES, D. Novel efficient technologies in Europe for axle bearing condition monitoring—The MAXBE Project. **Transp. Res. Procedia.** v. 14, p. 635-644, 2016.

[16] DIAS JÚNIOR, L. T. **Uso de autocodificadores variacionais para a detecção de danos estruturais.** 111 p. Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Juiz de Fora, Brasil, 2022.

[17] FINOTTI, R.P.; GENTILE, C.; BARBOSA, F.d.S.; CURY, A.A. Structural novelty detection based on sparse autoencoders and control charts. **Structural Engineering and Mechanics.** v. 81, n. 5, p. 647 – 664, 2022.

2 BIBLIOGRAPHICAL REVIEW

In recent years, SHM methodologies have advanced significantly, driven by the incorporation of ML techniques. The integration of these approaches has led to substantial improvements in the detection and assessment of structural damage in various types of infrastructure. Among these techniques, AutoEncoders have emerged as a promising solution, providing greater efficiency and autonomy in structural monitoring processes. This literature review explores the concept of damage detection, key studies and the techniques developed in this field over time, subsequently giving particular focus to its application in railway engineering and in the detection of OOR wheels.

2.1 Damage Detection

The concept of damage, in general, refers to a change or alteration occurring in a structure, which negatively impacts its current or future performance. This definition assumes that damage can only be quantified or assessed by comparing two distinct states of the structure: an initial state, considered intact or as a reference, and a subsequent state, which reflects the presence of alterations or degradations [18].

Most damage detection methods are based on the premise that the occurrence of damage in a structural system leads to changes in its physical properties, such as stiffness, mass, or energy dissipation capacity. These changes, in turn, affect the dynamic properties of the structure, such as natural frequencies, vibration modes, and damping [19]. However, mass variation tends to be insignificant compared to stiffness loss and can therefore be disregarded in most cases, as demonstrated in the studies by Adams *et al.* (1978) [20] and Hearn & Testa (1991) [21].

Thus, conducting dynamic tests at different moments throughout the service life of a structure allows for monitoring possible changes in its condition. In other words, if changes in the structure's dynamic pattern are observed at regular time intervals, this may indicate the existence of a problem. In this sense, identifying behavioral patterns plays an important role in structural integrity assessment, as the presence of damage directly affects the structure's modal characteristics. For this process to be effective, the precise extraction of these parameters is essential, as inaccuracies in the data may compromise damage detection and diagnosis [19].

The application of damage detection methods in real structures is often limited by practical factors. The spatial discretization required to capture localized phenomena, combined with the inevitable presence of noise in data collection, demands the development of robust signal processing techniques. Additionally, the appropriate selection of sensors and the definition of an effective instrumentation strategy are critical for obtaining reliable data, especially in large-scale structures with complex geometries [22].

Methods Based on Neural Networks and Artificial Intelligence

Conventional methodologies for damage detection in structures, based on time-domain analysis of modal characteristics, have demonstrated their effectiveness in various applications, as previously discussed. However, as also observed, these techniques have presented some limitations. The main drawback lies in their sensitivity to noise and variations in environmental conditions, which can lead to false positives. Additionally, the modal identification process, by filtering the data, may result in the loss of relevant information for an accurate assessment of the structure's condition [16].

Given these limitations, with the advent of new technologies and increased computational capacity, the development of innovative approaches for damage detection has been driven forward. Artificial Neural Networks and genetic algorithms, for example, have shown great potential in this context [23]. ANNs can learn complex and nonlinear patterns in the data, allowing damage detection in less detailed structural models [19]. However, training these networks requires a representative dataset covering different damage conditions, which can be a challenge in practical applications.

In recent studies on SHM methodologies, techniques based on Machine Learning have been widely explored due to their robustness in analyzing large volumes of data and automatically extracting relevant features. ML encompasses neural architectures composed of multiple layers, enabling a hierarchical representation of data and enhancing the detection of complex patterns. These approaches offer significant advantages, such as greater resistance to high-dimensional data and the ability to generalize across different application domains [24]. Wang & Cha (2021) [25], for example, proposed a hybrid methodology that combines AutoEncoders and one-class support vector machines to detect damage in numerical and scaled models. Complementarily, Finotti (2022) [24] implemented Sparse AutoEncoders for unsupervised detection of structural changes in different experimental contexts, including a monitored laboratory frame, the Z24 bridge, and an instrumented tower in Italy, achieving satisfactory results in

anomaly identification. Pollastro *et al.* (2023) [26] explored semi-supervised methods based on dynamic data for structural irregularity detection, employing ANNs and Variational AutoEncoders. Meanwhile, Spínola Neto *et al.* (2024) [27] analyzed the effectiveness of four AutoEncoder architectures (AE, SAE, CAE, and VAE) in combination with Hotelling's T^2 control chart for identifying and quantifying structural variations in different engineering systems.

Considering the relevance and promising results of techniques based on ANNs and AEs, which stand out among ML methods applied to structural change detection through vibration analysis, this work adopts SAE, VAE, and CAE architectures for developing an approach in the SHM context. The choice of these models is based on their proven effectiveness, as observed in the aforementioned studies, and their ability to capture structural patterns and detect subtle changes in the dynamic behavior of structures, contributing to the advancement of structural monitoring and diagnostic methodologies.

2.2 State of the Art of SHM Strategies in the Railway Context

Railway instrumentation dates back to the 19th century, with the development of the first track monitoring circuit by William Robinson in 1872, a system still widely used today [28]. Since then, sensor technology has evolved significantly, enabling the precise monitoring of parameters such as track deflection and variations in superstructure stiffness, which were previously assessed using invasive methods [29]. Although modern and non-destructive systems, such as the track inspection train, provide detailed data, their high acquisition and maintenance costs limit large-scale application [30]. Consequently, the demand for more accessible and efficient solutions is increasing, driving research on distributed sensors and Artificial Intelligence (AI)-based monitoring systems.

Traditional railway monitoring methods include visual inspections and test vehicles equipped with sensors, allowing the identification of both apparent and non-apparent structural failures [31,32,33]. However, these approaches have limitations, such as evaluator subjectivity, high operational costs, and the need to interrupt railway operations [34,35,36]. Additionally, they are not part of continuous monitoring strategies that provide real-time data, reinforcing the need for more efficient alternatives [37,38]. In this context, technological advancements have enabled the development of smaller, more accurate, and cost-effective sensors [28], fostering research on Structural Health Monitoring (SHM), particularly in high-speed railway infrastructure [29]. As a result, current investigations explore SHM applied to tracks, wheels, bridges, and tunnels, aiming for greater efficiency and reliability in structural monitoring.

Although railway wheels have historically received less attention than other structural components, recent studies highlight their relevance to the safety and efficiency of railway systems. Advances in monitoring technologies have enabled the detection and characterization of failures in these components, contributing to predictive maintenance and track integrity. Guedes *et al.* (2023) [1] investigated the identification of polygons in railway wheels using artificial intelligence, analyzing the dynamic responses induced in the track by the passage of freight wagons. The study compared feature extraction techniques, such as the AutoRegressive with Exogenous input (ARX) model and Continuous Wavelet Transform (CWT), using Principal Component Analysis (PCA) to reduce external interference. Additionally, a data fusion model based on Mahalanobis distance and outlier analysis was implemented to distinguish wheels in normal conditions from those with defects.

Complementing these advances, Jorge *et al.* (2024) [39] developed an unsupervised artificial intelligence-based method for the early detection of wear phenomena such as polygonization and flats in railway wheels. The system analyzes dynamic responses generated in the tracks by freight railway vehicles, covering stages such as data collection and preprocessing, feature extraction, and information fusion. The methodology stands out for using a Stacked Sparse AutoEncoder for feature compression and extraction, combined with Mahalanobis distance to enhance fault detection accuracy. The results indicate significant potential for reducing maintenance costs and increasing operational safety in the railway sector.

Based on the previously discussed studies and supported by prior investigations, it has been found that the use of ML techniques has significantly expanded within SHM applications in the railway sector. Given the promising results obtained so far, this research aims to employ SAE, VAE, and CAE autoencoders in SHM analysis of railway wheels with different stages of polygonal wear. These algorithms are recognized for their ability to provide greater autonomy and accuracy to Structural Health Monitoring methods based on AI. Table 1 presents a summary of some of the main studies developed so far in the context of monitoring and identifying OOR wheels.

Table 1: Summary of the main studies involving SHM techniques in railway wheels.

Year	**Authors**	**Contribution**
1988	Kaku & Yamashita [40]	Investigation of the human response to impact noise generated by flattened wheels and rail joints of electric trains in a laboratory experiment. Compared the sound of the train

		noise containing impact components with the noise of the train without impact components.
1991	Kumagai *et al.* [41]	Pioneering study on the identification of preventive measures against wheel flattening, including the influence of factors affecting adhesion and the relationship with mileage. Proposal of a new classification to assess the severity of the problem.
1997	Bracciali & Cascini [42]	Development and validation of an original experimental and numerical procedure for detecting corrugation and wheel misalignment in trains through the processing of track acceleration signals using combined energy criteria and cepstral analysis.
1999	Morys [43]	Analysis of the impact of radial wheel deviations on the short-term dynamics of a high-speed wagon, considering its interaction with the track. Identification of the impact of OOR wheel shapes on normal force and their potential effect on vertical wheel acceleration and wear.
2000	Nielsen & Johansson [6]	State-of-the-art analysis of the causes and consequences of railway wheels with OOR defects.
2003	Johansson & Nielsen [44]	Analysis of the influence of different types of wheel out-of-roundness (OOR) on the vertical wheel-rail dynamic contact force and track response.
2007	Stratman *et al.* [45]	Proposal of two quantitative criteria for the removal of wheels from service, with SHM trends in real-time, developed using data collected from trains in operation. The data are gathered using Wheel Impact Load Detectors (WILDs).
2012	Wei *et al.* [46]	Development of a real-time monitoring system for defects in railway wheels using fiber Bragg grating sensors. Measurement and processing of rail track deformation response during wheel-rail interaction to generate a condition index that directly reflects wheel status.
2017	Jing & Han [47]	Detailed numerical simulation of the dynamic wheel-rail interaction response due to the impact of a flat defect on the

		wheel, considering a three-dimensional rolling model. Factors such as wheel and rail structural inertia, material deformation rate, and thermal stress due to sliding friction were considered.
2020	Wang *et al.* [48]	Development of a methodology for detecting defects in high-speed train wheels in real-time, based on the Linear Dynamic Bayesian Model (DLM). The methodology includes logic for: (i) prognosis; (ii) detection of possible outliers; (iii) identification of occurrence.
2021	Ni *et al.* [49]	Development of a probabilistic Bayesian method for online quantitative assessment of wheel conditions in railways, using track stress monitoring data.
2023	Guedes *et al.* [1]	Presentation of a methodology for detecting polygonal wheels in freight trains based on artificial intelligence techniques ARX model and CWT, using dynamic responses from a train-track interaction model that simulates train passage over a set of accelerometers installed on the rail and sleepers.
2024	Jorge *et al.* [39]	Identification and classification of defects using an unsupervised methodology for detecting flat and polygonal wheels in freight rail vehicles, based on numerical dynamic responses.

Based on the thorough review conducted, it is observed that although a wide range of studies have been carried out on OOR wheels and, specifically, on polygonalization, there remains a gap to be explored in this context. Therefore, this study presents a comprehensive comparative analysis of the main AutoEncoder models for damage identification in OOR wheels. The primary objective is to provide researchers with a clear reference for selecting the most suitable model for this application. While several studies have investigated individual autoencoder models, there is a lack of direct comparisons between them, making it challenging for researchers and professionals to determine the most effective approach.

2.3 References

[18] SOHN, H.; FARRAR, C. R.; HEMEZ, F. M.; CZARNECKI, J. J. **A review of structural health review of structural health monitoring literature 1996-2001.** Los Alamos National Laboratory Report LA-13976-MS, Los Alamos, New Mexico, EUA, 2003.

[19] MIGUEL, L. F. F. **Identificação de sistemas e avaliação da integridade de estruturas treliçadas.** 184 p. Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

[20] ADAMS, R.; CAWLEY, P.; PYE, C.J.; STONE, B. A vibration technique for non-destructively assessing the integrity of structure. **Journal Mechanical Engineering Science.** v. 20, n. 2, p. 93 – 100, 1978.

[21] HEARN, G.; TESTA, R.B. Modal analysis for damage detection in structures. **Journal of Structural Engineering.** v. 117, n. 10, p. 3042 – 3063, 1991.

[22] SOHN, H. Effects of environmental and operational variability on structural health monitoring. **PhilosophicalTransactions of the Royal Society a: Mathematical, Physical and Engineering Sciences.** v. 365, n. 1851, p. 539 – 560, 2006.

[23] CURY, A. **Techniques d'anormalité appliquées à la surveillance de santé structurale.** 392 p. Thesis (PhD), Université Paris-Est, France, 2010.

[24] FINOTTI, R. P. **Inteligência artificial aplicada ao monitoramento de estruturas: detecção de alterações mecânico-estruturais baseada no uso de redes neurais autocodificadoras esparsas para a caracterização de respostas dinâmicas.** 169 p. Tese (Doutorado) — Universidade Federal de Juiz de Fora, Juiz de Fora, Brasil, 2022.

[25] WANG, Z.; CHA, Y. J. Unsupervised deep learning approach using a deep auto-encoder with a one-class support vector machine to detect damage. **Structural Health Monitoring.** v. 20, n. 1, p. 406 – 425, 2021.

[26] POLLASTRO, A.; TESTA, G.; BILOTTA, A.; PREVETE, R. Semi-supervised detection of structural damage using variational autoencoder and a one-class support vector machine. **IEEE Access.** vol. 11, p. 67098 - 67112, 2023.

[27] SPÍNOLA NETO, M.; FINOTTI, R.P.; BARBOSA, F.d.S.; CURY, A.A. Structural damage identification using autoencoders: A comparative study. **Buildings.** v. 14, n. 7, p. 2014, 2024.

[28] CASTILLO-MINGORANCE, J. M.; SOL-SÁNCHEZ, M.; MORENO-NAVARRO, F.; RUBIO-GÁMEZ, M. C. A critical review of sensors for the continuous monitoring of smart and sustainable railway infrastructures. **Sustainability**. v. 12. n. 22, 9428, 2020.

[29] KOUROUSSIS, G.; CAUCHETEUR, C.; KINET, D.; ALEXANDROU, G.; VERLINDEN, O.; MOEYAERT, V. (2015). Review of trackside monitoring solutions: From strain gages to optical fibre sensors. **Sensors**, v. 15, n. 8, p. 20115 – 20139, 2015.

[30] KOUROUSSIS, G.; KINET, D.; MOEYAERT, V.; DUPUY, J.; CAUCHETEUR, C. Railway structure monitoring solutions using fibre Bragg grating sensors. **International Journal of Rail Transportation,** v. 4, n. 3, p. 135 – 150, 2016.

[31] BIANCHI, G.; FANELLI, C.; FREDDI, F.; GIULIANI, F.; LA PLACA, A. Systematic review railway infrastructure monitoring: From classic techniques to predictive maintenance. **Advances in Mechanical Engineering,** v. 17, n. 1, 16878132241285631, 2025.

[32] UZARSKI, D. R.; GRUSSING, M. N. Beyond mandated track safety inspections using a mission-focused, knowledgebased approach. **International Journal of Rail Transportation**, v. 1, n. 4, p. 218 – 236, 2013.

[33] KAćUNIć, D. J.; LIBRIć, L.; Car, M. Application of unmanned aerial vehicles on transport infrastructure network. **Gradjevinar**, v. 68, n. 4, p. 287 – 300, 2016.

[34] PHARES, B. M.; WASHER, G. A.; ROLANDER, D. D.; GRAYBEAL, B. A.; MOORE, M. Routine highway bridge inspection condition documentation accuracy and reliability. **Journal of Bridge Engineering**, v. 9, n. 4, p. 403 –413, 2004.

[35] STAJANO, F.; HOULT, N.; WASSEL, I.; BENNET, P.; MIDDLETON, C.; SOGA, K. Smart bridges, smart tunnels: Transforming wireless sensor networks from research prototypes into robust engineering infrastructure. **Ad. Hoc. Networks**, v. 8, n. 8, p. 872-888, 2010.

