

Early identification of out-of-roundness damage wheels in railway freight vehicles using a wayside system and a stacked sparse autoencoder

T. Jorge, J. Magalhães, R. Silva, A. Guedes, D. Ribeiro, C. Vale, A. Meixedo, A. Mosleh, P. Montenegro & A. Cury

To cite this article: T. Jorge, J. Magalhães, R. Silva, A. Guedes, D. Ribeiro, C. Vale, A. Meixedo, A. Mosleh, P. Montenegro & A. Cury (2025) Early identification of out-of-roundness damage wheels in railway freight vehicles using a wayside system and a stacked sparse autoencoder, Vehicle System Dynamics, 63:2, 232-257, DOI: [10.1080/00423114.2024.2333771](https://doi.org/10.1080/00423114.2024.2333771)

To link to this article: <https://doi.org/10.1080/00423114.2024.2333771>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 27 Mar 2024.



Submit your article to this journal [↗](#)



Article views: 1355



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 6 View citing articles [↗](#)

Early identification of out-of-roundness damage wheels in railway freight vehicles using a wayside system and a stacked sparse autoencoder

T. Jorge^a, J. Magalhães^a, R. Silva^b, A. Guedes^a, D. Ribeiro ^a, C. Vale ^b, A. Meixedo^b, A. Mosleh ^b, P. Montenegro ^b and A. Cury^c

^aCONSTRUCT-LESE, School of Engineering, Polytechnic of Porto, Porto, Portugal; ^bCONSTRUCT-LESE, Faculty of Engineering, University of Porto, Porto, Portugal; ^cGraduate Program in Civil Engineering, Federal University of Juiz de Fora, Juiz de Fora, Brazil

ABSTRACT

Early identification of wheel defects can prevent serious damage to railways, considerably lowering maintenance costs for both railway administrations and rolling stock operators. Within this context, an unsupervised methodology based on artificial intelligence techniques is presented, which allows the detection and classification of out-of-roundness damage wheels, such as wheel flats and polygonal wheels, based on dynamic responses induced on the track by crossing freight railway vehicles. The methodology involves the following steps: (i) data collection and pre-processing, (ii) feature extraction (iii) data fusion and (iv) feature discrimination. In the first phase, an FFT algorithm is applied to the acceleration track responses. Then, features are extracted after training a Stacked Sparse Autoencoder, in which the main features of the responses are obtained after a compression stage using an encoder network. This lower dimensional layer forces the model to learn a compression of the input data. Then, these extracted features are merged using the Mahalanobis distance, which enhances the sensitivity to the damage recognition. Posteriorly, an outlier analysis is performed to distinguish a healthy wheel from a defective one and a cluster analysis to discriminate the two types of out-of-roundness (OOR) damage and classify the severity of each type of damage.

ARTICLE HISTORY

Received 1 November 2023
Revised 7 February 2024
Accepted 18 March 2024

KEYWORDS

OOR wheel damage; damage identification; freight trains; wayside condition monitoring; Stacked Sparse Autoencoder

1. Introduction

Rail vehicle wheels are subjected to constant wear and tear due to continuous use and inadequate maintenance plans. Over time, this wear and tear can generate more significant defects, so it is extremely important to use sensing to understand when maintenance of both the track and rolling stock is required. Over the last few decades, researchers have proposed two types of monitoring systems that are commonly referred as on-board and wayside systems, being the main difference in the location of the measuring devices. Most

CONTACT T. Jorge  tosjo@isep.ipp.pt

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

on-board techniques are based on acoustic, vibration, ultrasonic and image detection technologies [1] in which sensors are installed in the vehicle for management of the condition of both the rolling stock and railway infrastructure. However, this type of system entail high costs envisaging the vehicle wheel condition monitoring, since sensors need to be installed in every vehicle in order to achieve a good representativity [2]. On the other hand, wayside systems are a more viable solution for identifying wheel defects, since the condition of all wheels is indirectly estimated by measuring the in-service train responses on the railway track [3].

The most common wheel defects are known as OOR. As the name infers, this defect is related to deformations on the surface of the wheel circumference that make it not round. This deformation can have many forms and different causes, leading to faults on the vehicle and on the track due to the higher dynamic force values [4–6]. The OOR damages most frequently studied are wheel flats (discrete defects) and polygonal wheels (periodic irregularity).

Many authors have reported the impact of the flats on the measurement responses as a function of their shape and size, vehicle characteristics and speed [7–10]. Wheel flats induce high dynamic impact loads on the railway infrastructure, which causes significant damage to both tracks and vehicles, such as broken axles, damaged rolling bearings and cracks in the wheels, rails and sleepers [11]. Moreover, it also generates high vibrations transmitted to the rolling stock, inducing significant stresses and causing damage to the suspension system, frame, and carbody [12,13].

In the case of polygonal wheels, the formation mechanism can be roughly classified into three categories, namely, (1) initial wheel defects [14,15], (2) a natural vibration of the vehicle–track system [16,17] and (3) thermoelastic instability [18]. The formation mechanism is strongly contingent on the vehicle type and its operational characteristics, with the resultant effects on both the track and the vehicle intricately linked to the prevalence of dominant harmonic orders [19–22]. For example, Tao et al. [20] studied asymmetrical polygonal wear on metro wheels and obtained harmonic orders of 12–14, while another study by Tao et al. [19] found that the most common harmonic order in Chinese freight trains is 18–19. Although a polygonal wheel could generate less impact on the track than a wheel flat, this defect can cause serious damage to the track due to resonance effects in the wheelsets by increasing the vibration in these components [23].

To identify the occurrence of these OOR anomalies in train wheels, innovative methodologies have been developed combining signal processing and automatic learning techniques. These methodologies are based on the extraction of damage sensitivity features, making possible the distinction between undamaged and damaged situations. Recently, many studies have shown good results in detecting OOR damages using distinct approaches [24–29]. Mishra et al. [24] concluded that a single FBG (Fiber Bragg Grating) sensor, positioned along the rail can detect the presence of wheel flats. Lourenço et al. [25] created an automatic algorithm for diagnosing OOR defects. The proposed methodology consists of isolating specific defective wheels and then identifying the severity of the damage, using only acceleration and strain data measured on the track. In another study, Krummenacher et al. [29], proposed two machine-learning methods to automatically detect a defective wheel during operation by measuring the vertical wheel force and using a sensor system permanently installed on the railway track.

By extracting the features, it is possible to transform the measured data into an alternative form of information where the correlation with damage is more easily observed [30]. There are various techniques for feature extraction, the most common being CWT (continuous wavelet transform) [31–33], PCA (principal component analysis) [31,34,35], AR (autoregressive models) [31,36,37] and ARX (autoregressive models with exogenous) [27,31,38]. Guedes et al. [38] developed an unsupervised methodology that allows the detection of polygonal wheels on freight vehicles using CWT and ARX features, using 6 accelerometers, 3 on each side of the railway track. Mosleh et al. [32] introduced a methodology relying on a single-sensor wayside monitoring system for detecting wheel flats by extracting AR features from acceleration records. Moreover, Mohammadi et al. [31] adopted linear features, particularly, CWT, PCA, ARX and AR, and their research findings indicate that the AR and ARX extraction methods are more effective than CWT and PCA in the context of wheel flat damage detection. Although these studies have demonstrated robust results in detection, they show some limitations in dealing with the non-linear effects associated with operational effects, such as the train speed. Due to the inability of these linear techniques to attenuate operation effects, the inclusion of different speeds simultaneously becomes a challenge.