[36] VAGNOLI, M.; REMENYTE-PRESCOTT, R.; ANDREWS, J. Railway bridge structural health monitoring and fault detection: State-of-the-art methods and future challenges. **Structural Health Monitoring**, v. 17, n. 4, p. 971 – 1007, 2018.

[37] CHANDRAN, P.; ASBER, J.; THIERY, F.; ODELIUS, J.; RANTATALO, M. An investigation of railway fastener detection using image processing and augmented deep learning. **Sustainability**, v. 13, n. 21, 12051, 2021.

[38] AVSIEVICH, A.; AVSIEVICH, V.; AVSIEVICH, N.; OVCHINNIKOV, D.; IVASCHENKO, A. Railway track stress–strain analysis using high-precision accelerometers. **Applied Sciences**, v. 11, n. 24, 11908, 2021.

[39] JORGE, T.; MAGALHÃES, J.; SILVA, R.; GUEDES, A.; RIBEIRO, D.; VALE, C.; MEIXEDO, A.; MOSLEH, A.; MONTENEGRO, P.; CURY, A. Early identification of out-of-roundness damage wheels in railway freight vehicles using a wayside system and a stacked sparse autoencoder. **Vehicle System Dynamics**, v. 63, n. 2, p. 232 – 257, 2024.

[40] KAKU, J.; YAMASHITA, M. Impact noise from railroads. **Journal of Sound and Vibration**, v. 120, n. 2, p. 333 – 337, 1988.

[41] KUMAGAI, N.; ISHIKAWA, H.; HAGA, H.; KIGAWA, T.; NAGASE, K. Factors of wheel flats occurrence and preventive measures. **Wear**, v. 144, n. 1, p. 277 – 287, 1991.

[42] BRACCIALI, A.; CASCINI, G. Detection of corrugation and wheel flats of railway wheels using energy and cepstrum analysis of rail acceleration. **Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit,** v. 211, n. 2, p. 109 – 116, 1997.

[43] MORYS, B. Enlargement of out-of-round wheel profiles on high speed trains. **Journal of Sound and Vibration**, v. 227, n. 5, p. 965 – 978, 1999.

[44] JOHANSSON, A.; NIELSEN, J. C. O. Out-of-round railway wheels—wheel-rail contact forces and track response derived from field tests and numerical simulations. **Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit,** v. 217, n. 2, p. 135 – 146, 2003**.**

[45] STRATMAN, B.; LIU, Y.; MAHADEVAN, S. Structural health monitoring of railroad wheels using wheel impact load detectors. **Journal of Failure Analysis and Prevention,** v. 7, n. 3, p. 218 – 225, 2007.

[46] WEI, C.; XIN, Q.; CHUNG, W. H.; LIU, S.; TAM, H.; HO, S. L. Real-time train wheel condition monitoring by fiber bragg grating sensors. **International Journal of Distributed Sensor Networks,** v. 8, n. 1, 409048, 2012.

[47] JING, L.; HAN, L. Further study on the wheel–rail impact response induced by a single wheel flat: the coupling effect of strain rate and thermal stress. **Vehicle System Dynamics,** v. 55, n. 12, p. 1946 – 1972, 2017.

[48] WANG, Y. W.; NI, Y. Q.; WANG, X. Real-time defect detection of high-speed train wheels by using Bayesian forecasting and dynamic model. **Mechanical Systems and Signal Processing,** v. 139, 106654, 2020**.**

[49] NI, Y. Q.; ZHANG, H. Q. A Bayesian machine learning approach for online detection of railway wheel defects using track-side monitoring. **Structural Health Monitoring**, v. 20, n. 4, p. 1536 – 1550, 2021.

3 PAPER – OUT-OF-ROUNDNESS DAMAGE WHEEL IDENTIFICATION IN RAILWAY VEHICLES USING AUTOENCODER MODELS

Article published in Applied Sciences (IF: 2.5; Qualis Capes A3).
DOI: https://doi.org/10.3390/app15052662

Renato Melo[1]; Rafaelle Finotti[1]; António Guedes[2]; Vítor Gonçalves[2]; Andreia Meixedo[2]; Diogo Ribeiro[3]; Flávio Barbosa[1]; Alexandre Cury[1].

1 - Graduate Program in Civil Engineering, Federal University of Juiz de Fora, Juiz de Fora, 36036-900, Brazil
2 - CONSTRUCT-LESE, Faculty of Engineering, University of Porto, Porto, 4200-465, Portugal
3 - CONSTRUCT-LESE, School of Engineering, Polytechnic of Porto, Porto, 4200-465, Portugal.

Abstract

This study presents a comparative analysis of three AutoEncoder (AE) models – Variational AutoEncoder (VAE), Sparse AutoEncoder (SAE), and Convolutional AutoEncoder (CAE) - to detect and quantify structural anomalies in railway vehicle wheels, such as polygonization. Vertical acceleration data from a virtual wayside monitoring system serve as input for training the AE models, which are coupled with Hotelling's T^2 Control Charts to differentiate normal and abnormal railway component behaviors. The results indicate that the SAE-T^2 model outperforms its counterparts, achieving 16.67% higher accuracy than the CAE-T^2 model in identifying distinct structural conditions, although with a 35.78% higher computational cost. Conversely, the VAE-T^2 model is outperformed in 100% of the analyzed scenarios when compared to SAE-T^2 in identifying distinct structural conditions, while also exhibiting a 21.97% higher average computational cost. Across all scenarios, the SAE-T^2 methodology consistently provided better classifications of wheel damage, showing its capability to extract relevant features from dynamic signals for Structural Health Monitoring (SHM) applications. These findings highlight SAE's potential as an interesting tool for predictive maintenance, offering improved efficiency and safety in railway operations.

Keywords: Structural Health Monitoring, Railways, Damage Detection, Out-Of-Roundness, Sparse AutoEncoder, Convolutional AutoEncoder, Variational AutoEncoder.

1. Introduction

Under normal operation, railways are subjected to traffic loads that may lead to geometric wear and tear of the track and its components. When irregularities exceed design thresholds, the track is deemed defective, posing risks of failure or service disruption. Such incidents can result in high maintenance costs, economic losses, material damage, and even accidents that threaten human lives.

Railway wheels, which bear loads, transmit traction, and guide vehicles, are subjected to severe operating conditions that can lead to Out-Of-Roundness (OOR) or polygonization, characterized by tread irregularities along the wheel's circumference [1]. OOR affects wheel-rail interaction by causing sudden changes in normal forces, increasing rolling and impact noise, compromising comfort and safety, and potentially resulting in freight movement issues, wheel or rail cracks, and tie damage [2,3]. Previous research indicates that hot spots, hardening effects from brake blocks, wheel fragmentation, and eccentricity are major contributors to polygonal wear [4,5], and manufacturing processes can also induce eccentricities [3,6]. Furthermore, high-order polygonization at high speeds can significantly amplify axle box acceleration, wheel-rail normal forces, and even cause wheel-rail separation when resonance occurs [7].

Traditionally, verifications related to the assessment of the degradation process of railway structures are performed through visual inspections and manual measurements using devices such as tachometers and leveling instruments. These techniques present significant economic and technical challenges, particularly when applied on a large scale. Moreover, these methods often fail to detect internal and complex anomalies, notably in the case of initial damage [8], such as those associated with internal rail cracks and wheel polygonalization, compromising the safety and efficiency of railway operations [9]. Early detection of these problems is important, as the damage worsens, leading to gradual deterioration of structural functionality, potentially exceeding acceptable safety limits and significantly increasing the likelihood of critical failures or collapses [10].

Automated railway monitoring techniques have significantly advanced in recent years, enhancing safety and efficiency in rail operations. Advances in sensor technology [11] and modal identification [12], which allow for feature extraction from vibration signals and data

processing [13], have brought new possibilities for improving structural and railway monitoring, offering more precise and financially sustainable solutions [9,14,15]. A recent systematic review highlights the growing application of Wireless Sensor Networks (WSNs) in railway infrastructure monitoring, emphasizing their role in real-time data collection and analysis [16]. Wu et. al. (2023) [17] reviews three wheel-rail force measurement Methods - Instrumented Wheelset (IWS), ground-based sensors, and indirect estimation - highlighting IWS as the most mature technique, with future advancements focusing on non-destructive, wireless, and continuous measurement for improved accuracy. Furthermore, the development of AI-based trackside monitoring systems has enabled the unsupervised detection and classification of train wheel defects, contributing to more effective maintenance strategies [2,18,19].

Specifically, a significant amount of recent research focuses on the application of ML-assisted data-driven SHM techniques for railway wheel polygonalization detection [3,20-23]. Unsupervised Machine Learning techniques, particularly Physics-Informed Generative Adversarial Networks (PI-GANs), have shown significant promise in SHM of Vehicle-Bridge Interaction (VBI) systems. For instance, Zhou et al. (2025) [24] proposed a Physics-Constrained Generative Adversarial Network (PC-GAN) for accurately estimating bridge surface roughness from vehicle vibration responses, demonstrating its adequacy under challenging conditions and its potential for rapid bridge pavement inspection. Moreover, Lee et al. (2025) [25] proposed an image-to-image Generative Adversarial Network (GAN) to estimate temporal frequency variations in Vehicle-Bridge Interaction systems, with validation through experiments confirming its effectiveness for bridge condition assessment.

Among those, particularly AutoEncoders (AEs) [13,26] have shown significant promise. AEs are a type of unsupervised ANN that learns efficient data representations by compressing input into a lower-dimensional space and subsequently reconstructing it. By employing AEs as dimensionality reduction techniques and combining them with appropriate metrics, it becomes possible to identify structural behaviors indicative of damage.

According to Spínola Neto et al. (2024) [27], the three most prominent AE approaches in SHM problems are Sparse AutoEncoders (SAEs) [28], characterized by sparsity constraints; Variational AutoEncoders (VAEs) [29], featuring probabilistic representations of the data; and Convolutional AutoEncoders (CAEs) [30], which incorporate convolutional layers. In their work, Spínola Neto et al. (2024) [27] compared the performance of these AEs in SHM problems applied to structures such as frames and bridges, concluding that VAEs exhibited superior performance for the analyzed structural types.

This paper presents a comprehensive comparative study of the main AutoEncoder models for OOR wheel damage identification. Our primary goal is to provide researchers with a clear reference for selecting the most suitable model for this type of application. While numerous studies have explored individual AutoEncoder models, there is a lack of direct comparisons among them, making it challenging for researchers and practitioners to determine the most effective approach. To address this gap, this study evaluates the effectiveness of SAE, VAE, and CAE for detecting stages of polygonization. Furthermore, to enhance the detection capabilities, this paper employs an approach that combines these AE models with Hotelling's T^2 Control Chart to capture and quantify structural variations in railway wheels at different stages of polygonization.

This paper utilizes simulated data generated using computational models implemented in ANSYS® [31], based on the Finite Element Method (FEM). This methodology is commonly adopted in research [3,32,33] due to the significant challenges associated with instrumenting trains with sensors such as accelerometers. As highlighted by Castillo-Mingorance *et al.* (2020) [34], the harsh railway environment, characterized by vibrations, temperature fluctuations, humidity, and dust, poses significant challenges to the deployment and reliable operation of electronic sensors. Moreover, the installation of these sensors often requires access to difficult-to-reach locations and necessitates robust mounting mechanisms to ensure their integrity. Considering this scenario, numerical studies can also provide valuable insights into the robustness of the methodology at different levels of damage or noise in the data.

2. Theoretical Conceptualization

The SHM strategy implemented in this work encompasses two main techniques:

- **Unsupervised Machine Learning using AutoEncoders:** This step involves the application of VAE, SAE, and CAE models to the dynamic data. The models extract features that effectively characterize the monitored vibration signals and serve as a dimensionality reduction tool. It is important to mention that the application of AEs in SHM tasks has been showing promising results, as reported in [27].
- **Hotelling's T^2 Control Chart:** This chart displays the values of the T^2 statistical metric, which is computed from the reduced parameters obtained through the VAE, SAE, and CAE models. The present study uses the T^2 Control Chart to visually and objectively assess the AutoEncoder models' capability to detect structural changes in railway wheels.

AutoEncoders are neural network models designed to learn compact latent representations of input data to reconstruct the original data accurately. Their architecture consists of two main parts: an encoder and a decoder. This design enables AEs to capture the data's intrinsic features and statistical patterns. The encoder maps the input to a lower-dimensional latent space $\mathbf{h} = f(\mathbf{x})$, trying to preserve the most relevant data features. On the other hand, the decoder reconstructs the original data from this latent representation, denoted as $\mathbf{r} = g(\mathbf{h})$. Essentially, it learns to reverse the transformation performed by the encoder [35], as shown in Eq. 1:

$$L(\boldsymbol{x}, g\big(f(\boldsymbol{x})\big) \tag{1}$$

where is L a loss function that penalizes the output $g(f(\mathbf{x}))$, for deviating from the original input $\boldsymbol{x}$.

The following subsections will provide the main principles of the aforementioned techniques. For further information on AutoEncoders, the reference Goodfellow *et. al.* (2016) [35] is advised. Additional details on the Shewhart Control Chart (T^2) can be found in Montgomery (2009) [36].

2.1. Sparse AutoEncoder

Sparse AutoEncoders incorporate a sparsity penalty $\Omega(\mathbf{h})$ into their cost function, as shown in Eq. 2, to create sparser latent representations and increase network robustness by focusing on distinct data features. Goodfellow *et al.* (2016) [35] note that this sparsity enhances anomaly detection. Moreover, Finotti *et al.* (2023) [37] show that SAEs effectively detect behavioral changes in structural systems, simplifying data analysis and improving sensitivity to relevant patterns.

$$L_{SAE} = L(\boldsymbol{x}, g\big(f(\boldsymbol{x})\big) + \Omega(\mathbf{h}) \tag{2}$$

As described in [35], the Mean Squared Error (MSE) in Eq. 3 computes the average of the squared differences between the original values (x_i) and their reconstructions (y_i). Its quadratic penalization of larger errors makes it particularly effective for detecting anomalies, where identifying significant deviations is fundamental.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2 \tag{3}$$

A typical SAE network (Figure 1) features a symmetric arrangement of two encoding and two decoding layers, with a central layer defining the latent representation **h**. Both the first (encoder) and last (decoder) layers contain *M* neurons, matching the length of each sampled signal, while the central layer has *k* neurons ($k \leq M$), thus reducing the dimensionality. The feature vector (**h**) captures essential information about structural behavior and is subsequently used as input to the T^2 Control Chart for structural novelty detection.

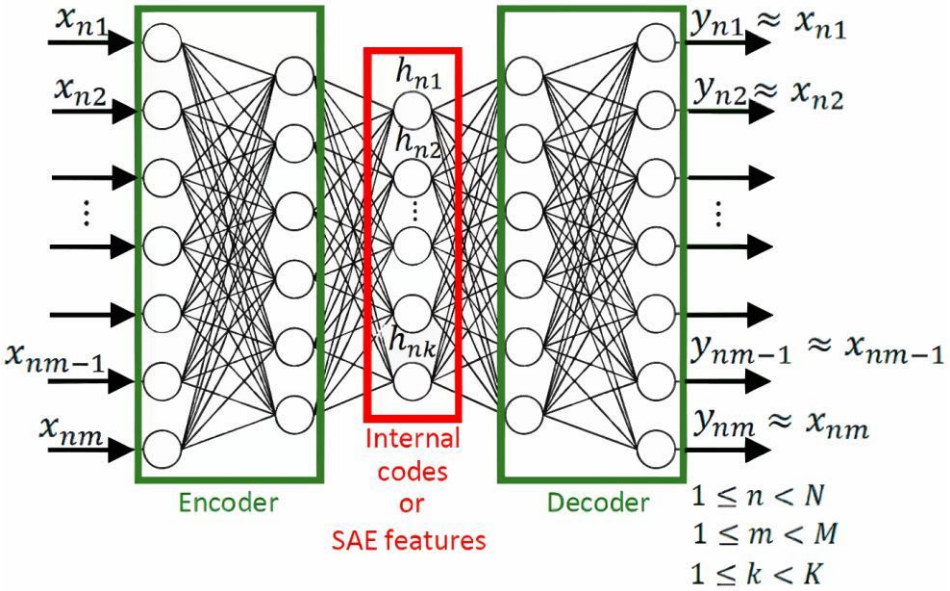

Figure 1: Example of an SAE network structure (Finotti et al. (2021) [13]).

2.2. Variational AutoEncoder

A Variational AutoEncoder extends a standard AutoEncoder by employing a regularized latent space based on principles of probabilistic Machine Learning. Instead of encoding an input $\boldsymbol{x}$ as a single point, the encoder $q_\phi = (f(\boldsymbol{x})|\,\boldsymbol{x})$ receives the input as a distribution characterized by its mean (μ_x) and standard deviation (σ_x), from which a sample is drawn $f(\boldsymbol{x})$ and subsequently decoded. The decoder $p_\phi = (\boldsymbol{x} \mid f(\boldsymbol{x}))$ receives the latent variables $f(\boldsymbol{x})$ and the probabilistic density function provides the matrices containing the mean and covariance of the model $g(f(\boldsymbol{x}))$. The weights ϕ need to be learned by the ANN so that it can configure a loss function and generate both the latent space and the output. As shown in Eq. 4, the loss

function combines a reconstruction term, $L(\boldsymbol{x}, g(f(\boldsymbol{x}))$, and a Kullback-Leibler (KL) divergence term L_{KL}. This regularization component organizes the latent space by aligning it with a standard normal distribution. Similar to the SAE, in this work, the VAE models also incorporate MSE as its loss function for comparative reasons. Figure 2 shows the structure of a general VAE network.

$$L_{VAE} = L(\boldsymbol{x}, g(f(\boldsymbol{x})) + L_{KL} \tag{4}$$

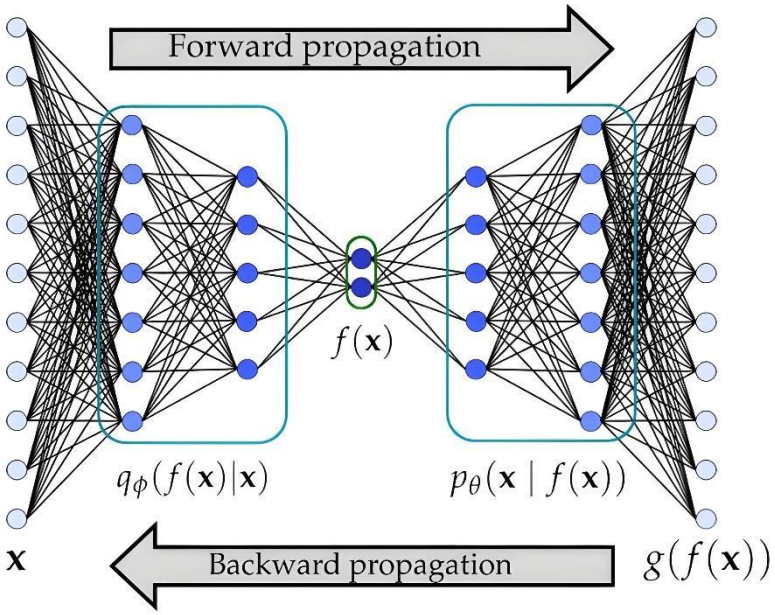

Figure 2: Example of an VAE network structure.

2.3. Convolutional AutoEncoder

Convolutional Neural Networks (CNNs) are specially tailored for grid-structured data, including time series and images, by replacing at least one layer's matrix multiplication with convolution [35,38]. A Convolutional AutoEncoder applies this principle to encode and decode data, achieving success not only in image-based tasks but also in time series analysis [27]. Through 1D convolutional layers and max-pooling (Conv + Max), the encoder extracts and compresses relevant features; flattened and dense layers (Flat + Dense) then prepare the data for reconstruction. The decoder reverses this procedure using up-sampling and convolutions (Up + Conv) to restore the original data. Feature extraction in CAEs is performed through 1D convolutions applied to each input channel, and max-pooling preserves the most informative

responses while reducing dimensionality [27,39]. Figure 3 provides a schematic representation of a CAE network, illustrating its encoding and decoding processes.

Like other AutoEncoders, CAE training iteratively minimizes the discrepancy between the reconstructed and original data. The Mean Absolute Error (MAE) is commonly used as the loss function to quantify this difference. As expressed in Eq. 5, the MAE calculates the mean of the absolute differences across all samples in the dataset. As highlighted by Spínola Neto *et al.* (2024) [27] and Jun Qi *et al.* (2020) [41], the MAE has been widely employed in CAE applications, demonstrating its effectiveness in capturing complex patterns within the data.

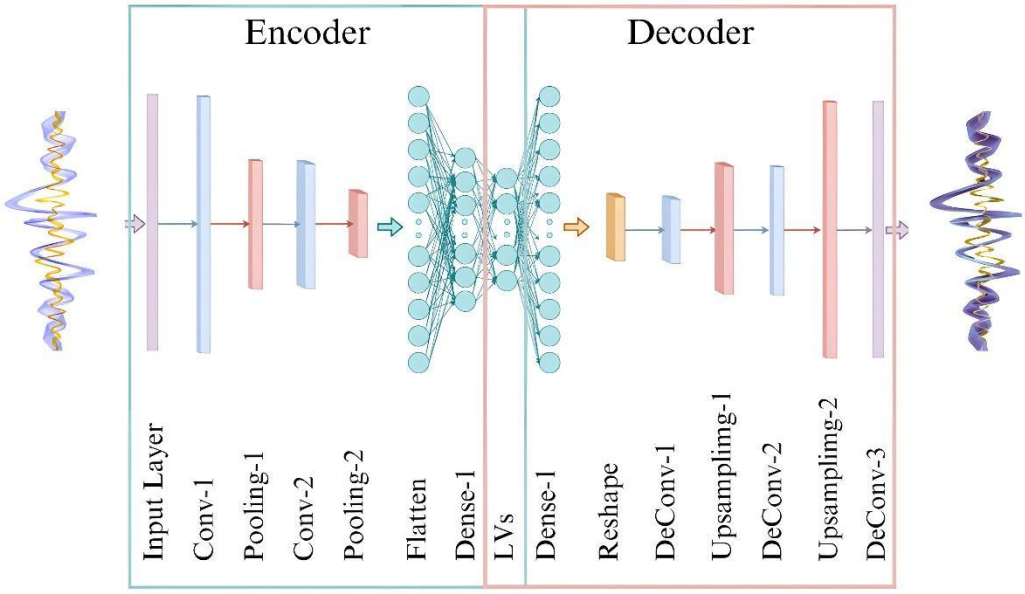

Figure 3: Example of a CAE network structure (Adapted from Cataltas e Tutuncu (2023) [40]).

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - x_i| \tag{5}$$

To conclude the AE background, Table 1 provides an overview of the characteristics that distinguish the three AutoEncoder models, allowing for a direct comparison.

Table 1: Feature analysis across different AutoEncoder implementations evaluated.

Model	**Use in SHM**	**Advantages**	**Disadvantages**	**Key Features of SHM**

SAE	- Feature extraction from vibration, acoustic, or strain signals. - Identifying relevant features indicative of damage through sparse activations.	- Promotes a compact, interpretable feature space. - Reduces overfitting by forcing many hidden units to stay near zero activation. - Relatively simple to implement and train.	- Requires careful tuning of sparsity penalties (too strong can lose key information, too weak can yield irrelevant representations). - Not specifically tailored to spatial/ temporal correlations.	- Particularly good for applications where explainability or interpretability of sensor features is desired. - Works well in combination with other models for dimension reduction.
VAE	- Anomaly detection by learning a probabilistic latent space. - Estimating uncertainty in sensor signals. - Generative modeling of normal structural behavior.	- Provides a probabilistic framework. - Good for assessing the likelihood of damage scenarios. - The latent space can capture underlying structural variability. - Can generate synthetic "normal" data for comparison or data augmentation.	- More complex to train due to KL divergence term & reparameterization trick. - May produce smoother/less sharp reconstruction outputs, especially for highdimensional signals / images.	- Enables thresholding on reconstruction probability (not just reconstruction error). - Helpful for tasks where uncertainty quantification is important in decision-making (e.g., high-stakes infrastructure).
CAE	-Image-based SHM (e.g., thermographic images, structural surface crack detection). - Potentially applied to 2D/3D data	- Learns translationally invariant features, good for imaging tasks. - Handles highdimensional data more efficiently (weight sharing, local receptive fields).	- Less straightforward if data is purely 1D (e.g., single vibration signals), unless converted to 2D representations (spectrograms,	-Particularly wellsuited for visual inspection or 2D representations of structural signals (e.g., spectrogrambasemethods). - Commonly used in damage segmentation or crack detection tasks.

	(e.g., wavefields, damage scans).	- Often provides higher-fidelity reconstructions for visual data.	wavelet tran forms). - Network can become large, requiring more computational resources.	

2.4. Shewhart T² Control Chart

The Shewhart (T^2) Control Chart is a visual statistical tool that monitors the stability of parameters over time. Its graphical representation, consisting of data points and horizontal control limits, allows deviations from the normal condition to be identified. When a point exceeds these limits, it indicates the presence of external factors affecting data variability, signaling an out-of-control condition. In addition to its early detection capability, T^2 is widely recognized in the scientific community for its clear and intuitive interpretation, facilitating decision-making. Its applicability across various fields makes it a valuable tool for identifying structural changes in data, contributing to quality and reliability improvements in SHM systems [36].

In this study, the T^2 statistic is employed to evaluate the integrity of railway wheels based on features extracted from AE models. This metric quantifies the distance of a new observation from the historical data mean, allowing the detection of anomalies that may indicate damage. The magnitude of T^2 is directly related to data variability and correlation. By comparing the calculated value to a statistically established Upper Control Limit (UCL), it is possible to determine whether the structure operates within its normal conditions [36,37].

Although other metrics exist, such as the MSE and the Original to Reconstructed Signal Ratio (ORSR), T^2 proved to be more effective in identifying anomalies in the analyzed data. Eq. 6 presents the mathematical formulation of T^2:

$$T^2 = R\left(\overline{\boldsymbol{h}} - \overline{\overline{\boldsymbol{h}}}\right)^T \boldsymbol{S}^{-1}\left(\overline{\boldsymbol{h}} - \overline{\overline{\boldsymbol{h}}}\right), \quad (6)$$

where $\overline{\boldsymbol{h}}$ represents the sample mean vector, while $\overline{\overline{\boldsymbol{h}}}$ and $\boldsymbol{S}$ are the reference mean vector and the reference covariance matrix, respectively. The UCL was defined as the value that exceeds

95 % of the T^2 values of the training data, establishing a threshold for anomaly detection [27,37].

Thus, T^2 values above the UCL indicate significant changes in the data, suggesting the presence of structural damage or other irregularities. This approach allows for continuous monitoring of the structure's condition and the implementation of corrective actions before catastrophic failures occur. It is important to emphasize that, by definition, 5% of the T^2 values from the training data will exceed the UCL, which is not a cause for concern if they remain close to the threshold [27,37].

3. Methodology

This work proposes an unsupervised learning approach based on three different AE models to investigate their ability to detect and isolate simulated vibration signals from different stages of railway wheel polygonization. These models are trained with a subset of the data to learn the most informative latent representation for each class. When confronted with new data, the models reconstruct a given signal and compare it with its original input. The magnitude of the difference between these two representations, known as reconstruction error, serves as a metric to quantify the dissimilarity between the analyzed signal and each reference class, e.g., ``undamaged wheel", ``damaged wheel level 1", etc.

As with any unsupervised AutoEncoder model-based approach for SHM, it is essential to first establish a baseline behavior - whether normal or not - by assessing the structure's initial health condition. Once this baseline is defined, the monitoring procedure in our approach operates in an unsupervised manner. As new data are acquired, the Shewhart T^2 statistic is evaluated. If its value exceeds a predefined threshold, the new data is considered significantly different from the baseline, indicating a potential change in the structure's behavior. To evaluate the effectiveness of the proposed methodology, the data are divided into three sets: training, validation, and monitoring. Specifically, 70% of the data representing the undamaged structural state are used for training, enabling the AE models to learn to accurately reconstruct the vibration signals associated with this condition. The remaining 30% of the undamaged state data form the validation set, which is used to fine-tune the model's hyperparameters and assess its generalization capability. Finally, the model performance is evaluated using a monitoring set composed of data from two damaged structural states. The T^2 statistic is employed as a metric to quantify the distance between the latent representation new set of signals and the distribution learned during training. The model is expected to yield significantly higher T^2 values for the

monitoring data, indicating the presence of damage in the structure and demonstrating the AEs capability to distinguish among different structural states.

For all cases analyzed in this study, the simulated dynamic signals are directly used in the time domain as input for the AE models due to its simplicity and effectiveness. Preliminary tests showed that reconstruction errors in the time domain were similar to those in the frequency domain, indicating that both approaches perform equally well. Consequently, using time domain data was preferred, as it eliminates additional transformations, easing the data preprocessing without compromising signal compression and reconstruction accuracy. Besides, the time-domain vibration signals are standardized using the *z*-score normalization to ensure comparability between different samples and facilitate the mapping performed by AEs. The *z*-score transforms the original data into a common scale with zero mean and unit standard deviation, and it is applied before AEs process data. Eq. 7 mathematically describes this transformation, where $\mathbf{z}$ represents the standardized vector, $\boldsymbol{x}$ the original vector, μ the mean of the data, and σ the standard deviation.

$$\mathbf{z} = \frac{\boldsymbol{x} - \mu}{\sigma} \tag{7}$$

The hyperparameters for the AutoEncoders - such as the learning rate, the number of layers, and the latent representation dimensions - were optimized using Optuna, a Python library that applies Bayesian optimization [42]. Optuna explores the hyperparameter space through probabilistic methods, guiding its search based on previous results to identify the best configuration for minimizing the loss function. Optimizing hyperparameters in AE models is particularly challenging due to architectural complexity and the non-linear interaction among parameters. The Bayesian approach that Optuna uses differs from traditional methods, such as grid search and random search, once it constructs iterative probabilistic models that guide the search toward more promising regions of the hyperparameter space. This capability is enhanced by the Tree-structured Parzen Estimator (TPE) algorithm, which employs non-parametric distributions to estimate the probability of superior performance across different configurations [43]. Minimizing the difference in T^2 values between the training and validation sets ensured that the model could effectively discriminate between different structural states, promoting adequate generalization. Optuna performs 50 iterations to minimize the reconstruction error.

Figure 4 provides an overview of the proposed methodology. This graphical representation facilitates the understanding of the process of analyzing and quantifying structural modifications in railway wheels.

All the simulations performed in this study were conducted in a local environment set up in Jupyter Notebook, implemented in version 1.93 of Visual Studio Code [44]. The code used was developed in Python, version 3.10.11, ensuring compatibility with the necessary libraries and tools for performing the analyses. The machine used to run the simulations is equipped with an AMD Ryzen 5 5500 processor, 6 cores, and 12 threads, operating at 3.6 GHz (with a boost of up to 4.2 GHz), a 19 MB cache, and compatibility with the AM4 socket. The installed RAM consisted of a 16 GB DDR4 Mancer Dantalion Z module, with a frequency of 3200 MHz and C19 latency. In addition, the simulations were accelerated by a Mancer Radeon RX 5500 XT Streaky graphics card, equipped with 8 GB of GDDR6-128-bit memory.

The choice of a local infrastructure, as opposed to cloud-based platforms, was motivated by several reasons. Firstly, cloud environments often require specific subscriptions for access to more powerful Graphical Processing Units (GPUs) and Central Processing Units (CPUs), which can result in significant recurring costs for long-term projects. Additionally, such platforms may experience disconnections after prolonged continuous use, compromising the execution of lengthy simulations and requiring manual restarts. With the local setup, it was possible to ensure greater control over the execution environment, continuity in data processing, and resource optimization while maintaining the reproducibility and computational efficiency required for this study.