For most of these studies, a feature normalisation stage is used when it is necessary to remove the operational effects. Usually, latent variable methods are used, such as PCA [39] or multiple linear regression (MLR) [40]. Thus, to deal with the non-linear phenomena and reduce the large amounts of data without relying on the aforementioned methods, the first artificial neural networks were created to inherently execute this task [41]. Over the years, new and more advanced networks have been developed, with the Autoencoder (AE) being one of the main types. Later, Olshausen et al. [42] proposed the creation of sparse coding and Ng et al. [43] introduced a sparse penalisation term in the hidden layer, to acquire more compact and effective features from lower-dimensional data. For some datasets with complex relationships between characteristics, employing more than one SAE is necessary, which is why stacked sparse autoencoders (SSAE) are commonly used [44]. Nowadays, these types of AE are widely used in various scientific fields, but specifically for structural engineering only a few studies are found, and most are about damage detection on bridges [45–47]. Amaral et al. [45], proposed an approach for detecting structural anomalies based on SAE and Shewhart's T control chart. The characteristics extracted by the SAE models directly from the accelerations allow detecting early abnormal behaviours in structures. In another similar study, Amaral et al. [46], applied the same approach to a numerical model of a highway viaduct in Brazil reaching more than 99% of successful classifications of structural damage with the combination of SAE and Support Vector Machine. Pathirage et al. [47] proposed a framework utilizing SAE's potential for dimensionality reduction, which facilitated the assessment of damage detection performance on a pre-stressed concrete bridge.

Many of these AE algorithms already performed feature classification in order to discriminate the different types of damages and severities. However, the AE methods can also be combined with other techniques commonly used to classify the data. They can be used unsupervised methods, in which the model is trained without any labels provided to the training dataset, such as k-means [36], self-organising maps (SOM) [48], isolation forest [25], gaussian mixture model [49], k-medoids [50] and cluster analysis [51], or supervised methods, which assume that the hierarchy of database categories is known, such as Naive

Bayes classifiers [52], k-Nearest Neighbours classifiers [53], Support Vector Machine [54] and decision trees [55].

This study aims to develop an automatic methodology for detecting and classifying two types of OOR damages, flats and polygonal wheels, based on acceleration responses obtained from a virtual wayside monitoring system. The methodology comprises seven stages: (i) data collection, (ii) data pre-processing, (iii) training on a Stacked Sparse Autoencoder, (iv) feature extraction, (v) data fusion, (vi) outlier analysis and (vii) cluster analysis. Thus, this numerical study aims to make innovative contributions to the current state of the art, with special emphasis on the following aspects:

- Use of a Autoencoder neural network for feature extraction. This technique has been applied in the structural engineering area, but the application in the railway engineering area is very scarce;
- The application of Autoencoder in the railway area can be very valuable, instead of the classical linear feature techniques, since allows an efficient removal of the operational effects, such as train speed, from the data;
- The proposed methodology allows to detect and classify two types of OOR damages within a wide range of train speeds, unlike most methodologies that show results for only one speed;
- Additionally, the proposed methodology allows for a reduction in the number of accelerometers compared to other studies [25,38], showing good efficiency in damage classification using only two accelerometers. The application of sensors on both sides of the rail allows the methodology to be applied without knowing on which side is located the damage.

2. Numerical simulations

To obtain the responses from the railway track, it is necessary to apply sensors on the track that can measure the dynamic responses of the crossing trains. As it is not always possible to install them due to the high costs involved, simulations have been carried out based on calibrated models of the track [56], vehicle [57] and considering their dynamic interaction, as well as including the track irregularities. Simulations include cases with healthy wheels and with OOR wheels (polygonal wheels and wheel flats).

2.1. Vehicle-track interaction

To simulate the dynamic interactions between the vehicle and the track, an in-house software VSI – Vehicle-Structure Interaction Analysis is used, which was initially developed by Neves et al. [58] and later improved by Montenegro et al. [59,60]. This numerical tool was validated with numerical examples and experimental data (see [38], [59], [60] and [61]).

The train and the track are initially modelled separately in ANSYS® [62] and then coupled using a wheel-rail contact model implemented in MATLAB® [63]. The track comprises multiple-layers with rail and sleepers are represented by beam elements, while the ballast is represented by discrete mass point elements. Elastic elements are used to connect the various layers, simulating the interface between them.

Table 1. Main properties adopted for vehicle and track models.

Vehicle		Track	
Parameter	Value	Parameter	Value
Mass of carbody, m_{cb} (ton)	13.5	Area of rail, A_r (m^2)	7.67×10^{-4}
Roll moment of inertia in carbody, $I_{cb,x}$ ($t.m^2$)	49	Inertia of rail, I_r (m^4)	30.38×10^{-6}
Pitch moment of inertia in carbody, $I_{cb,y}$ ($t.m^2$)	673	Longitudinal stiffness of the rail pad, K_p (N/m)	20.00×10^6
Yaw moment of inertia in carbody, $I_{cb,z}$ ($t.m^2$)	665	Modulus of elasticity of the sleeper, E_s (N/m^2)	40.90×10^9
Longitudinal stiffness of suspensions, $K_{1,x}$ (kN/m)	44,981	Vertical stiffness of the ballast, $K_{b,z}$ (N/m)	30.00×10^6
Vertical stiffness of suspensions, $K_{1,z}$ (kN/m)	1860	Lateral damping of the ballast, $C_{b,y}$ ($N \cdot s/m$)	15.00×10^3
Vertical damping of suspensions, $C_{1,z}$ ($kN \cdot s/m$)	16.7	Vertical stiffness of the foundation, $K_{f,z}$ (N/m)	20.00×10^6

For the train model, a Laagrass model was used, specifically designed for container transport and capable of reaching speeds up to 120 km/h [57]. The vehicle model integrates spring elements to simulate suspensions, mass elements to represent the mass and inertia at the centre of gravity of the wagon components. For the structural connection of these elements, rigid beam elements are adopted. Table 1 outlines the properties of the track and vehicle model.

The Hertzian theory is applied to calculate the normal contact forces, while the USETAB routine [64] allows evaluating the tangential forces resulting from the rolling contact phenomenon.

The modelling of track irregularities and OOR wheel defects is also performed in MATLAB® [63]. These models are then included in the constraint equations that couple the train to the track in the VSI software. Using Lagrange multipliers, this tool connects wheel displacements with the track nodal displacements, incorporating track irregularities and OOR profiles through a set of constraint equations. These equations are solved alongside equilibrium equations, forming a unified system that couples the two subsystems (see [60,61] for a full description of the coupling model). Including the track irregularities directly in the constraint equations are beneficial since it avoids explicitly considering them in the FEM model. Moreover, it also allows to include the wheel and OOR profiles as a periodic rail irregularity, simplifying the analysis. Thus, before starting the analysis, the VSI software imports track irregularities and OOR profiles from a MATLAB file, which contains one column related to the position of the track and four columns corresponding to the vertical and lateral irregularities in each rail in each respective position, and adds it to the constraint equations of the system throughout the dynamics analysis. Then, the system of dynamic equations is solved using the typical Newmark integration method.

The impact of track irregularity on wheel defect classification is significant, as a lower track quality tends to result in a higher likelihood of misclassification. A comprehensive exploration of this influence was previously undertaken by Mosleh et al. as documented in their research study [56]. Hence, this study does not delve into the impact of track quality; rather, it primarily centres on the utilisation of AI-based methodologies for the detection and classification of wheel defects. For this, four different profiles are used, based on European Standard EN 13848–2 [65]. In this case, power spectral density (PSD) curves

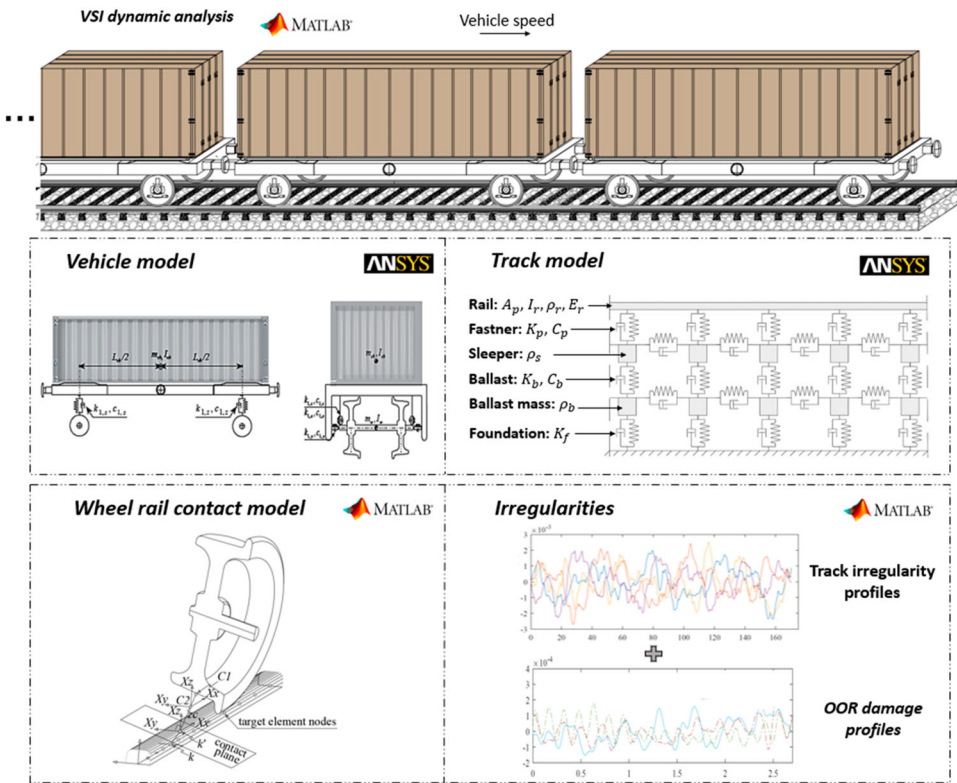


Figure 1. Vehicle-track interaction model schematisation.

are generated and artificial irregularity profiles are created in the vertical and transverse directions.

For OOR defects, several profiles are defined based on real measurements in representative profiles of the various severities. Subsequently, the structural matrices are imported and integrated into the VSI software using a fully coupled strategy.

More details on track and train properties, including additional information on train-track interaction, can be found in previous publications [59,66]. Figure 1 shows the vehicle-track interaction model schematisation and some details about the train and track subsystems, as well as the track irregularities and wheel-rail contact model.

2.2. Virtual wayside monitoring system

According to the conclusions reached in other studies with this type of damage [31,38], the most promising location of the sensors is on the rail. To reduce installation and maintenance costs, it is important to reduce the number of sensors without compromising the quality of the results. Therefore, based on a sensitivity analysis performed during the study, only two vertical unidirectional accelerometers were used on each rail, 4 in total (A1 to A4), since they provided a higher cost-benefit ratio. These sensors are mounted on the rail at mid-span between two sleepers, as illustrated in Figure 2.

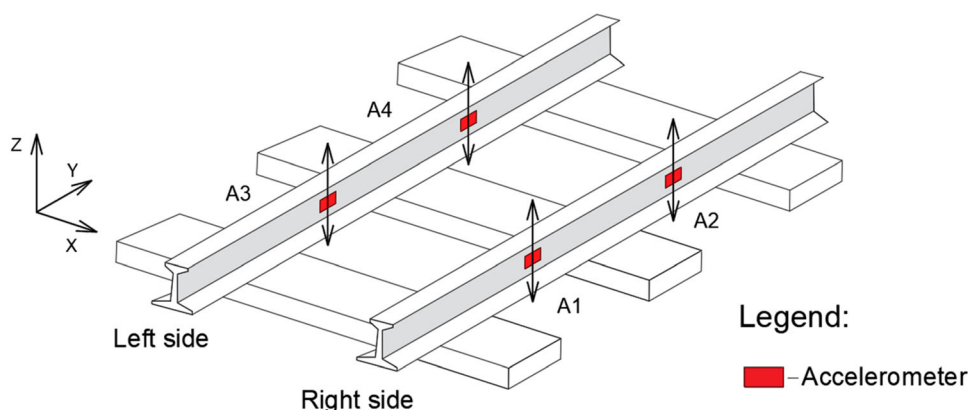


Figure 2. 3D wayside monitoring system.

2.3. Scenarios

2.3.1. Overview

To test and validate the automatic methodology for the detection and classification of wheel defects, different simulation scenarios with perfect wheels (baseline) and OOR wheels (damage) were defined for a five-wagon Laagrss freight train. Based on different crossing speeds, loading schemes and rail irregularities profiles, 120 baseline and 180 damage scenarios (90 for polygonal wheels and 90 for wheel flats) were defined, as summarised in Table 2.

For baseline scenarios, six loading schemes with five different speeds (40, 60, 80, 100 and 120 km/h) and four rail irregularity profiles were considered, giving a total of 120 scenarios. The loading schemes used are: empty, half-loaded, full-loaded and, additionally, three other loading schemes with unbalanced loads within the limits recommended by UIC [67].

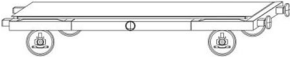
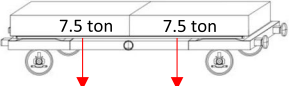
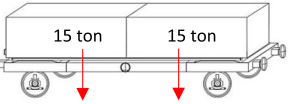
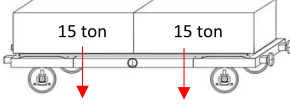
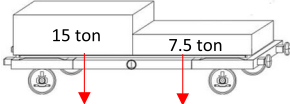
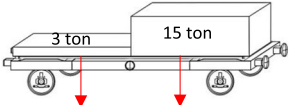
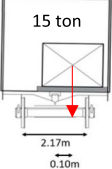
For the damage cases, a total of 90 scenarios were generated for polygonal wheels. These scenarios including 1 load x 1 irregularity x 3 speeds x 3 severities (H1, H2 and H3) x 10 different irregularity profiles within each specified harmonic interval. In these 10 irregularity profiles, based on the defect amplitude, two amplitudes are considered, five simulations for A1 and five simulations for A2. Similarly, 90 scenarios were created for wheel flats, involving 1 load x 1 irregularity x 3 speeds x 3 severities (F1, F2 and F3) x 10 different irregularity profiles within each flat length interval.

The two types of defects are simulated on different wheels and wagons for full load cases and for three crossing speeds (60, 80 and 100 km/h).

2.3.2. Polygonal wheel

A polygonal wheel is a wheel that shows a periodic radial tread deviation from the mean wheel radius [67]. This wave phenomenon can be characterised by the different wavelengths formed (harmonics θ), and by the levels of irregularity (Lw) associated with each of the harmonics. According to Peng [68], it is considered a polygonal wheel when the wear amplitude assumes values above 0.2 mm and wavelengths between 140 mm and the wheel perimeter.

Table 2. Baseline and damage scenarios.

	Baseline Scenarios		Damage Scenarios	
	Perfect Wheel		Polygonal wheel	Wheel flat
Load schemes	Empty			
	Half			
	Full		Full	
	UNB 1			
	UNB 2			
	UNB 3			
Irregularity profiles	4		1	
Speeds	40, 60, 80, 100, 120 km/h		60, 80, 100 km/h	
Localization	-	-	1 st wagon (First wheel right side)	3 rd wagon (Second wheel left side)
Amplitude of the defect (W)	-	-	A1: 0.25 – 0.35 mm A2: 0.65 – 0.75 mm	F1: F10–20 mm F2: F25–50 mm F3: F55–100 mm
Order of harmonics (H)	-	-	H1: H6-8 H2: H12-14 H3: H17-18	-
Total analyses	120		90	90

To define scenarios with polygonal wheels for the Laagrss wagon, a literature survey on experimental measurements in defective wheels was conducted [20,69,70]. Three experimental irregularity profiles were collected, illustrated in Figure 3. The first group (H6-8) corresponds to a study carried out by Cai et al. [69] with dominant harmonics from 6 to 8;

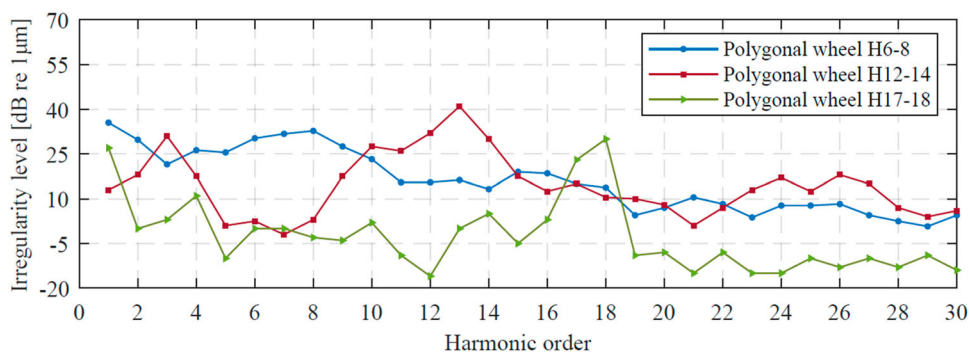


Figure 3. Wheel irregularity amplitude spectra (L_w) and harmonic order (θ) of polygonal wheels with dominant harmonic orders: (a) H6-8, (b) H12-14, (c) H17-18.