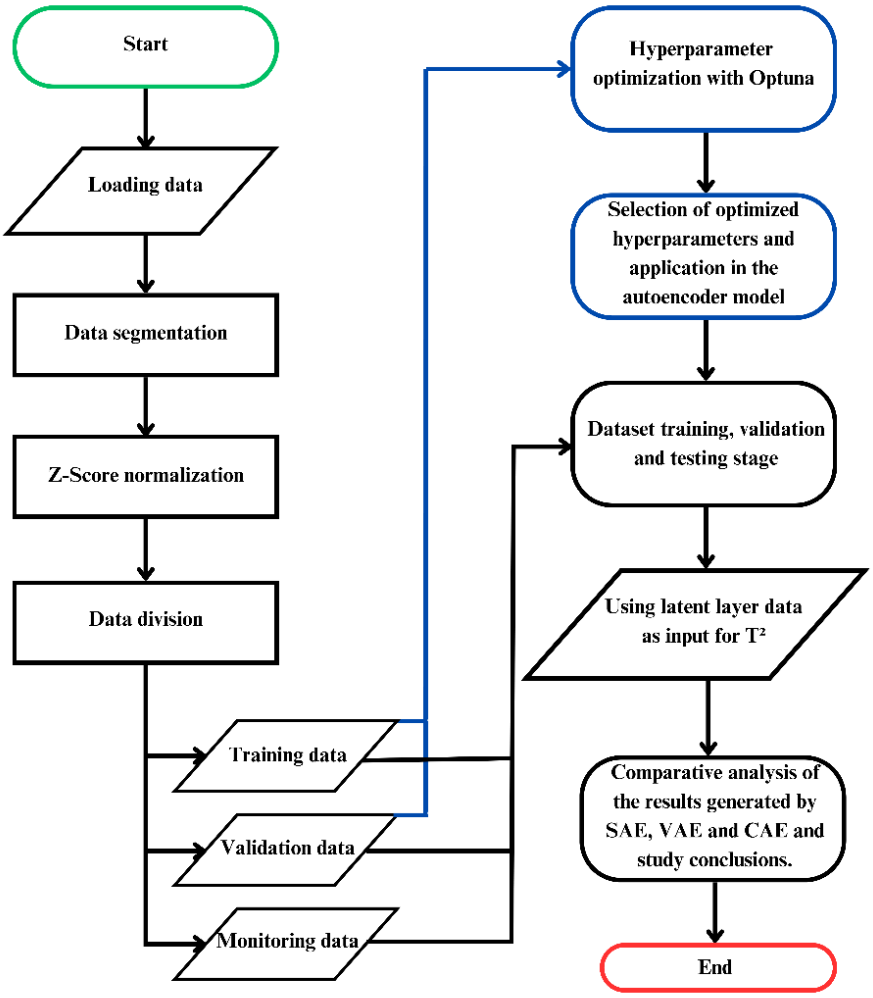

Figure 4: Flowchart of the proposed methodology.

4. Characterization of the Simulated Data

This study uses numerical simulations to investigate the identification of polygonal wear in railway wheels. A detailed model of a Laagrss freight train interacts with a three-dimensionally modeled track, considering different levels of wheel [45]. The dynamic responses of the system are obtained and analyzed to extract features that enable the identification of polygonal defects.

Section 4.1 describes the vehicle model, while Section 4.2 details the track modeling. The interaction between the two is presented in Section 4.3, with emphasis on the wheel-rail contact model. In Section 4.4, the simulation scenarios and the virtual instrumentation used for data collection are described. Finally, Section 4.5 presents the polygonal wear profiles considered in the studies, based on real wheel data.

4.1.Railway vehicle

The present study employed a three-dimensional numerical model of the Laagrss freight train, consisting of five wagons and capable of reaching speeds of up to 120 km/h [2,45]. The ANSYS® platform (2018) [31] was used to develop a detailed 3D multi-body model, simulating the vehicle dynamics through spring-damper elements and lumped masses. The geometry and material properties of the model, including a tare weight of 27 t per wagon and a load capacity of 52 t, (detailed in Table 2 and Figure 5). The vehicle was calibrated based on real data and modal parameters outlined in the work of Ribeiro *et al.* (2013) [46].

The interaction between wagon components was represented using rigid beam elements, enabling an accurate analysis of the vehicle's dynamic response. The modeling methodology adopted, as detailed in Bragança *et al.* (2022) [47], allows for a precise evaluation of the interaction between vehicle components and its influence on the system's dynamic behavior.

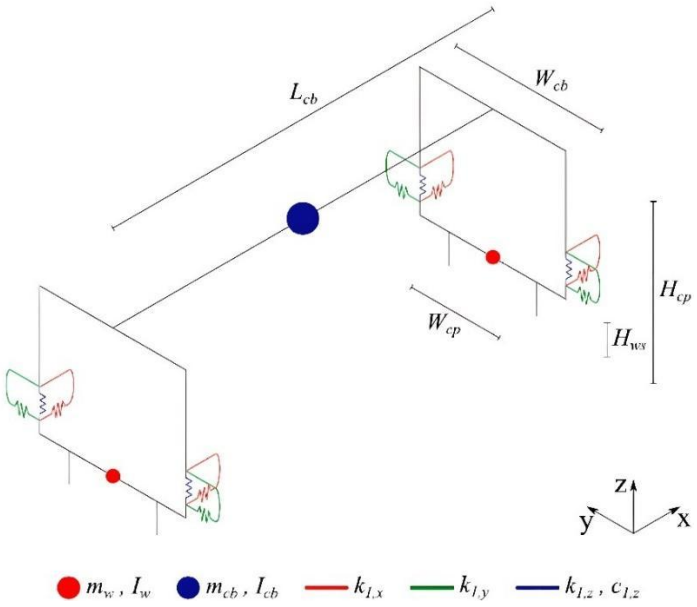

Figure 5: Graphical representation of the numerical model of the railway vehicle considered (Guedes *et al.* (2023) [2]).

Table 2: Parameters and adopted values for the car-body, wheel-set, and suspensions.

Parameter	Symbol (Unit)	Adopted Value
Car-body		
Mass	m_{cb} (t)	41,1

Roll moment of inertia	$I_{cb,x}$ (t . m²)	49
Pitch moment of inertia	$I_{cb,y}$ (t . m²)	673
Yaw moment of inertia	$I_{cb,z}$ (t . m²)	665
Length	L_{cb} (m)	10
Wheel-set		
Mass	m_w (kg)	1247
Roll moment of inertia	$I_{w,x}$ (kg . m²)	312
Yaw moment of inertia	$I_{w,z}$ (kg . m²)	312
Suspensions		
Longitudinal stiffness	$k_{l,x}$ (kN / m)	44,981
Lateral stiffness	$k_{l,y}$ (kN / m)	30,984
Vertical stiffness	$k_{l,z}$ (kN / m)	1860
Vertical damping	$c_{l,z}$ (kN . s /m)	16,7

4.2. Railroad

The numerical modeling of the track employed in this study is based on the research conducted by Montenegro *et al.* (2020) [48] and is also implemented in the ANSYS® software [31]. The model adopts a multilayer approach to represent the interaction between the track components, including ballast, sleepers, rails, and foundation [2,45].

Figure 6 illustrates the discretization of the model, where the rails and sleepers are modeled using beam elements with appropriate properties. The interfaces, such as ballast and pads/fasteners below and above the sleepers, are considered using spring and pressure point elements in all three directions. Finally, the foundation properties were considered by a simplified approach based on a Winkler model (presented in detail in Mosleh *et al.* (2020) [49]). The 3D ballast track numerical model is validated with modal parameters as presented in Ribeiro *et al.* (2021) [50]. The material and geometric parameters considered in the simulation are detailed in Table 3.

Table 3: Parameters for the numerical model of the railway track structure.

Parameter	Symbol (Unit)	Value
Rail Area	A_r (m²)	$7{,}67 \times 10^{-4}$
Rail Density	ρ_r (kg / m^3)	7850

Rail Second Moment of Area	I_r (m^4)	30,38 × 10^{-6}
Rail Young's Modulus	E_r (N / m²)	210 × 10^9
Rail Pad, Longitudinal Stiffness	$k_{p,x}$ (N / m)	20 × 10^6
Rail Pad, Longitudinal Damping	$C_{p,x}$ (N . s / m)	50 × 10^3
Rail Pad, Lateral Stiffness	$k_{p,y}$ (N / m)	20 × 10^6
Rail Pad, Lateral Damping	$C_{p,y}$ (N . s / m)	50 × 10^3
Rail Pad, Vertical Stiffness	$k_{p,z}$ (N / m)	500 × 10^6
Rail pad, Vertical Damping	$C_{p,z}$ (N . s / m)	200 × 10^3
Sleeper Density	ρ_s (kg / m^3)	2590
Ballast, Longitudinal Stiffness	$k_{b,x}$ (N / m)	900 × 10^3
Ballast, Longitudinal Damping	$C_{b,x}$ (N . s / m)	15 × 10^3
Ballast, Lateral Stiffness	$k_{b,y}$ (N / m)	2250 × 10^3
Ballast, Lateral Damping	$C_{b,y}$ (N . s / m)	15 × 10^3
Ballast, Vertical Stiffness	$k_{b,z}$ (N / m)	30 × 10^6
Ballast, Vertical Damping	$C_{b,z}$ (N . s / m)	15 × 10^3
Foundation, Longitudinal Stiffness	$k_{f,x}$ (N / m)	20 × 10^6
Foundation, Lateral Stiffness	$k_{f,y}$ (N / m)	20 × 10^6
Foundation, Vertical Stiffness	$k_{f,z}$ (N / m)	20 × 10^6

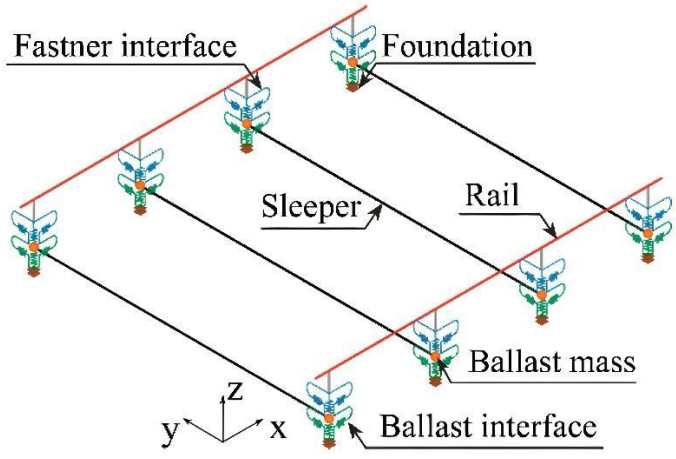

Figure 6: 3D graphical representation of the numerical model of the track adopted in the study (Guedes *et al.* (2023) [2]).

The geometric specifications and simulation parameters used for the tracks and trains analyzed in this study are also presented in detail by Mosleh *et al.* (2020, 2021) [49,51]. It is recommended to consult these works for additional information on the models adopted.

4.3. Dynamic Train-Rail Interaction

The methodology used for vehicle-structure interaction analysis, detailed and validated in Montenegro *et al.* (2015) [52] to take into consideration the lateral dynamics alongside vertical interaction, has been widely applied in various contexts [49,51,53-55].

The coupling is established through a validated wheel-rail contact model [52], which is formulated using a specially designed contact finite element. This element employs a structured approach consisting of three main stages:

- **Geometric Contact Determination:** The first stage involves identifying the precise contact point between the wheel and rail at each computational step. This is achieved using an online method [56], where a set of nonlinear equations is solved to ensure compatibility between the contact surfaces. Further details on the parameterization of these surfaces and the associated geometric equations can be found in Montenegro *et al.* (2015) [52].
- **Normal Contact Force Calculation:** Once the contact point is determined, the second stage focuses on computing the normal forces exerted at the interface. These forces are obtained using Hertzian contact theory [57], which provides an analytical approach to estimating pressure distribution and deformation within the contact patch.
- **Tangential Contact and Creep Force Computation:** The final stage deals with the forces generated due to rolling friction between the wheel and rail. The creep forces, which arise from relative motion at the contact interface, are evaluated using the USETAB routine [58]. In this method, precomputed values for longitudinal and lateral tangential forces are stored in a lookup table, allowing efficient interpolation during dynamic simulations based on creepage values and the aspect ratio of the Hertzian contact ellipse.

For the dynamic analysis solver, the system's governing equilibrium equations are supplemented by constraint equations that link the displacements of the vehicle's contact nodes to the corresponding nodal displacements of the track structure. These equations form a unified system, with both displacements and contact forces as unknowns, which is then solved using an optimized block factorization algorithm (further details available on Montenegro *et al.*

(2015) [52]). Since the numerical framework is built on the finite element method, it allows for the modelling of both structures and vehicles of varying complexity. The current formulation has been implemented in MATLAB® [59], whereas the vehicle and track structures are modelled in ANSYS® [31], as illustrated in Figure 7. A more in-depth discussion of the train-track interaction tool and the wheel-rail contact model, along with their validation, is provided by Montenegro *et al.* (2015) [52].

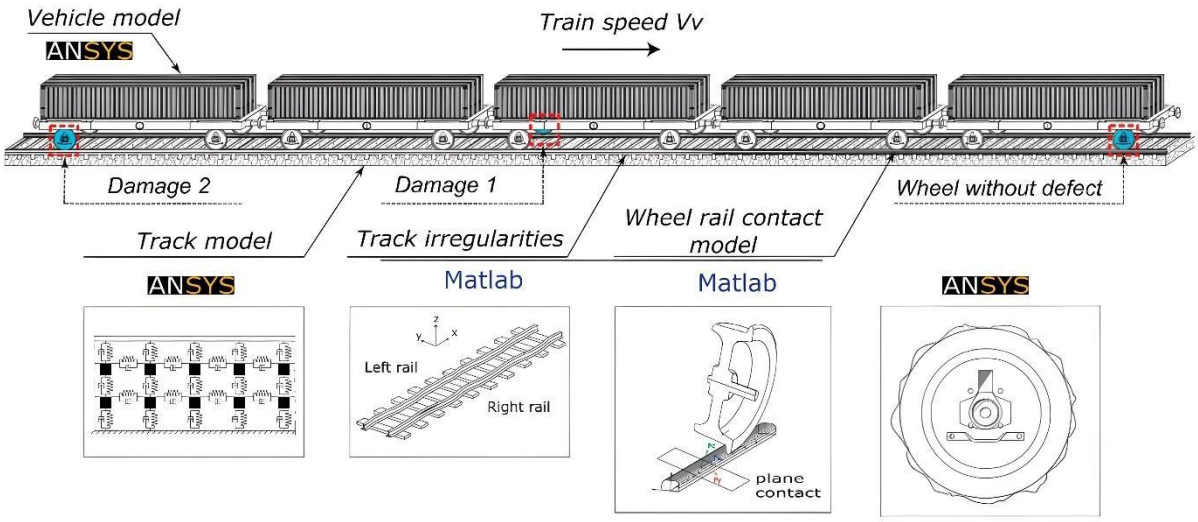

Figure 7: Representation of the numerical modeling steps that form the study model, and the wheels considered in the analysis highlighted. (Adapted from Mohammadi *et al.* (2023) [60]).

4.4.Simulated Scenarios

For the early detection of wheel polygonization stages, a virtual monitoring system was implemented to acquire the dynamic data evaluated in this study. Following the recommendations of Mosleh *et al.* (2022) [61,62], the instrumentation focused on the track, specifically the rail, as sensors placed near the rails are more promising in achieving satisfactory results. Six accelerometers were strategically positioned at the midpoint between sleepers, as shown in Figure 8, to capture vibrations induced by train passages. This configuration enables the precise identification of vibrational signatures characteristic of defective wheels, allowing intervention before more severe damage occurs.

To evaluate the effectiveness of the proposed method for detecting polygonalized wheels, numerical simulations were conducted under a variety of operating conditions. The simulation dataset included both reference scenarios, characterized by the absence of defects, and scenarios with two distinct stages of wheel wear. This approach enabled the validation of

the method's ability to accurately reproduce experimental data obtained under real-world conditions, including cases involving multiple stages of polygonalization. The simulations were performed considering a nominal speed of 80 km/h.