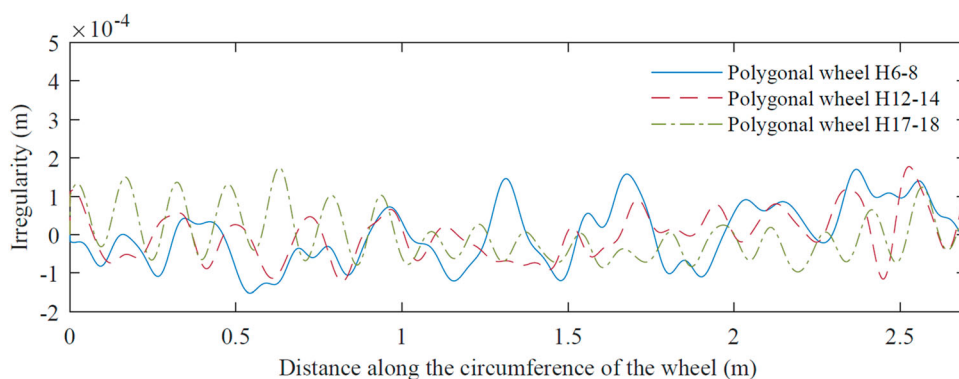


Figure 4. Examples of polygonal damage profiles for each of the dominant harmonic intervals.

the second group (H12-14) to a study authored by Tao et al. [20] with dominant harmonics from 12 to 14; and the third profile (H17-18) to a study performed by Wu et al. [70] with dominant harmonics from 17 to 18.

Based on the wheel irregularity spectra (L_w) taken from the respective measurement profiles, shown in Figure 3, and considering the contribution of the first 30 harmonics, several irregularity profiles were generated, according to Figure 4. For this, phase angles (ψ_0) are randomly assigned to sinusoidal functions in the range $0-2\pi$, which is affected by a value referring to the amplitude of the sine function of each wavelength (A_0).

One of the crucial parameters to be aware of before obtaining irregularity profiles is the wheel diameter, which, in the case of Laagrss, is approximately 0.86 m.

More details on the process of generation of polygonal wheel profiles can be found in the study by Guedes et al. [38].

For each harmonic group, five profiles with amplitudes ranging between 0.25 and 0.35 mm (A1) and other five profiles with amplitudes ranging between 0.65 and 0.75 mm (A2) were simulated.

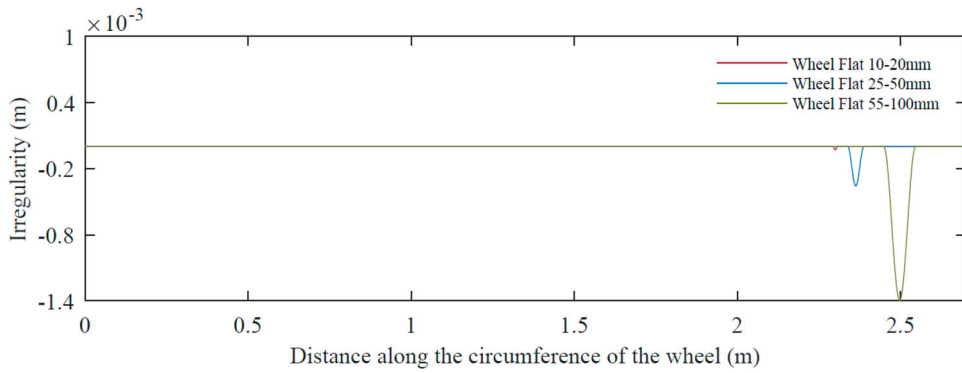


Figure 5. Examples of wheel flat profiles for each length.

2.3.3. Wheel flat

A wheel flat is a discrete defect on the wheel surface often caused when the wheelsets are locked and slide along the rails due to poorly adjusted or defective brakes. Since we are dealing with a discrete defect, the wheel profile is typically represented by a peak in the contact zone between the defect and the track.

Thus, three severity groups based on flat length intervals were defined based on the study of Mosleh et al. [37]. A less severe group, with a 10–20 mm length, an intermediate group, with a 25–50 mm length and a more severe case, with a 55–100 mm length. The irregularity profiles are generated randomly, but in this case, in function of two geometric variables, the length and the depth of the flat. More details about the process of generation of flat profiles can be found in the study performed by Mosleh et al [37]. Figure 5 shows the irregularity profiles for each severity, clearly showing the influence of flat length and depth.

As the flat is a punctual defect that occurs in the wheel from perimeter to perimeter and is characterised by a smooth area on the wheel, the irregularity profile is marked by a peak in the zone where the flat occurs.

2.4. Track dynamic responses

Figure 6 presents the acceleration records measured on the rail (A1) for the baseline and damage scenarios. All time series are filtered based on a low-pass Chebyshev type II digital filter [38,61] with a cut-off frequency of 1500 Hz. To have a more realistic track response, an artificial noise of 5% is considered in the numerical measurements. The sampling frequency for data acquisition is considered 10 kHz in order to ensure a good problem integration and allow numerical signals to have a clear definition.

Figure 6a) shows two baseline cases for two speeds (40 and 100 km/h), where it is possible to observe the influence of speed on the dynamic response. The amplifying effect of speed can influence the identification of damage, which is why a robust baseline with a wide range of scenarios is very important. At a speed of 40 km/h, lower accelerations around 0.5 m/s^2 are observed. In contrast, at a speed of 100 km/h (considerably higher), accelerations increase substantially to around 3.5 m/s^2 . This illustrates the significant dynamic effect exhibited by this vehicle emphasizing its role in amplifying the responses at

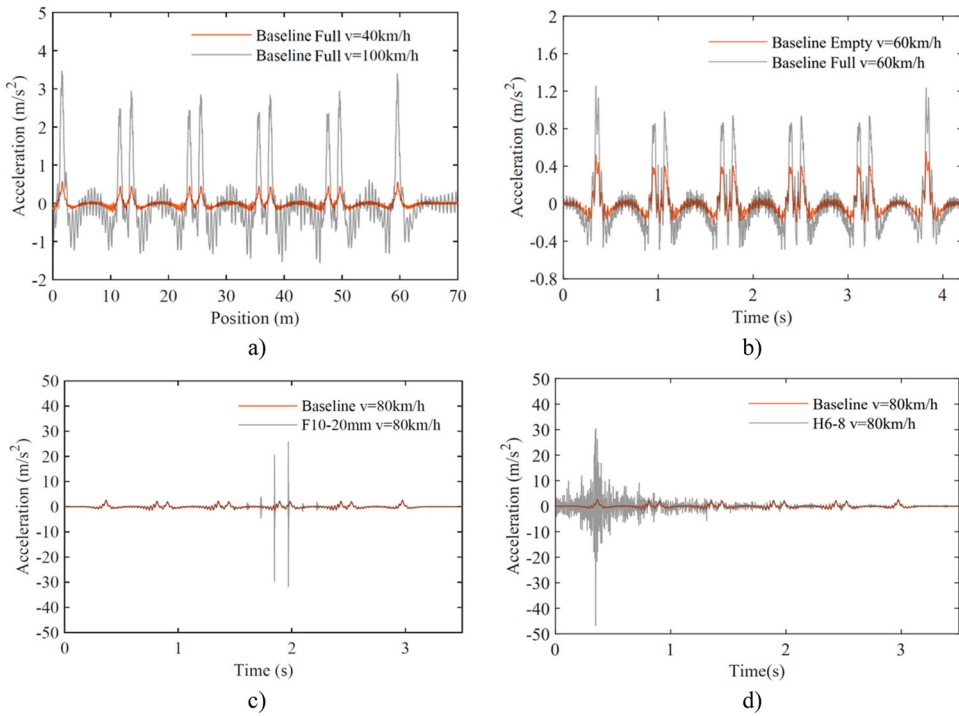


Figure 6. Acceleration time-series in sensor A1 for Laagrss vehicle: (a) influence of speed, (b) influence of load schemes, (c) wheel flat defect, (d) polygonal wheel defect.

higher speeds. In Figure 6b), the influence of the load scheme is observed in the baseline responses, considering the same speed (60 km/h) and the same type of irregularity (Irregularity 2). Lastly, Figure 6c) and d), show two acceleration responses for the two types of simulated OOR damages. Compared to a perfect wheel, there is a significant increase of the peak values, more localised in the case of flats (Figure 6c)) and more distributed in the case of polygonal wheels (Figure 6d)).

3. Methodology

3.1. Overview

The automatic methodology proposed in this study for the detection and classification of the wheel condition is based on Machine Learning (ML) and comprises several phases, summarised in the flowchart presented in Figure 7. After obtaining the acceleration measurements from the numerical simulations, an FFT algorithm is applied to the time series allowing the conversion of the signal from the time domain to the frequency domain. Then, a Stacked Sparse Autoencoder (SSAE) is used to train the input data based on a set of predefined hyperparameters of the algorithm. The set hyperparameters are encoding and decoding transfer functions, hidden layer size, maximum number of epochs, L2 weight regularisation term, sparsity regularisation term and a sparsity proportion term. Afterwards,

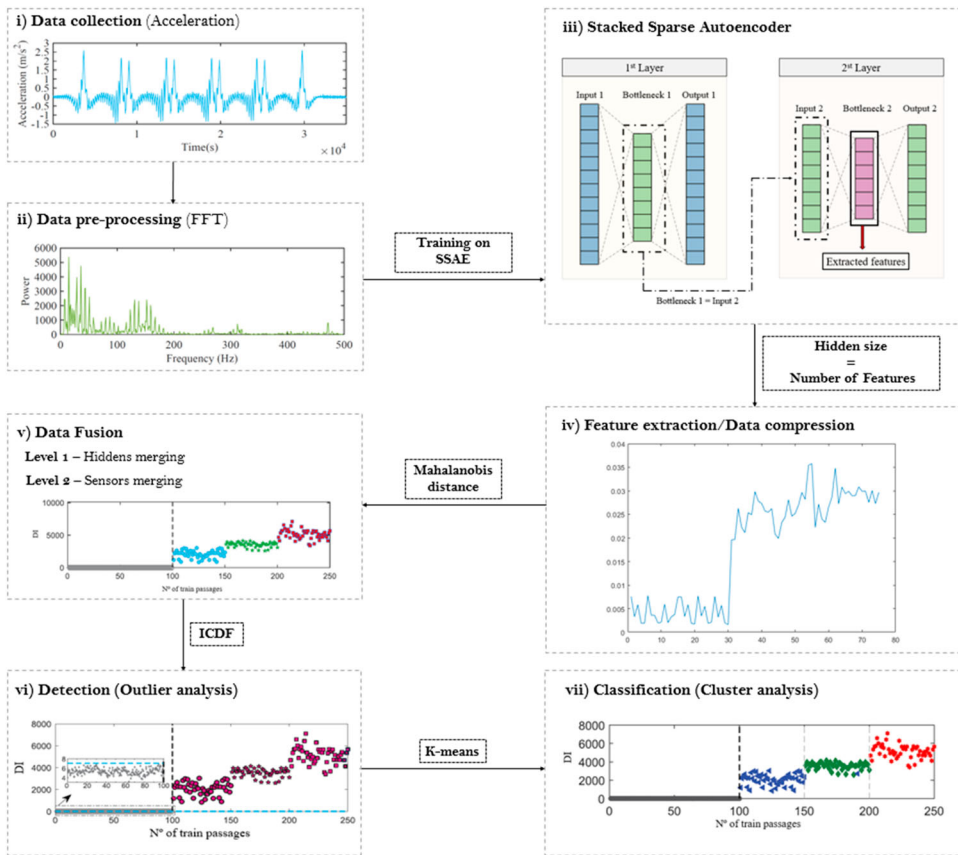


Figure 7. Flowchart with the main phases of the methodology for detection and classification of OOR wheel damages on railway vehicles.

the features are extracted in the more compressed layer of the SAE, the so-called bottleneck layer. The number of features obtained is equal to the number of hidden layers which is usually considerably lower than the data size, allowing a compression of the input data.

Furthermore, to enhance the sensitivity to damage, two levels of data fusion are applied to the bottleneck features using the Mahalanobis distance. In the first level, all features are fused into a damage index (DI) for each sensor and, in the second level, all the sensors are merged resulting in only one DI for each simulation. Finally, automatic damage detection is achieved by performing an outlier analysis, in which a statistical confidence limit (CB) is used, based on the Inverse Cumulative Gaussian Distribution Function (ICDF). For feature classification, an automatic clustering process based on the k-means technique is implemented. This algorithm allows for the unsupervised and automatic classification of data, based on the calculation of statistical distances between objects. In k-means technique given a vector of data and chosen k clusters, the key is to find k centroids that minimise the distance of each vector to its nearest centroid [72]. Since it is necessary to define the number of clusters in advance, a global Silhouette index (SIL) [33] was adopted to make an automatic process. At the end, the clusters will allow the classification of the two types of OOR damages and their respective severities.

The evaluation of the accuracy of the classification was performed through the ratio between the number of correct classifications and the total number of cases.

3.2. Data collection and pre-processing

The transformation of acceleration records using an FFT algorithm is frequently used in damage detection problems, as it allows a signal to be converted from the time domain to the frequency domain, to capture the excitation frequencies induced by the wheel defects. In addition, a data reduction is carried out at this stage, since the initial records contain a maximum of 74,000 data points, which are reduced to 7,000.

The results obtained in this first step of the methodology are plotted in Figure 8 for a crossing speed of 80 km/h. When comparing the FFT spectra for a train passage with flats, Figure 8a), with polygonal wheels, Figure 8b), there are visible differences in terms of amplitude peak values and excited frequency ranges. For flats, the amplitude peak values are higher and they induce frequencies in the range between DC and 1500 Hz, while for polygonal wheels they are restricted to the range between 50 and 300 Hz, which is in accordance with some authors reports [24, 68]. In terms of severities, for flats, the more severe group (55–100 mm) is the one with the higher frequency content, as expected, with a peak ranging essentially between frequencies DC and 500 Hz. On the other hand, the less severe group is the one inducing less frequency content and is very similar to a baseline response. For polygonal wheels, from the analysis of Figure 8b), it can be seen the frequency peaks for the H12–14 and H17–18 profiles around 100 and 150 Hz, respectively, while for the profile H6–8 a more flat spectrum is obtained without showing any significant peak. This can be explained by the irregularity level profile of this harmonic group (Figure 3) which has a less evident harmonic peak than the other two profiles.

These two graphs can show which frequencies have higher content in the original signal.

3.3. Stacked sparse autoencoder

3.3.1. Autoencoder (AE)

Autoencoders are artificial neural networks trained using an unsupervised approach, which aim first to learn coded representations of the data, compressing the input into a latent space representation, and second, through this compressed information, reconstructing the information. The basic architecture of this type of network consists of three parts [71]: the Encoder, the Latent Space and the Decoder.