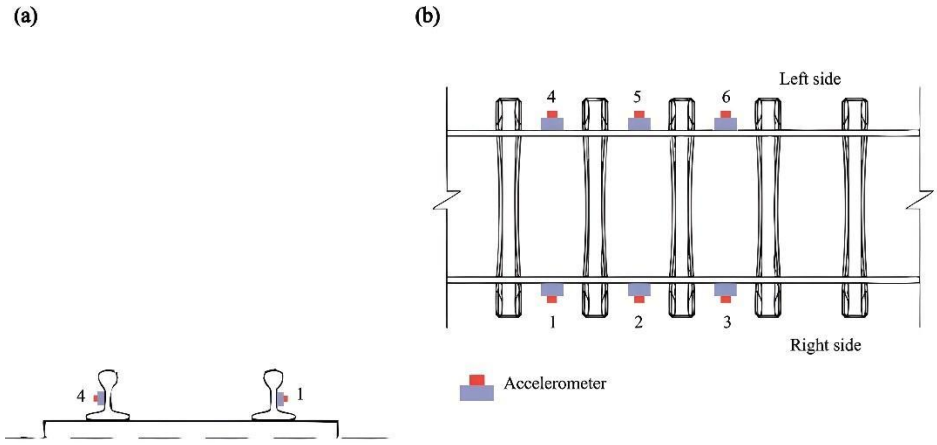

Figure 8: Virtual wayside monitoring system with accelerometers positioned mid-span between sleepers: (a) Back view; (b) Top view. (Adapted from Magalhães *et al.* (2024) [3]).

The analysis of the simulation data focused on three distinct cases: (i) a reference scenario, characterized by the absence of defects in the front right wheel of the first wagon; (ii) a polygonalization scenario in the rear left wheel of the third wagon (Damage 1); and (iii) a polygonalization scenario in the rear right wheel of the fifth wagon (Damage 2). For each case, 10 numerical simulations were performed. The selection of these specific cases was driven by the need to assess the method's ability to detect damage in different positions along the train. Figure 7 provides a graphical visualization of the various wheel positions considered in the simulated scenarios.

To prepare the simulation data for training the AE model, a preprocessing step was carried out. Each complete simulation, consisting of 39.169 data points, represented the full dynamic signal of the train in a given scenario. However, the high data volume and the limited number of samples per scenario (only 10 simulations) could compromise training effectiveness and increase computational costs. Given this context, each simulation of each of the scenarios was segmented into 10 smaller parts, each containing 3.916 points. This strategy significantly expanded the training dataset from 30 simulations to 300 simulations without compromising the representativeness of the original data. Data segmentation is a common practice in Machine Learning, especially when dealing with time series, as it aims to optimize the training process and enhance the model's generalization capability. Table 4 presents a comparison of the number

of signals per scenario before and after segmentation, highlighting the substantial increase in the number of samples available for training, validation, and monitoring.

Table 4: Comparison of data before and after segmentation for each scenario.

Scenario	Number of Samples (Before)	Number of Samples (After)	Points per Sample (Before)	Points per Sample (After)
No Damage (Reference)	10	100	39.169	3.916
Damage 1 (Left Rear Wheel, 3rd Wagon)	10	100	39.169	3.916
Damage 2 (Right Rear Wheel, 5th Wagon)	10	100	39.169	3.916

The acceleration data, collected in baseline and damaged scenarios, were prepared for numerical simulation through the application of a second-order Chebyshev type II low-pass filter. With a cutoff frequency of 500 Hz, this filter attenuates high-frequency components, isolating the dynamic characteristics of interest. To simulate real operating conditions, random noise with a normal distribution was added to the filtered signals, representing the inherent uncertainties and disturbances in the system. The 10 kHz sampling rate ensures adequate temporal resolution for capturing the system's fast dynamics.

4.5. Wheel polygonization profiles

The wheels of railway vehicles, although designed to be circular, often present irregularities that significantly affect the dynamic performance of the system. One such imperfection, known as polygonization, consists of radial undulations along the tread. The severity of these undulations, when greater than 0.2 mm, can induce considerable variations in the wheel-rail contact

forces, generating vibrations and noise. The geometric characterization of polygonization involves the calculation of the wavelength (λ) of the irregularities, which is a function of the wheel radius (R_w) and the harmonic order (θ) of the undulation, as expressed by Eq. 8, where θ takes positive integer values.

$$\lambda = \frac{2\pi R_w}{\theta}, \qquad \theta = 1, 2, 3, \dots, n \tag{8}$$

To cover a wide range of operating conditions, two sets of wheel profiles were generated for the numerical simulations. These profiles were constructed from real wheel data, both under initial polygonization damage conditions (Damage 1) and under more severe polygonization damage conditions (Damage 2), with their respective irregularity spectra presented in Figure 9.

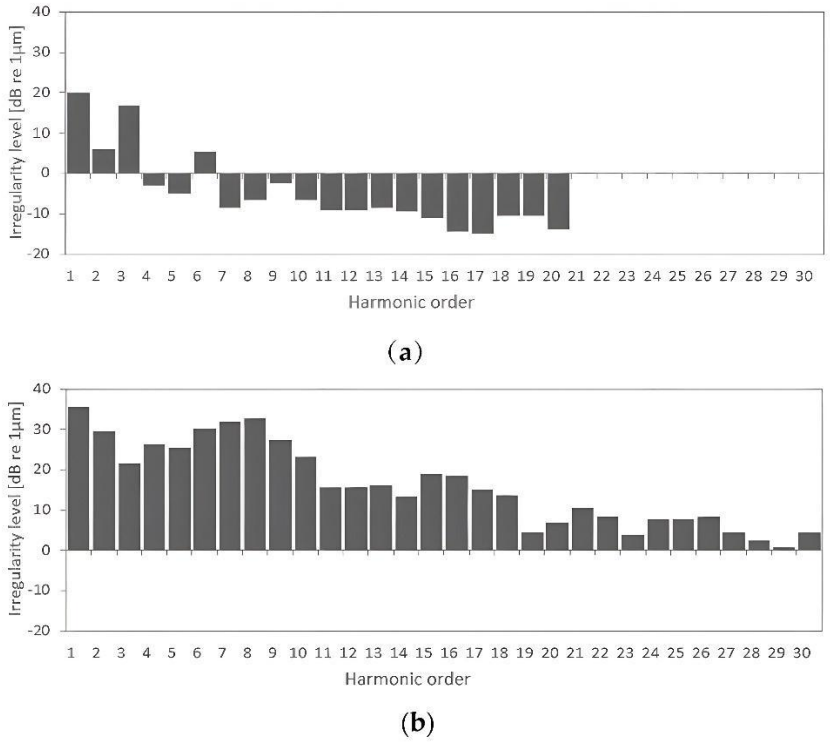

Figure 9: Amplitude spectra of wheel irregularity (Lw) and harmonic order (θ) in: (a) undamaged polygonal wheel (Johansson *et al.* (2005) [63]), (b) damaged polygonal wheel (Cai *et al.* (2019) [64]).

The measurement data used to characterize the wheel profiles were obtained from two main sources: {i} Johansson *et al.* (2005) [63] provided data for intact polygonal wheels (damage 1), with a wavelength spectrum extending across 20 harmonics, ranging from 0.135 to 2.7

meters (Figure 9a); {ii} Cai et al. (2019) [64] provided data for polygonal wheels with damage (damage 2), where the first 30 harmonics were considered. For these latter data, the 6th to 8th harmonic orders were predominant, indicating a specific wear pattern.

A variety of wheel profiles with different levels of irregularities were numerically synthesized. The methodology employed involved the superposition of multiple sinusoidal functions (H=30), as described in Eq. 9.

$$\omega(x_\omega) = \sum_{\theta=1}^{H} A_\theta \sin(\frac{2\pi}{\lambda}\, x_\omega + \psi_\theta) \tag{9}$$

The amplitude of each sinusoidal component, A_θ, was determined using Eq. 10, where ω_{ref}, represents a reference wavelength.

$$A_\theta = \sqrt{2}\,.\,10^{\frac{L_\omega}{20}}\,.\,\omega_{ref} \tag{10}$$

Figure 9 illustrates the irregularity spectra used to calibrate the numerical models, both for the initial polygonization condition (Figure 9a) and for the severe polygonization condition (Figure 9b). For each spectrum, a set of profiles was generated by randomly varying the initial phases ψ_θ of the sinusoidal functions within the range of 0 to 2π. This approach enabled the simulation of the inherent randomness associated with wheel manufacturing and wear processes. The high similarity between the acceleration signals obtained for healthy wheels and those with different levels of polygonization can be verified in Figure 10. This figure illustrates the complexity of the task of classifying defects in railway wheels using Machine Learning techniques. The small variation in signal amplitude and the similarity between them make the distinction between the classes a complex task, requiring the extraction of high-order features for accurate classification.

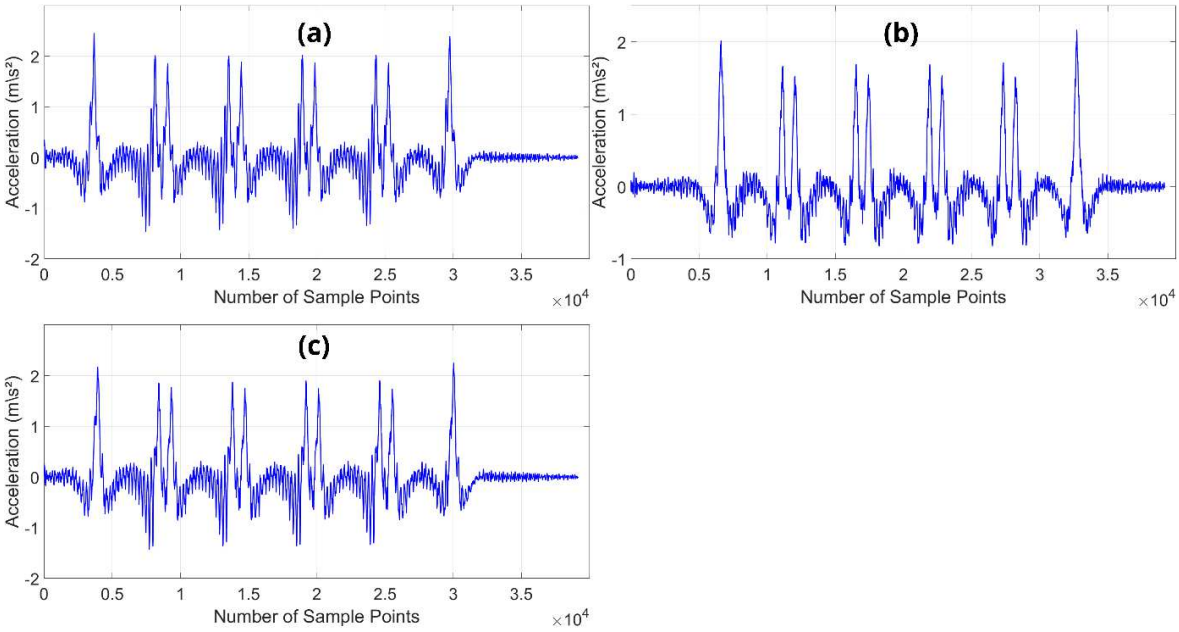

Figure 10: Original signals from accelerometer 1 for the conditions: (a) No damage; (b) Damage 1; (c) Damage 2.

5. Application of the Proposed SHM Methodology to Numerical Data

5.1. Analysis of Reconstruction Results

The original vibrational response and its respective reconstruction made by VAE, SAE, and CAE models for the first accelerometer of each simulated scenario are presented in Figures Figure *11*, Figure *12*, and Figure *13*, respectively. The results of the other accelerometers were similar to those presented and thus, were omitted. The optimal hyperparameters for the AutoEncoders, determined using Optuna, are listed in Table 5. They are applied in both the training and validation of the models.

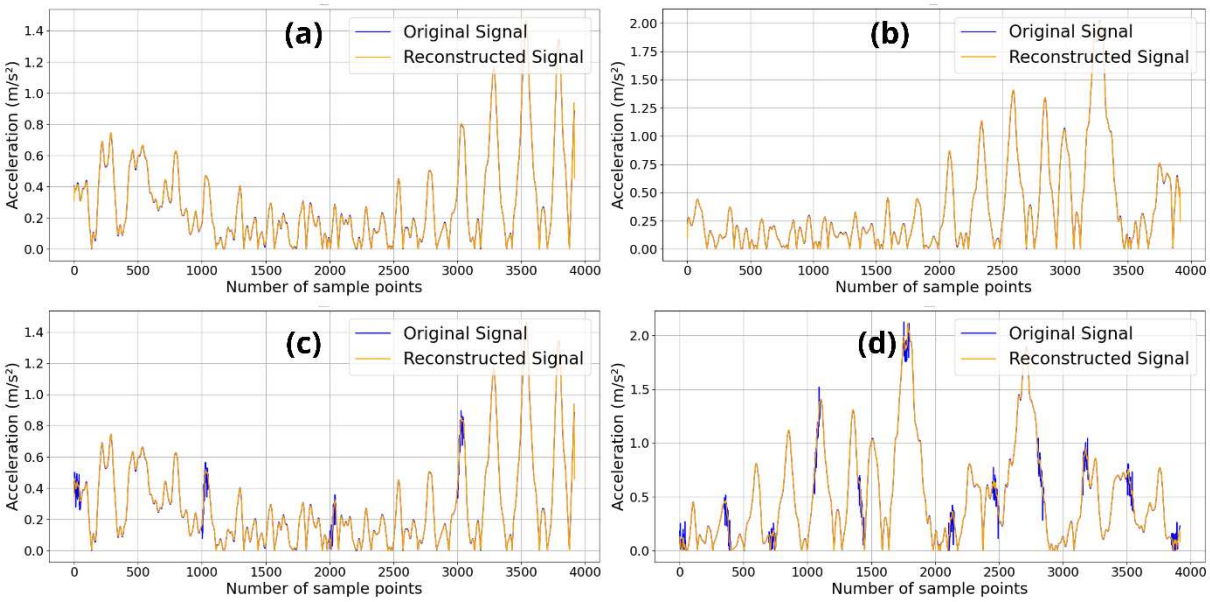

Figure 11: Original and reconstructed signals by the CAE model for accelerometer 1: (a) Training of damage class 0; (b) Validation of damage class 0; (c) Testing of damage class 1; (d) Testing of damage class 2.

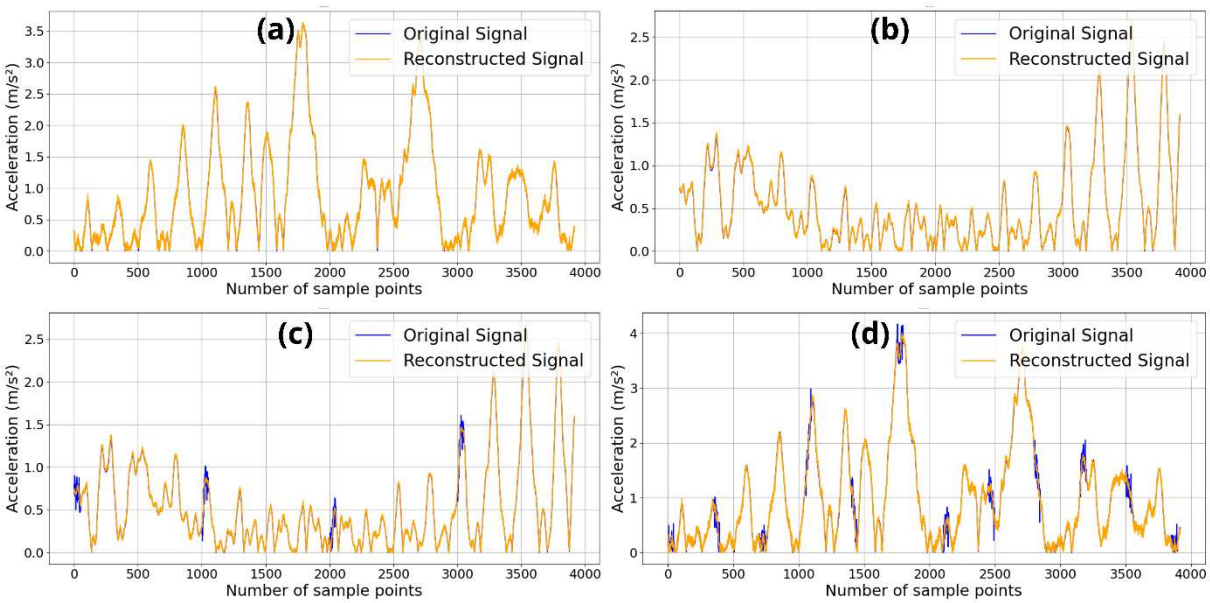

Figure 12: Original and reconstructed signals by the SAE model for accelerometer 1: (a) Training of damage class 0; (b) Validation of damage class 0; (c) Testing of damage class 1; (d) Testing of damage class 2.