The encoding process (Equation (1)) and the decoding process (Equation (2)) can be defined by:

$$H = \sigma_1(W_{ij} \cdot X + b_1) \quad (1)$$

$$\hat{X} = \sigma_2(W_{jk} \cdot H + b_2) \quad (2)$$

where W_{ij} , W_{jk} are the weights connection matrices between input layer-hidden layer and hidden layer-output layer, $X = (X_1, X_2, \dots, X_n)$ is the input data vector, $H = (H_1, \dots, H_m)$ is the low-dimensional vector and $\hat{X} = (\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n)$ is the output

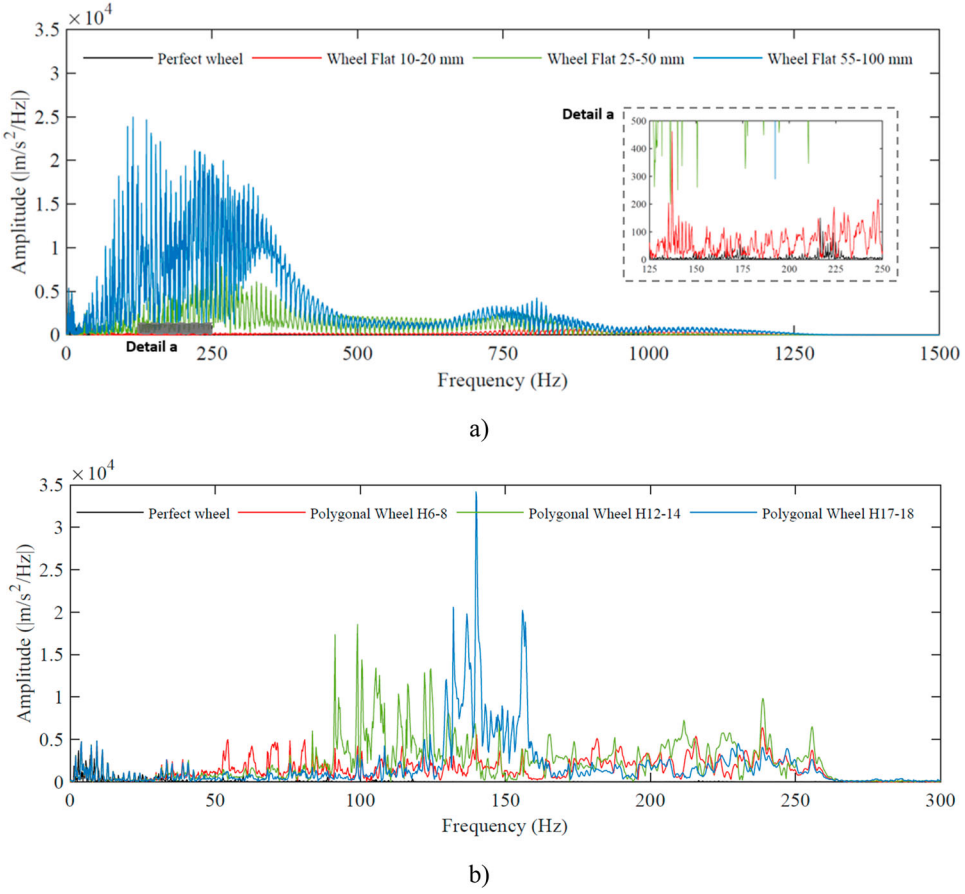


Figure 8. Frequency domain responses on position 1 for the crossing speed of 80 km/h: (a) wheel flats, (b) polygonal wheels.

data vector. $\mathbf{b}_1, \mathbf{b}_2$ are bias vectors and σ_1, σ_2 are the activation function used in each of the parts.

The optimisation of the weight matrix and bias vectors is performed through the minimisation of an error function that expresses the difference between the input X and the reconstruction of the input $\hat{X}$, as shown in Equation (3). This function uses the mean squared-error function, where N is the number of input samples.

$$E(W, \sigma) = \frac{1}{N} \sum_{j=1}^N \| \hat{X} - X \|^2 \quad (3)$$

3.3.2. Sparse autoencoder (SAE)

In the Sparse Autoencoder, to make the coding more efficient, two regularisation terms, sparse penalty term and weight regularisation term, are added to the error function, causing only a subset of neurons to be activated. The average activation of neurons in the hidden

layer $\hat{\rho}_i$ is given by [42]:

$$\hat{\rho}_i = \frac{1}{N} \sum_{j=1}^N z_i(x(j)) = \frac{1}{N} \sum_{j=1}^N \sigma(w_i^T \cdot x_j + b_i) \quad (4)$$

As the average activation $\hat{\rho}$ is intended to be close to a constant ρ (sparsity proportion) and it is usually a small positive number near 0, the Kullback-Leibler (KL) divergence is used [42].

$$\Omega_{sparse} = KL(\rho \parallel \hat{\rho}_i) = \rho \cdot \log\left(\frac{\rho}{\hat{\rho}_i}\right) + (1 - \rho) \cdot \log\left(\frac{1 - \rho}{1 - \hat{\rho}_i}\right) \quad (5)$$

In addition to this, a weight regularisation term, called L2 regularisation, is also added to avoid overfitting the network [42].

$$\Omega_{weights} = \frac{1}{2} \sum_{l=1}^L \sum_{j=1}^{n_l} \sum_{i=1}^{k_l} (w_{ji}^l)^2 \quad (6)$$

where L is the number of hidden layers, n_l the output size of layer l , k_l the input size of layer l .

Thus, the cost function of a Sparse Autoencoder is given by:

$$E_{sparse}(W, \sigma) = E(W, \sigma) + \mu \cdot \Omega_{sparse} + \beta \cdot \Omega_{weights} \quad (7)$$

where μ is the coefficient for the L2 regularisation term and β the coefficient for the sparsity regularisation term.

3.3.3. Training of the sparse autoencoder

After data pre-processing, the acceleration responses in the frequency domain, are then used as input for training a stack of several sparse autoencoders.

In this work, two layers of SAE's were used to form the network, as shown in Figure 9. The architecture of an SSAE involves training each autoencoder individually in successive layers, with the outputs of the previous autoencoder serving as input for the next one. After incorporating various inputs into the SSAE training, they are integrated into the so-called 'Input 1'. As training progresses, the data is compressed, and the main characteristics of the signals are obtained and retained in 'Bottleneck 1'. Subsequently, these characteristics are used as input for training the second sparse autoencoder, which, after further compression, enables the extraction of the final features. Depending on the type of data under analysis, the compression level differs, so multiple tests should be conducted to optimise these layer sizes. This optimisation was achieved by trial and test, changing both the size of the first and second layers.

It should be noted that this network architecture, where the dimensions of the layers gradually decrease along the network until reaching the bottleneck layer [73], has proven to be highly robust for the type of data used. It allows for effective compression of results without compromising efficiency in data classification.

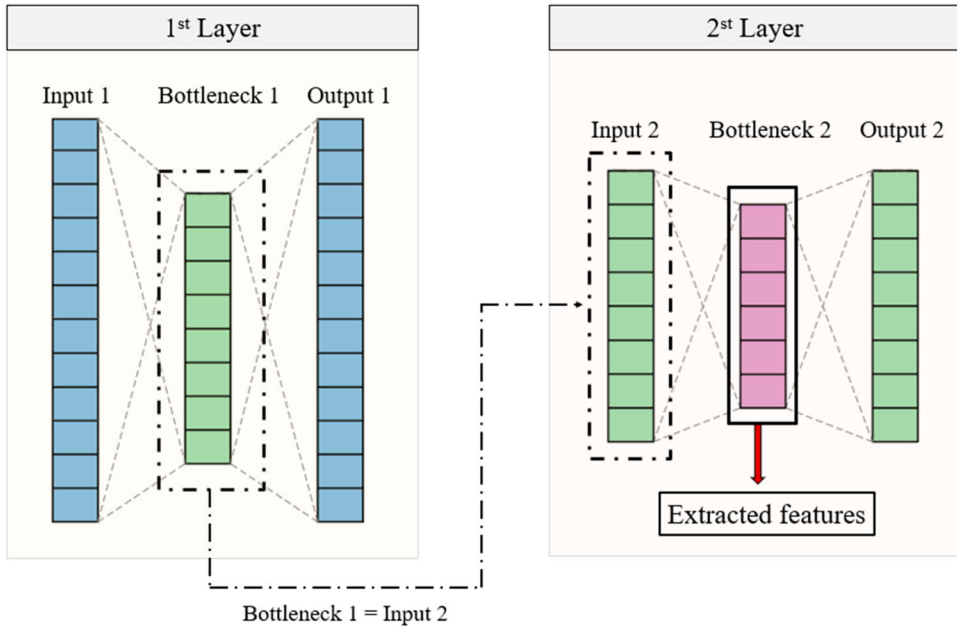


Figure 9. Schematic of the structure of the SSAE neural network used in this study.

3.4. Data fusion

After extracting the features from the bottleneck layer of the SSAE, the difference between the healthy wheels scenarios and the damage wheels scenarios is still not sufficient. Hence a data fusion is performed. With this fusion, it is possible to increase the sensitivity of the features, and as a result, to obtain a damage index (DI) for each simulation. Thus, the Mahalanobis distance (MD) is used to reduce the volume of extracted data and preserve the most relevant information [37,38,72]. Briefly, MD is a metric that measures the distance between two points in a feature space consisting of two or more variables, being independent of the scale of the features, thus allowing an increase in the sensitivity of the features to cases with damage. The MD, denoted here as DI, is calculated for each simulation using the following expression:

$$DI = \sqrt{(x_i - \bar{x}) \cdot C^{-1} \cdot (x_i - \bar{x})^T} \quad (8)$$

in which, the inverse covariance matrix of the baseline simulation is defined by C^{-1} , and $\bar{x}$ is a mean vector of the features from the baseline simulation. The test vector of f features representing the potential damage is defined by x_i . The MD is calculated for each sensor and for each simulation, obtaining a matrix of n distances per k number of sensors.

3.5. Feature discrimination

Finally, the discrimination of the features consists in distinguishing the features as corresponding to a healthy or damaged case. At this stage of the methodology, an analysis of abnormal values is carried out, which will allow the condition of the train wheels to be

automatically monitored and these values to be classified into different classes. This phase can be divided into two main ones:

3.5.1. Outlier analysis

The outlier analysis consists of grouping, by probabilistic distribution processes, the data of normal conditions and verifying if the new extracted data have similar characteristics to the previously created group. For this, a statistical confidence limit (CB) is used, based on the inverse cumulative Gaussian distribution function (ICDF), considering the mean value, $\bar{\mu}$, and standard deviation, σ , of the baseline feature vector.

$$CB = invF_x(1 - \alpha) \quad (9)$$

$$F(x \setminus \bar{\mu}, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\alpha}^x e^{-\frac{1}{2} \left(\frac{x - \bar{\mu}}{\sigma} \right)^2} dy, \text{ for } x \in \mathbb{R} \quad (10)$$

Whenever a value of a DI indicator is equal to or greater than CB, it is considered an outlier (the null hypothesis is rejected), thus representing a case of damage.

3.5.2. Cluster analysis

For feature classification, a clustering process is used to divide the dataset into distinct clusters that should be as compact and separated as possible [33]. For clustering, the k-means technique uses the city-block distance and aims to reorganise the data into k clusters based on the distances of their vectors. Recent studies showed that the k-means algorithm has shown promising results [37, 72]. The main limitation of this technique is that it requires knowledge of the number of clusters in advance. For this purpose, the global silhouette index [33] is used, making this an automatic process.

4. Results

This section presents the main results obtained through the application of the methodology proposed in section 3 to the calculated track responses. The methodology is applied using damaged wheel passages for the three simulated train speeds (60, 80 and 100 km/h), simultaneously.

4.1. Training hyperparameters in sparse autoencoder

The application of a Sparse Autoencoder (SAE) introduces constraints by using sparsity hyperparameters (presented in section 3). By adjusting these hyperparameters, it is possible to make the algorithm more efficient and improve the generalisation capabilities, as only the most important features are retained. The optimisation of these hyperparameters is performed through an iterative process that requires several tests and depends on the type of data under analysis. Based on the principles of the SAE already outlined in other works employing this type of neural network [47,73], in this study, the hyperparameters were exclusively calibrated using the training of baseline data.

A total of 90 baseline cases are used for training, while the remaining 30 are saved for testing, representing around 70% for training and 30% for testing, following the recommendations of other study [73]. Table 3 shows the optimised hyperparameters obtained for the two types of damage and considering the three speeds.

Table 3. Hyperparameters used to train the SSAE.

Training hyperparameters in Sparse Autoencoder	Type of damage	Polygonal wheel	Wheel flat
Encode transfer function	logsig	logsig	logsig
Decode transfer function	logsig	logsig	logsig
Size of hidden layer (Bottleneck 1/Bottleneck 2)	3500/2000	1500/500	1000/500
Max. number of epochs (Bottleneck 1/Bottleneck 2)	5/10	50/300	50/300
L2 weights regularisation (Bottleneck 1/Bottleneck 2) (λ)	$1 \times 10^{-4}/1 \times 10^{-6}$	$1 \times 10^{-4}/1 \times 10^{-6}$	$1 \times 10^{-4}/1 \times 10^{-6}$
Sparsity Regularisation (Bottleneck 1/Bottleneck 2) (β)	0.1/1	0.1/8	0.1/8
Sparsity Proportion (Bottleneck 1/Bottleneck 2) (ρ)	$0.01/1 \times 10^{-3}$	$0.01/1 \times 10^{-3}$	$0.01/1 \times 10^{-3}$

Additionally, a Scaled Conjugate Gradient algorithm (SCG) was considered with a stopping criterion when the model reached the value of 1×10^{-6} in the loss function (E_{sparse}) or the maximum number of epochs.

In the context of this study, the non-linear transfer functions (logsig) proved to be more effective than the linear ones, being the ones used in all the studied cases. After analysing Table 3, there is a similarity in the hyperparameters when comparing the two types of OOR damage. It is observed that the data compression level is the same for both cases, around 93%, as 500 features are extracted in Bottleneck 2 from the initial input data with 7000 points.

Concerning the type of damage situation, the size of the ‘Bottleneck 1’ and ‘Bottleneck 2’ was increased to bring them even closer to the initial amount of input data. This can be explained by the fact that two different types of damage are being evaluated in the same analysis, making it challenging for the autoencoder to define relationships between the data, hence requiring a larger number. On the contrary, the maximum number of epochs was decreased since after several tests was found that the influence of the epochs was minimal and have a negative influence on the calculation time.

It is also important to emphasise that of the various hyperparameters presented, the ones that had the greatest impact on the results were the size of the hidden layer and the maximum number of epochs.

Once the hyperparameters have been defined, all the steps of the methodology are applied to the new dataset. For simplicity, in this section, only the last phase of the methodology will be presented, namely an outlier analysis for damage detection and cluster analysis for damage classification. For each sensor, a matrix $n \times m$ is created, where n is the number of test passages (baseline scenarios + damage scenarios), and m represents the number of extracted features. So, the size of the matrices is 210×2000 in case of the type of damage, 75×500 for each amplitude in polygonal wheel, and 120×500 for wheel flat.

4.2. Damage detection

Despite the visual distinction between healthy and damaged wheels based on the data processing presented in subsection 3.2, unsupervised learning must use strategies capable of automatically detecting damage. To this end, an outlier analysis is employed, which consists of defining a confidence threshold using a significance level of 1% [36], represented

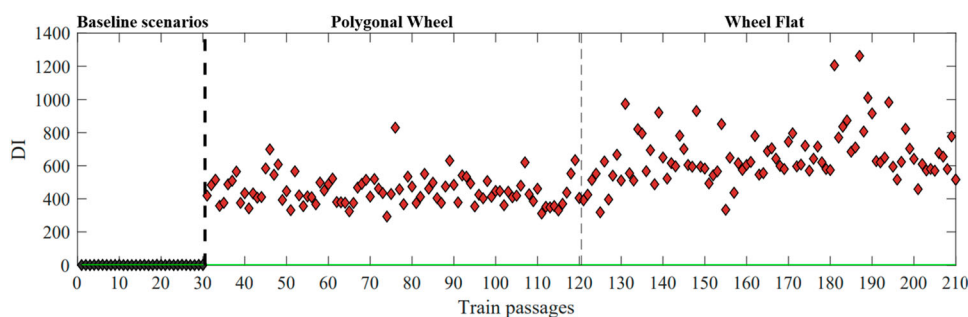


Figure 10. Automatic OOR damage detection based on outlier analysis.

by the green line in Figure 10. This approach makes it possible to successfully detect all damage cases without any false positives since the damage index values are all above the threshold line.