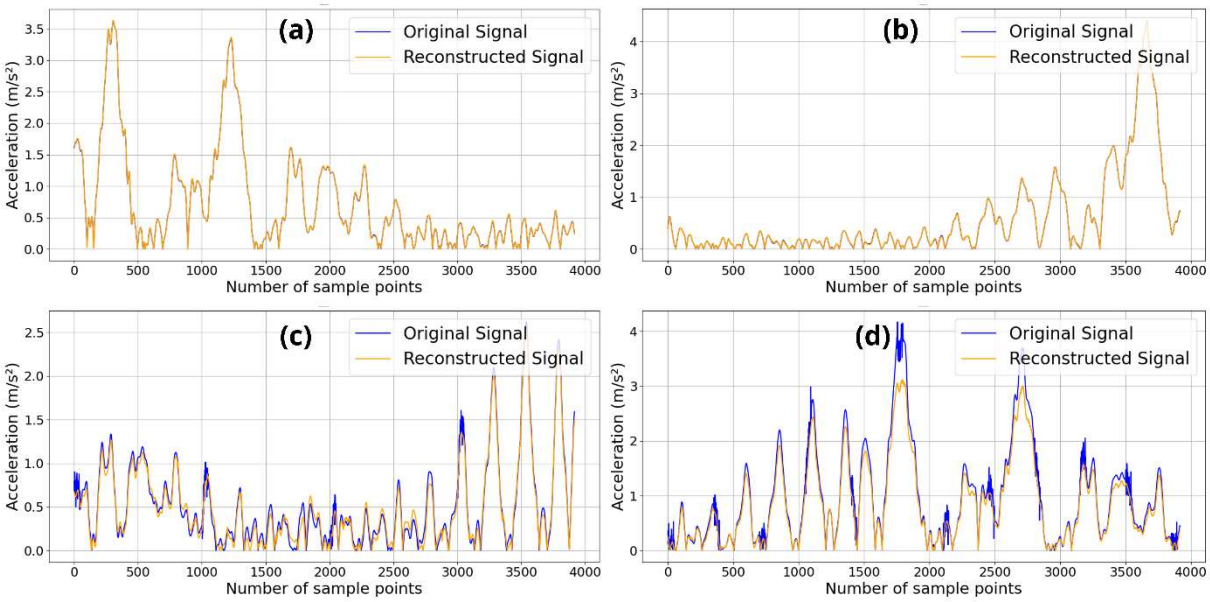

Figure 13: Original and reconstructed signals by the VAE model for accelerometer 1: (a) Training of damage class 0; (b) Validation of damage class 0; (c) Testing of damage class 1; (d) Testing of damage class 2.

Except for the VAE, which exhibited difficulties in fully reconstructing all signal classes, both SAE and CAE demonstrated a good ability to represent the original characteristics of the data. A comparative analysis of the latent representations reveals a clear gradient in reconstruction similarity relative to the original signal, with classes closer to Class 0 exhibiting higher fidelity. This trend suggests that both the SAE and CAE can capture the structural nuances of the data, enabling the accurate identification of anomalies and changes in behavioral patterns.

Table 5: Optimized hyperparameters used in autoencoder models.

Hyperparameter	CAE	SAE	VAE
learning_rate	0.000688171	0.000690141	0.00000104471
epochs	283	306	140
batch_size	5	32	12
original_dim	3916	3916	3916
intermediate_dim	-	-	2170
latent_dim	157	345	214
optimizer	adam	adam	adam
lambda_sparse	-	0.00003972	-

Visualization of the results obtained using Shewhart's T^2 Control Chart (Figure 14 to Figure 16) confirms the effectiveness of the autoencoders in detecting anomalies. Analysis of the data from accelerometers 1 to 4 reveals a clear separation among classes, indicating that the models can discriminate between different operating conditions. This separability is particularly evident for the SAE and CAE models, corroborating the results of the qualitative analysis of the reconstructions.

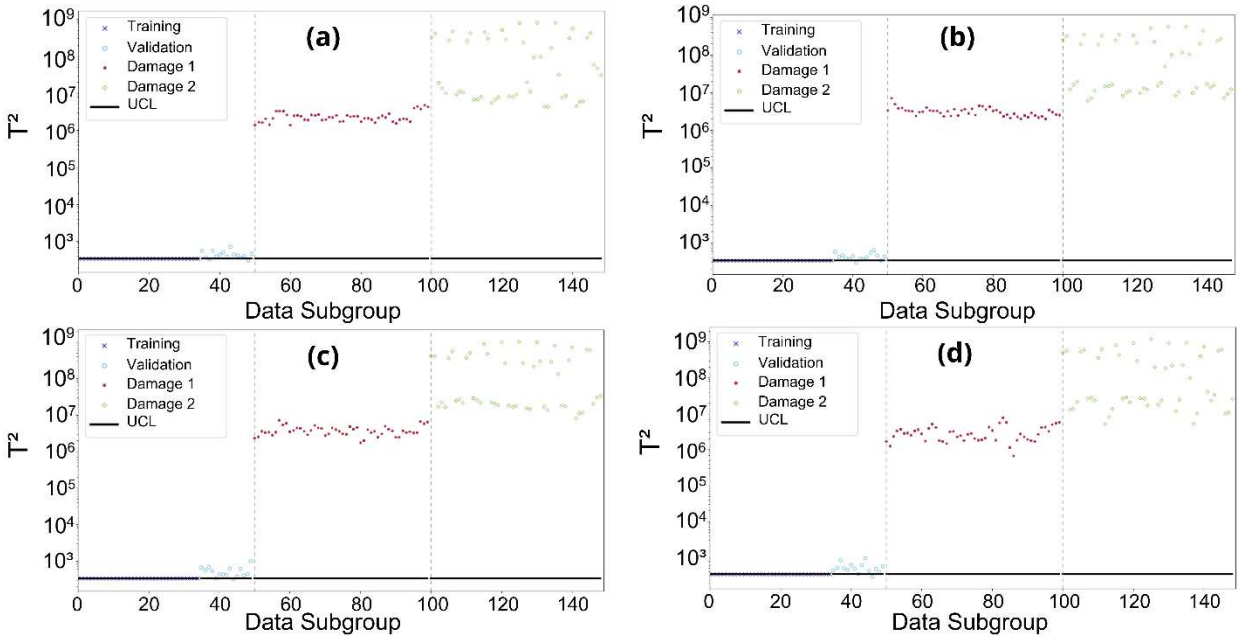

Figure 14: Shewhart T^2 Control Charts for the analysis of the CAE model. (a) Accelerometer 1; (b) Accelerometer 2; (c) Accelerometer 3; (d) Accelerometer 4.

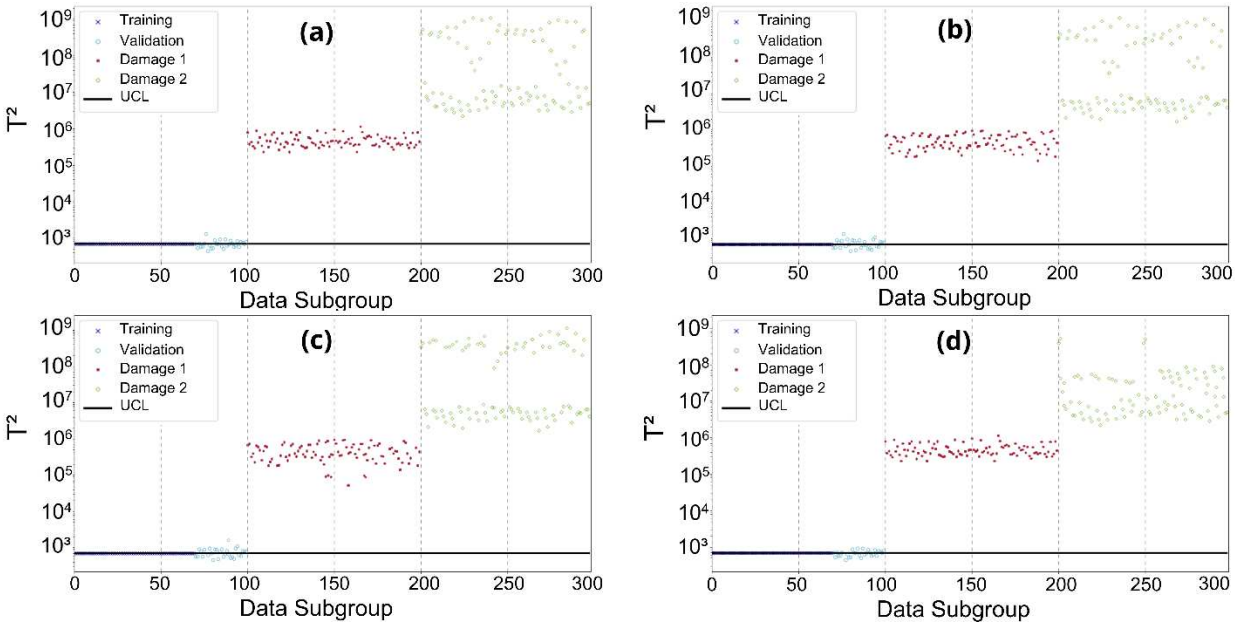

Figure 15: Shewhart T² Control Charts for the analysis of the SAE model. (a) Accelerometer 1; (b) Accelerometer 2; (c) Accelerometer 3; (d) Accelerometer 4.

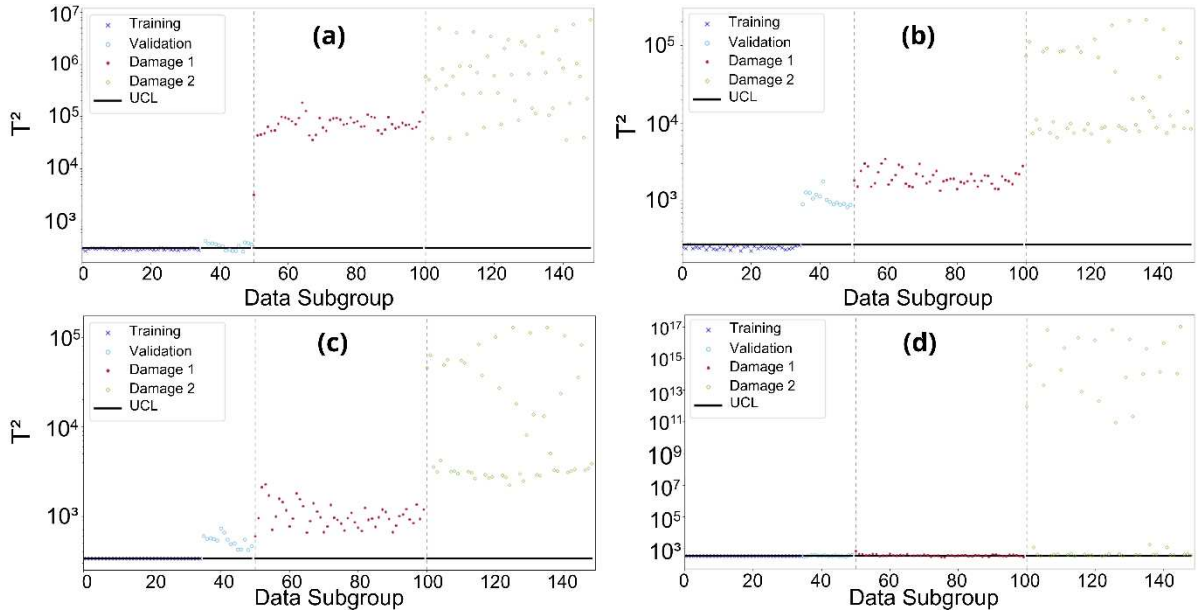

Figure 16: Shewhart T² Control Charts for the analysis of the VAE model. (a) Accelerometer 1; (b) Accelerometer 2; (c) Accelerometer 3; (d) Accelerometer 4.

The analysis of Figure 14 and Figure 15 reveals a good agreement between the T² results obtained by the CAE and SAE models in both training and validation classes. This similarity indicates a generalization capability of the models, i.e., their ability to accurately classify data not used during training. The VAE model, as illustrated in Figure 16 and corroborated by Figure 19, exhibits mixed results. In some cases, a less evident separation between validation and monitoring data is observed, as indicated in Figure 16b and Figure 16c. In other cases, it is

possible to identify an overlap between damage classes 1 and 2, as evidenced in Figure 16a and Figure 19a. Furthermore, the model suggests a convergence of damage classes 1 and 2 with the training and validation data, as demonstrated in Figure 16d and Figure 19d. These mixed behaviors and inconsistencies highlight the difficulty of the VAE in capturing the intrinsic characteristics of the data and, consequently, in generalizing to new datasets.

Except for the VAE model, all proposed methodologies demonstrated sensitivity to structural changes, as evidenced by crossing the Upper Control Limit in the monitoring classes. The VAE model, in none of the six evaluated scenarios, exhibited sensitivity to structural changes, as previously discussed, showing varied behaviors and being outperformed in 100% of the cases by the SAE methodology and in 83.33% of the cases by the CAE methodology. However, the comparative analysis of the results indicates that the SAE methodology stands out in precisely quantifying these changes. When comparing the performances of the CAE and SAE models, the latter proved to be more effective in quantifying the magnitude of structural changes. Although both models show an increase in T^2 values as damage severity increases, the SAE provides a clearer separation between damage classes 1 and 2, indicating higher sensitivity to subtle variations in the data. Furthermore, among the six accelerometers evaluated, the SAE methodology did not present any overlap of the whiskers in the box-plots for monitoring data, whereas the CAE methodology exhibited whisker overlap for accelerometer 4, as shown in Figure 14d. These results reveal a 100% accuracy rate for the SAE model compared to 83.33% for the CAE model.

For a more detailed visualization of the T^2 value distributions and to facilitate model comparisons, box-plots were constructed (Figure 17, Figure 18, and Figure 19). These graphs enable the identification of the variability in T^2 values for each model and scenario, as well as highlight potential outliers. The analysis of the box-plots supports the previous conclusions, emphasizing the superiority of the SAE methodology in detecting and quantifying structural changes.

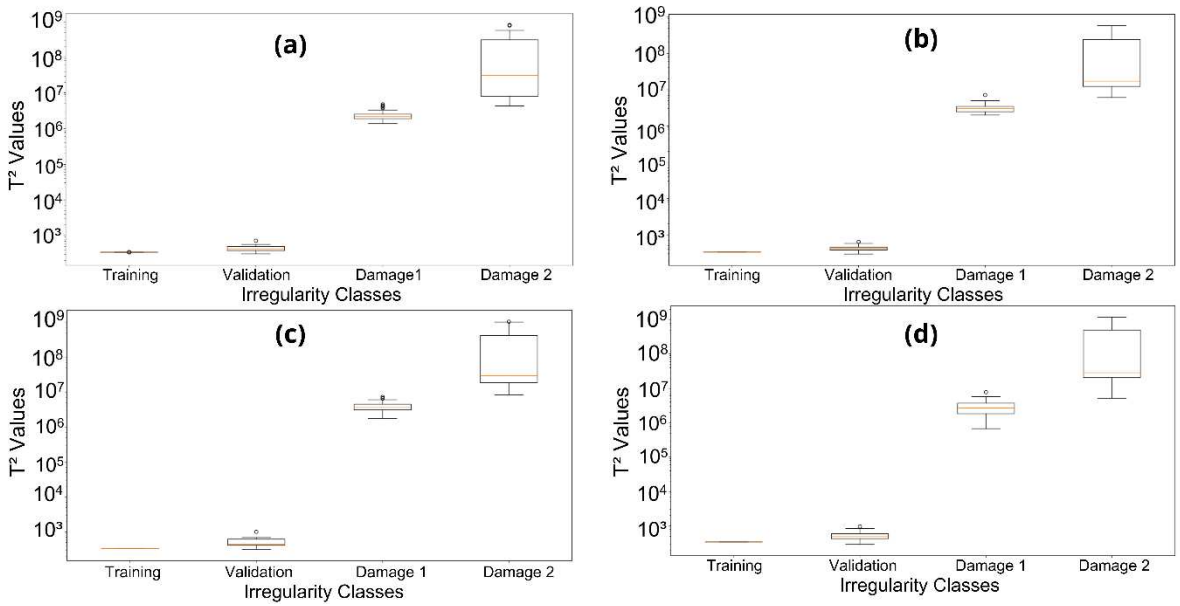

Figure 17: Boxplots of T² values from the CAE model. (a) Accelerometer 1; (b) Accelerometer 2; (c) Accelerometer 3; (d) Accelerometer 4.

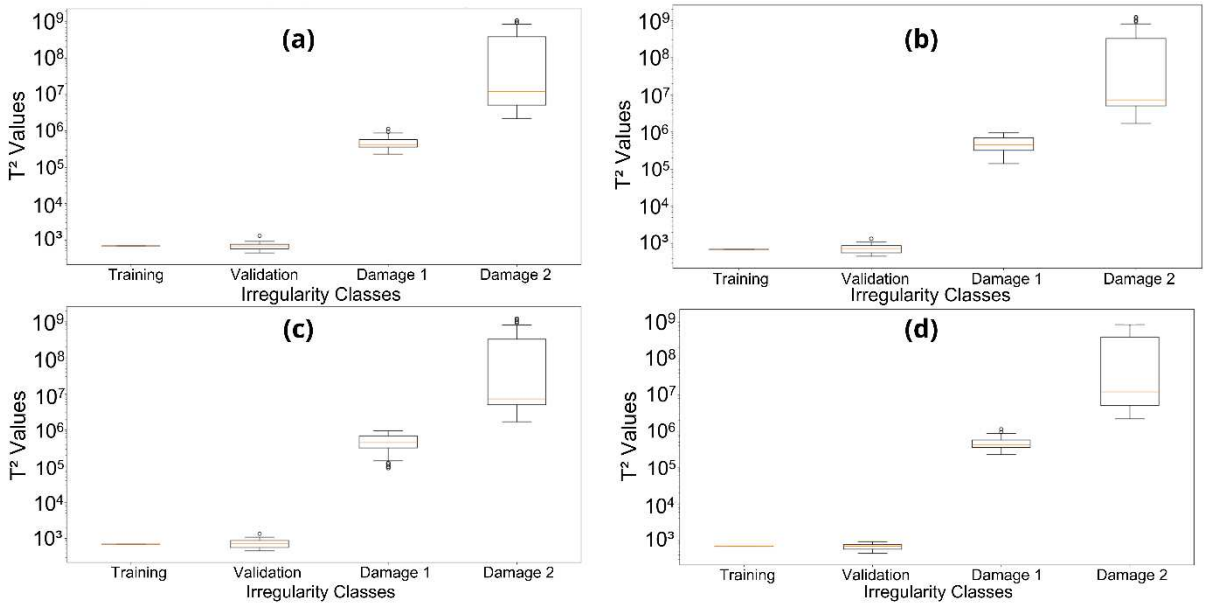

Figure 18: Boxplots of T² values from the SAE model. (a) Accelerometer 1; (b) Accelerometer 2; (c) Accelerometer 3; (d) Accelerometer 4.

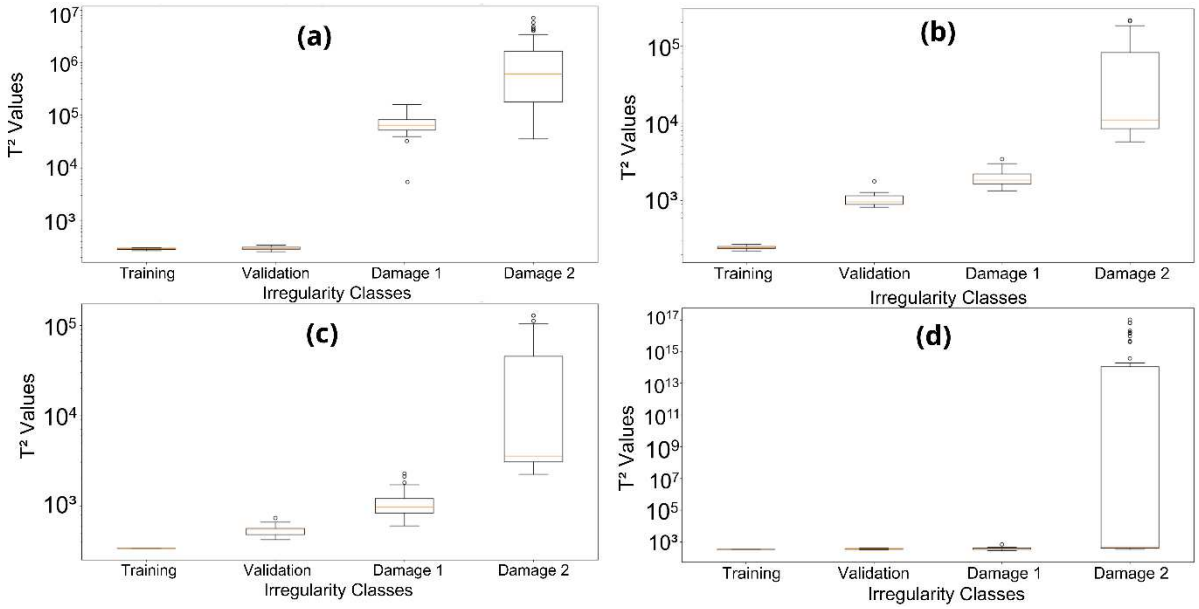

Figure 19: Boxplots of T2 values from the VAE model. (a) Accelerometer 1; (b) Accelerometer 2; (c) Accelerometer 3; (d) Accelerometer 4.

When submitting an AutoEncoder trained with data representative of a healthy structural state to new data from the same state, it is expected that the latent representations exhibit small variations around a central value, as evidenced by T^2 values close to the mean. This observation reflects the model's ability to capture the intrinsic characteristics of the healthy state. However, when the model is exposed to data from a damaged state, the latent representations tend to deviate significantly from the learned pattern, manifesting in high T^2 values. This divergence results from the model's inability to adequately reconstruct anomalous data, revealing the presence of damage or anomalies in the structure.

The results obtained in this study support the theoretical expectations, with the exception of the VAE. Specifically, for the SAE and CAE models, damage classes 1 and 2, associated with structural damage in railway wheels, exhibited significantly higher T^2 values, indicating a clear correlation between the severity of the damage and the distance of the latent representations from the mean. This trend demonstrates the effectiveness of AutoEncoders in quantifying structural damage severity, with SAE and CAE showing greater sensitivity to changes in data characteristics.

The comparative analysis between the CAE and SAE models revealed good performance in discriminating between different damage classes, with no overlap between the whiskers of the box-plots corresponding to distinct classes. However, the SAE model proved to be more sensitive to damage gradations, showing greater separation between damage classes 1 and

2. This observation suggests that SAE may be more suitable for accurately quantifying degradation stages.

In contrast to the results obtained for the CAE and SAE models, the VAE faced challenges in distinguishing between different levels of irregularities, as shown in Figure 19. The significant overlap between the box-plots corresponding to validation data and damage class 1, along with the deviation between training and validation data, indicates that the model failed to effectively capture the distinctive features of these classes. Although the results for accelerometers 2 and 3 were relatively acceptable, the same does not apply to accelerometers 1 and 4, as their box plots showed significant overlap. This limitation may be attributed to various factors, such as problem complexity, model architecture, and even the range of hyperparameters used in optimization performed by Optuna.

5.2. Computational Processing Times

To ensure reproducibility, result accuracy, and feasibility verification for model application, each stage of the experiment — from data loading to the final damage assessment using the T^2 index — was timed. The systematic repetition of the training, validation, and testing stages, along with the application of T^2, the generation of box-plots, and damage index quantification, guaranteed the reproducibility of the results. For each AE model, procedures were executed six times using the available accelerometers. The average execution times obtained for each model are presented in Table 6.

Table 6: Individual time per accelerometer and average execution time for each autoencoder.

Accelerometer	CAE	SAE	VAE
1	3h27min33s	5h28min17s	5h45min55s
2	3h34min11s	5h27min52s	6h53min19s
3	3h33min12s	5h35min50s	6h55min36s
4	3h31min01s	5h19min23s	8h19min21s
5	3h36min28s	5h45min34s	7h20min29s
6	3h24min27s	5h15min52s	6h54min23s
Average Time	**3h31min9s**	**5h28min48s**	**7h01min21s**

The evaluation of execution times across different methodologies revealed a clear hierarchy of computational complexity. Variational AutoEncoder models exhibited the longest processing time, averaging 7h01min21s. This characteristic is intrinsic to the probabilistic nature of VAEs, which require inference of complex latent distributions. In contrast, Sparse AutoEncoder models demonstrated an average execution time of 5h28min48s, indicating a considerable reduction compared to VAEs. Convolutional AutoEncoders emerged as the most efficient approach, with an average execution time of 3h31min9s, representing a 35.78% reduction compared to SAE and 50.11% compared to VAE. This computational performance advantage can be attributed to the convolutional architecture, which weights the hierarchical structure of the data and reduces the number of parameters to be estimated.

Despite the differences observed in execution times, it is interesting to highlight that, in a practical context, the average values obtained for all methodologies were quite reasonable. Furthermore, although the computational cost is relevant, it does not represent a hindering barrier to the application of any of these techniques. Consequently, the choice between VAE, SAE and CAE models should be guided mainly by the quality of the results produced and the suitability of each model to the specific problem under investigation.

6. Conclusions

In this study, three AE-based methodologies, combined with Hotelling's T^2 statistics, were compared to detect and quantify structural changes in railway wheels. SAE-T^2 stood out for its accurate anomaly detection and computational efficiency, outperforming VAE-T^2 and CAE-T^2. VAE, on the other hand, faced challenges in generalizing and correctly classifying data. While CAE-T^2 produced promising results, SAE-T^2 proved to be slightly more accurate in distinguishing between different levels of damage analyzed. Hyperparameter optimization proved to be important in enhancing models' performances, as evidenced by the variation in the results obtained for each model and structure.

The choice between SAE and CAE methodologies depends on factors such as runtime and available computational resources. SAE-T^2 is more accurate, while CAE-T^2 is faster. Considering current technological advances and available data processing capacity, runtime does not necessarily constitute a significant impediment for SAE-T^2. VAE, however, proved to be less effective due to difficulties in adjusting the suggestion range of its hyperparameters.

For future research, we propose the development of real-time monitoring systems coupling IoT (Internet of Things) with cloud computing to detect anomalies in an automated manner. Additionally, exploring advanced machine learning techniques, such as reinforcement learning and generative adversarial networks, may lead to significant advancements in SHM. Incorporating time-frequency analysis techniques may also enrich data analysis and enable the detection of subtle changes in structural behavior. Furthermore, to validate the methodology developed in this work, an experimental campaign will be planned, in which vehicles with wheels defects properly characterized will pass through an instrumented section of the track.

Author Contributions: Conceptualization, R.F., F.S., A.C.; Methodology, R.F., F.S., A.C.; Validation, R.M.; Formal analysis, R.M. and A.C.; Investigation, A.M., D.R.; Resources, D.R.; Data curation, A.M., A.G. and V.G.; Writing—original draft, R.M.; Writing—review \& editing, F.B., A.C., R.F., and D.R.; Supervision, A.C.; Project administration, F.B. and A.C.; Funding acquisition, F.B., A.C., D.R., A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financed by CAPES (Finance Code 0001), Conselho Nacional de Desenvolvimento Científico e Tecnológico — CNPq (Brazil) — Grants CNPq/FNDCT/MCTI 407256/2022-9, 402533/2023-2, 303982/2022-5, and 308008/2021-9 and Fundação de Amparo à Pesquisa do Estado de Minas Gerais—FAPEMIG — Grants BPD-00080-22 and APQ-00032-24. Coauthors Andreia Meixedo, Diogo Ribeiro, António Guedes e Vítor Gonçalves acknowledge the financial support from Base Funding - UIDB/04708/2020 with DOI 10.54499/UIDB/04708/2020 (https://doi.org/10.54499/UIDB/04708/2020) and Programmatic Funding - UIDP/04708/2020 with DOI 10.54499/UIDP/04708/2020 (https://doi.org/10.54499/UIDP/04708/2020) of the CONSTRUCT - Instituto de I\&D em Estruturas e Construções - funded by national funds through the FCT/MCTES (PIDDAC).

Data Availability Statement: The dataset described in this study was also addressed in the works of Magalhães et al. (2024) [3] and Guedes et al. (2023) [2] and is available a https://bit.ly/SHM-Data**.** All codes used in this study are available at https://bit.ly/SHM-CoDes**.**

Conflicts of Interest: The authors declare no conflicts of interest.

7. References

[1] Nielsen, J.C.O.; Johansson, A. Out-of-round railway wheels-a literature survey. Proc. of the Inst. of Mechanical Engineers, Part F: J. of Rail and Rapid Transit 2000, 214, 79–91. https://doi.org/10.1243/0954409001531351

[2] Guedes, A.; Silva, R.; Ribeiro, D.; Vale, C.; Mosleh, A.; Montenegro, P.; Meixedo, A. Detection of wheel polygonization based on wayside monitoring and artificial intelligence. Sensors **2023**, 23, 2188. https://doi.org/10.3390/s23042188

[3] Magalhaes, J.; Jorge, T.; Silva, R.; Guedes, A.; Ribeiro, D.; Meixedo, A.; Mosleh, A.; Vale, C.; Montenegro, P.; Cury, A. A strategy for out-of-roundness damage wheels identification in railway vehicles based on sparse autoencoders. Railway Engineering Sciences **2024**, 32, 421–443. https://doi.org/10.1007/s40534-024-00338-4

[4] Kaper, H. Wheel corrugation on Netherlands railways (NS): Origin and effects of "polygonization" in particular. Journal of Sound and Vibration **1988**, 120, 267–274. https://doi.org/10.1016/0022-460X(88)90434-8

[5] Snyder, T.W.; Stone, H.; Kristan, J. Wheel flat and out-of round formation and growth. Proceedings of the 2003 IEEE/ASME Joint Railroad Conference, 2003. **2003**, pp. 143–148. https://doi.org/10.1109/RRCON.2003.1204660

[6] Staskiewicz, T.; Firlik, B. Out-of-round tram wheels – current state and measurements. Archives of Transport **2018**, 45, 93–103. https://doi.org/10.5604/01.3001.0012.0946

[7] Wu, X.; Rakheja, S.; Qu, S.; Wu, P.; Zeng, J.; Ahmed, A. Dynamic responses of a high-speed railway car due to wheel polygonalisation. Vehicle System Dynamics **2018**, 56, 1817–1837. https://doi.org/10.1080/00423114.2018.1439589

[8] Vale, C.; Simoes, M.L. Prediction of Railway Track Condition for Preventive Maintenance by Using a Data-Driven Approach. Infrastructures **2022**, 7, 34. https://doi.org/10.3390/infrastructures7030034

[9] Malekjafarian, A.; OBrien, E.J.; Quirke, P.; Cantero, D.; Golpayegani, F. Railway Track Loss-of-Stiffness Detection Using Bogie Filtered Displacement Data Measured on a Passing Train. Infrastructures **2021**, 6, 93. https://doi.org/10.3390/infrastructures6060093

[10] Ngamkhanong, C.; Kaewunruen, S.; Costa, B.J.A. State-of-the-Art Review of Railway Track Resilience Monitoring. Infrastructures 2018, 3, 3. https://doi.org/10.3390/infrastructures3010003

[11] Du, J.; Zhao, X.; Su, J.; Li, B.; Duan, X.; Dong, T.; Lin, H.; Ren, Y.; Miao, Y.; Radamson, H.H. Review of Short-Wavelength Infrared Flip-Chip Bump Bonding Process Technology. Sensors **2025**, 25, 263. https://doi.org/10.3390/s25010263

[12] Finotti, R.P.; Barbosa, F.d.S.; Cury, A.A.; Gentile, C. A novel natural frequency-based technique to detect structural changes using computational intelligence. Procedia Engineering **2017**, 199, 3314–3319. https://doi.org/10.1016/j.proeng.2017.09.438

[13] Finotti, R.P.; Barbosa, F.d.S.; Cury, A.A.; Pimentel, R.L. Numerical and Experimental Evaluation of Structural Changes Using Sparse Auto-Encoders and SVM Applied to Dynamic Responses. Appl. Sci. **2021**, 11, 11965. https://doi.org/10.3390/app112411965

[14] Wei, Z.; Nunez, A.; Li, Z.; Dollevoet, R. Evaluating Degradation at Railway Crossings Using Axle Box Acceleration Measurements. Sensors **2017**, 17, 2236. https://doi.org/10.3390/s17102236

[15] Barke, D.; Chiu, W.K. Structural Health Monitoring in the Railway Industry: A Review. Struct. Health Monit. **2005**, 4, 81–93. https://doi.org/10.1177/1475921705049764

[16] Bianchi, G.; Fanelli, C.; Freddi, F.; Giuliani, F.; Placa, A.L. Systematic review railway infrastructure monitoring: From classic techniques to predictive maintenance. Advances in Mechanical Engineering **2025**, 17, 16878132241285631. https://doi.org/10.1177/16878132241285631

[17] Wu, P.; Zhang, F.; Wang, J.; Wei, L.; Huo, W. Review of wheel-rail forces measuring technology for railway vehicles. Advances in Mechanical Engineering **2023**, 15, 16878132231158991. https://doi.org/10.1177/16878132231158991

[18] Ni, Y.Q.; Zhang, Q.H. A Bayesian machine learning approach for online detection of railway wheel defects using track-side monitoring. Struct. Health Monit. **2021**, 20, 1536–1550. https://doi.org/10.1177/1475921720921772

[19] Mohammadi, M.; Mosleh, A.; Vale, C.; Ribeiro, D.; Montenegro, P.; Meixedo, A. Smart railways: AI-based track-side monitoring for wheel flat identification. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit **2025**, 0, 09544097251313570. https://doi.org/10.1177/09544097251313570

[20] Amini, A.; Entezami, M.; Huang, Z.; Rowshandel, H.; Papaelias, M. Wayside detection of faults in railway axle bearings using time spectral kurtosis analysis on high-frequency acoustic emission signals. Advances in Mechanical Engineering **2016**, 8. https://doi.org/10.1177/1687814016676000

[21] Wu, Y.;Wang, J.; Liu, M.; Jin, X.; Hu, X.; Xiao, X.;Wen, Z. Polygonal wear mechanism of high-speed wheels based on full-size wheel-rail roller test rig. Wear **2022**, 494-495, 204234. https://doi.org/10.1016/j.wear.2021.204234

[22] Ye, Y.; Wei, L.; Li, F.; Zeng, J.; Hecht, M. Multislice Time-Frequency image Entropy as a feature for railway wheel fault diagnosis. Measurement **2023**, 216, 112862. https://doi.org/10.1016/j.measurement.2023.112862

[23] Ye, Y.; Bin, Z.; Ping, H.; Bo, P. OORNet: A Deep Learning Model for on-Board Condition Monitoring and Fault Diagnosis of out-of-Round Wheels of High-Speed Trains. Measurement **2022**, 199, 111268. https://doi.org/10.1016/j.measurement.2022.111268

[24] Zhou, J.; Zhang, Z.; Jin, Z.; Kong, X.; Wang, X.; Liu, H. Indirect measurement of bridge surface roughness using vibration responses of a two-axle moving vehicle based on physics-constrained generative adversarial network. Journal of Sound and Vibration **2025**, 595, 118763. https://doi.org/10.1016/j.jsv.2024.118763

[25] Lee, J.; Song, S.H.; Kim, R.E.; Lee, J.H. A Generative adversarial network model for estimating temporal frequency variation of vehicle-bridge interaction using modified Stockwell transform. Journal of Sound and Vibration **2025**, 594, 118655. https://doi.org/10.1016/j.jsv.2024.118655

[26] Finotti, R.P.; Gentile, C.; Barbosa, F.d.S.; Cury, A.A. Structural novelty detection based on sparse autoencoders and control charts. Structural Engineering and Mechanics **2022**, 81, 647–664. https://doi.org/10.12989/sem.2022.81.5.647

[27] Spínola Neto, M.; Finotti, R.P.; Barbosa, F.d.S.; Cury, A.A. Structural Damage Identification Using Autoencoders: A Comparative Study. Buildings 2024, 14, 2014. https://doi.org/10.3390/buildings14072014

[28] Finotti, R.P.; Barbosa, F.d.S.; Cury, A.A.; Pimentel, R.L. Novelty Detection Using Sparse Auto-Encoders to Characterize Structural Vibration Responses. Arabian Journal for Science and Engineering **2022**, 47, 13049–13062. https://doi.org/10.1007/s13369-022-06732-6

[29] Xu, H.; Chen, W.; Zhao, N.; Li, Z.; Bu, J.; Li, Z.; Liu, Y.; Zhao, Y.; Pei, D.; Feng, Y.; et al. Unsupervised Anomaly Detection via Variational Auto-Encoder for Seasonal KPIs in Web Applications. Proceedings of the 2018 World Wide Web Conference **2018**, pp. 187–196. https://doi.org/10.1145/3178876.3185996

[30] Ehsani, N.; Aminifar, F.; Mohsenian-Rad, H. Convolutional autoencoder anomaly detection and classification based on distribution PMU measurements. IET Generation, Transmission & Distribution **2022**, 16, 2816–2828. https://doi.org/10.1049/gtd2.12424

[31] ANSYSR. Academic Research; ANSYSR: Canonsburg, Pensilvania, USA, 2018.

[32] Costa, C.; Ribeiro, D.; Jorge, P.; Ruben, S.; Rui, C.; Arede, A. Calibration of the Numerical Model of a Short-span Masonry Railway Bridge Based on Experimental Modal Parameters. Procedia Engineering **2015**, 114, 846–853. https://doi.org/10.1016/j.proeng.2015.08.038

[33] Meixedo, A.; Ribeiro, D.; Calcada, R.; Delgado, R. Global and Local Dynamic Effects on a Railway Viaduct with Precast Deck, 2014. Paper presented at the Proceedings of the Second International Conference on Railway Technology: Research, Development and Maintenance, Ajaccio, Corsica, France, 8–11 April.

[34] Castillo-Mingorance, J.M.; Sol-Sanchez, M.; Moreno-Navarro, F.; Rubio-Gamez, M.C. A Critical Review of Sensors for the Continuous Monitoring of Smart and Sustainable Railway Infrastructures. Sustainability **2020**, 12, 9428. https://doi.org/10.3390/su12229428

[35] Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning; MIT Press: Cambridge, 2016.

[36] Montgomery, D.C. Introduction to Statistical Quality Control, Sixth Edition; John Wiley & Sons: Hoboken, 2009.

[37] Finotti, R.P.; Felicio, C.; Oliveira, P.H.E.; Barbosa, F.d.S.; Cury, A.A.; Silva, R.C. Novelty detection on a laboratory benchmark slender structure using an unsupervised deep learning algorithm. Lat. Am. J. Solids Struct. **2023**, 20, e512. https://doi.org/10.1590/1679-78257591

[38] Resende, L.; Finotti, R.; Barbosa, F.; Garrido, H.; Cury, A.; Domizio, M. Damage identification using convolutional neural networks from instantaneous displacement measurements via image processing. Struct. Health Monit. **2024**, 23, 1627–1640. https://doi.org/10.1177/14759217231193102

[39] Li, Z.; Liu, F.; Yang,W.; Peng, S.; Zhou, J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Transactions on Neural Networks and Learning Systems **2022**, 33, 6999–7019. https://doi.org/10.1109/TNNLS.2021.3084827

[40] Cataltas, O.; Tutuncu, K. Detection of protein, starch, oil, and moisture content of corn kernels using one-dimensional convolutional autoencoder and near-infrared spectroscopy. PeerJ Computer Science **2023**, 9, e1266. https://doi.org/10.7717/peerj-cs.1266

[41] Qi, J.; Du, J.; Siniscalchi, S.M.; Ma, X.; Lee, C.H. On Mean Absolute Error for Deep Neural Network Based Vector-to-Vector Regression. IEEE Signal Processing Letters **2020**, 27, 1485–1489. https://doi.org/10.1109/LSP.2020.3016837

[42] Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework. Association for Computing Machinery **2019**, p. 2623–2631. https://doi.org/10.1145/3292500.3330701

[43] Bergstra, J.; Yamins, D.; Cox, D. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures, 2013. Paper presented at the Proceedings of the 30th International Conference on Machine Learning, Atlanta, Geoorgia, USA, 16–21 June.

[44] Microsoft. Microsoft Visual Studio Code; Version 1.93: Redmond,Washington, USA, 2009.

[45] UIC. Code of Practice for the Loading and Securing of Goods on Railway Wagons; UIC: Paris, Ile-de-France, France, 2022.

[46] Ribeiro, D.; Calcada, R.; Delgado, R.; Brehm, M.; Zabel, V. Finite-element model calibration of a railway vehicle based on experimental modal parameters. Vehicle System Dynamics **2013**, 51, 821–856. https://doi.org/10.1080/00423114.2013.778416

[47] Braganca, C.; Neto, J.; Pinto, N.; Montenegro, P.A.; Ribeiro, D.; H., C.; Calçada, R. Calibration and validation of a freight wagon dynamic model in operating conditions based on limited experimental data. Vehicle System Dynamics **2022**, 60, 3024–3050. https://doi.org/10.1080/00423114.2021.1933091

[48] Montenegro, P.; Heleno, R.; Carvalho, H.; Calcada, R.; Baker, C. A comparative study on the running safety of trains subjected to crosswinds simulated with different wind models. Journal of Wind Engineering and Industrial Aerodynamics **2020**, 207, 104398. https://doi.org/10.1016/j.jweia.2020.104398

[49] Mosleh, A.; Montenegro, P.; Alves Costa, P.and Calcada, R. An approach for wheel flat detection of railway train wheels using envelope spectrum analysis. Structure and Infrastructure Engineering **2020**, 17, 1710–1729. https://doi.org/10.1080/15732479.2020.1832536

[50] Ribeiro, D.; Calcada, R.; Brehm, M.; Zabel, V. Calibration of the numerical model of a track section over a railway bridge based on dynamic tests. Structures **2021**, 34, 4124–4141. https://doi.org/10.1016/j.istruc.2021.09.109

[51] Mosleh, A.; Montenegro, P.A.; Costa, P.A.; Calcada, R. Railway Vehicle Wheel Flat Detection with Multiple Records Using Spectral Kurtosis Analysis. Appl. Sci. **2021**, 11, 4002. https://doi.org/10.3390/app11094002

[52] Montenegro, P.; Neves, S.; Calcada, R.; Tanabe, M.; Sogabe, M. Wheel–rail contact formulation for analyzing the lateral train–structure dynamic interaction. Computers & Structures **2015**, 152, 200–214. https://doi.org/10.1016/j.compstruc.2015.01.004

[53] Montenegro, P.A.; Calcada, R.; Pouca, N.V.; M., T. Running safety assessment of trains moving over bridges subjected to moderate earthquakes. Earthquake Engineering & Structural Dynamics **2016**, 45, 483–504. https://doi.org/10.1002/eqe.2673

[54] Montenegro, P.A.; Barbosa, D.; Carvalho, H.; Calcada, R. Dynamic effects on a train-bridge system caused by stochastically generated turbulent wind fields. Engineering Structures **2020**, 211, 110430. https://doi.org/10.1016/j.engstruct.2020.110430

[55] Goncalves, V.; Mosleh, A.; Vale, C.; Montenegro, P. Wheel Out-of-Roundness Detection Using an Envelope Spectrum Analysis. Sensors **2023**, 23, 2138. https://doi.org/10.3390/s23042138

[56] Sugiyama, H.; Araki, K.; Suda, Y. On-line and off-line wheel/rail contact algorithm in the analysis of multibody railroad vehicle systems. J. Mech. Sci. Technol. **2009**, 23, 991–996. https://doi.org/10.1007/s12206-009-0327-2

[57] Hertz, H. Ueber die Beruhrung fester elastischer Korper. De Gruyter **1882**, pp. 156–171. https://doi.org/10.1515/9783112342404-004

[58] Kalker, J. Book of Tables for the Herzian Creep-force Law; Faculty of Technical Mathematics and Informatics, Delft University of Technology: Delft, The Netherlands, 1996.

[59] MathWorks. MATLABR. version R2022a; The MathWorks Inc.: Natick, Massachusetts, USA, 2022.

[60] Mohammadi, M.; Mosleh, A.; Vale, C.; Ribeiro, D.; Montenegro, P.; Meixedo, A. An Unsupervised Learning Approach for Wayside Train Wheel Flat Detection. Sensors **2023**, 23, 1910. https://doi.org/10.3390/s23041910

[61] Mosleh, A.; Meixedo, A.; Ribeiro, D.; Montenegro, P.; Calcada, R. Early wheel flat detection: an automatic data-driven wavelet based approach for railways. Vehicle System Dynamics **2022**, 61, 1644–1673. https://doi.org/10.1080/00423114.2022.2103436

[62] Mosleh, A.; Meixedo, A.; Ribeiro, D.; Montenegro, P.; Calcada, R. Automatic clustering-based approach for train wheels condition monitoring. International Journal of Rail Transportation **2023**, 11, 639–664. https://doi.org/10.1080/23248378.2022.2096132

[63] Johansson, A.; Andersson, C. Out-of-round railway wheels—a study of wheel polygonalization through simulation of three-dimensional wheel–rail interaction and wear. Vehicle System Dynamics **2023**, 43, 539–559. https://doi.org/10.1080/00423110500184649

[64] Cai, W.; Chi, M.; Tao, G.q.; Wu, X.; Wen, Z.f. Experimental and Numerical Investigation into Formation of Metro Wheel Polygonalization. Shock and Vibration **2019**, 2019, 1538273. https://doi.org/10.1155/2019/1538273

4 FINAL CONSIDERATIONS

This study focused on the comparison of three AutoEncoder (AE) architectures for detecting and quantifying structural anomalies in railway vehicle wheels, highlighting the application of Machine Learning techniques in Structural Health Monitoring. The results demonstrated that the approach based on the Sparse AutoEncoder (SAE) model, combined with statistical analysis using the Hotelling T^2 control chart, exhibited superior performance in accurately classifying the structural conditions of the wheels, despite a higher computational cost compared to the Convolutional AutoEncoder (CAE) model.

The evaluation of the models revealed that VAE faced significant limitations in generalizing damage patterns, resulting in lower efficiency in distinguishing between normal and anomalous conditions. On the other hand, CAE showed potential for applications where computational efficiency is prioritized, although its accuracy was slightly lower than that of SAE. The optimization of hyperparameters proved to be a fundamental factor in enhancing model performance, highlighting the need for careful calibration for each specific application.

Considering technological advancements and the increasing availability of computational resources, adopting more precise methodologies, such as SAE-T^2, tends to be a viable alternative for real-time structural monitoring applications. However, the choice between SAE and CAE should consider specific processing demands and execution time. In particular, the practical implementation of these models in real railway systems requires a thorough analysis of operational conditions and the inherent challenges in acquiring vibration data in the field.

4.1 Future works

For future research, we propose:

- **Integrate IoT and cloud computing solutions:** Automate anomaly detection and enhance the efficiency of structural monitoring.
- **Explore advanced machine learning techniques:** Investigate reinforcement learning and generative adversarial networks to improve the identification of complex patterns.
- **Conduct experimental validation:** Perform controlled environment testing to consolidate the proposed methodology’s application in real-world scenarios.