4.3. Classification of type of damage OOR

To distinguish the two types of OOR damage, a cluster analysis is performed. Figure 11 shows a clear distinction between the damage, even when using a dataset with three different train speeds (60, 80 and 100 km/h).. In the case of Figure 11a), using all four accelerometers, it is possible to efficiently separate the set into two large clusters, despite 30 misclassifications (16 in the polygonisation and 14 in the flats). Even so, the accuracy percentage is around 85%, which is a very good result. To analyse the sensitivity to the number of sensors, only two sensors were used, one on each side (Figure 11b)). The results improved, with a decrease in the number of misclassifications from 30 to 21 (13 for polygonisation and 8 for flats). One explanation could be the disparity in sensor sensitivity to the damage, where some sensors are more sensitive than others due to their spatial positioning in relation to the damage. By merging the data from the less sensitive sensors with the more sensitive ones, this could have a negative impact on the results. Thus, the percentage of accuracy in this case is around 90%.

The purpose behind presenting these two plots was to demonstrate that, whether using 4 sensors or 2 sensors, the classification remains remarkably robust, showing low misclassifications in both scenarios.

4.4. Classification of severity

After identifying the type of OOR damage, a cluster analysis is performed once again to classify the severity of each damage separately. Regarding polygonal wheels, the classification involves both amplitude and harmonic orders, leading to the clusters depicted in Figure 12. Initially, amplitude classification is essential, as depicted in Figure 12a). Here, the obtained clusters reasonably reflect the two amplitude groups, with all misclassifications associated with the H6-8 harmonic cases (17 misclassifications). Once the amplitude is identified, harmonic classification must be carried out separately for each of the amplitudes. In Figure 12b), related to amplitude A1, it is observed that the obtained clusters

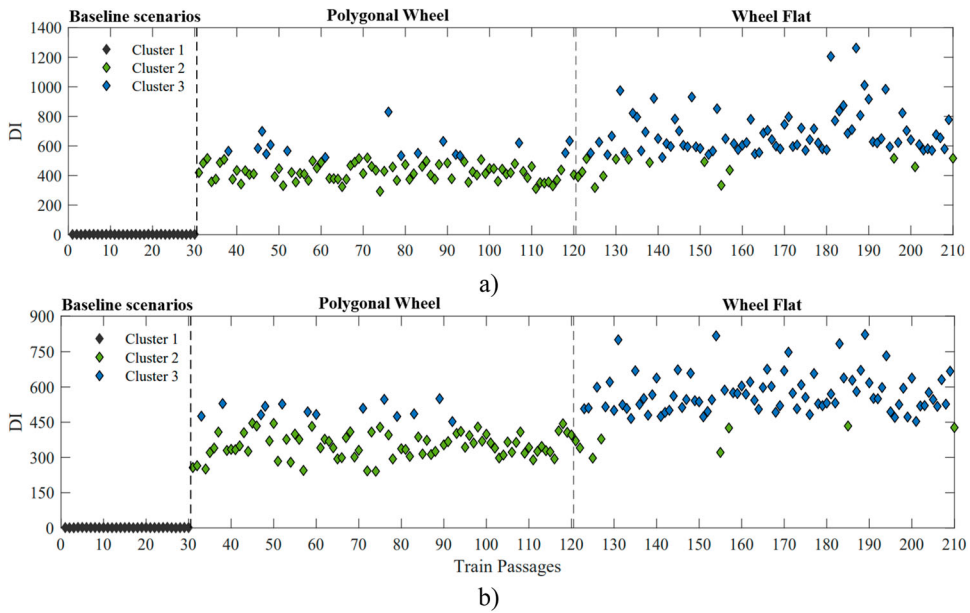


Figure 11. Automatic OOR damage classification based on cluster analysis using: a) 4 sensors; b) 2 sensors.

are nearly perfect, with only 3 misclassifications. In Figure 12c), there are only 3 misclassifications, totalling 8 misclassifications in the classification of the number of dominant harmonics.

The classification of harmonic order is only possible with the separation of amplitudes, otherwise, the algorithm becomes less accurate.

Figure 13 presents the clusters obtained for the wheel flat cases. The separation between the less severe group and baseline cases and intermediate (cluster 3) and severe cases (cluster 4) is reduced, resulting in a total of 19 misclassifications. Nevertheless, this result shows that the methodology efficiently distinguishes the three severity groups with an accuracy rate of about 84%. One possible reason for this significant number of misclassifications could be directly associated with the number of impacts captured by the sensor.

5. Conclusions

In order to contribute to the safety of railway transportation and service quality, an automatic methodology based on machine learning has been proposed. This methodology can accurately detect and classify two types of OOR wheel damage in railway vehicles allowing to mitigate problems at both track and vehicle level, such as derailments or material degradation. It includes a new technique for feature extraction based on a dedicated neural network, the Sparse Autoencoder. This non-linear technique is an upgrade concerning the commonly used linear techniques for feature extraction since it allows a better normalisation of the operational effects in the analysed data, as well as the reduction of large amounts of data without relying on other techniques, while keeping the most valuable information. These improvements are achieved through the use of different hyperparameters in

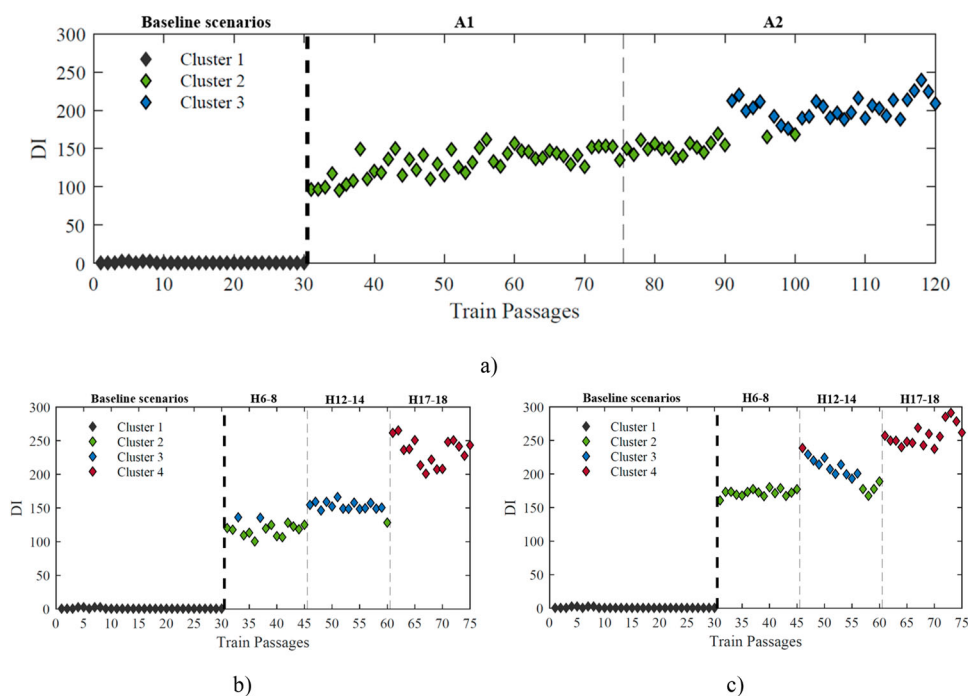


Figure 12. Automatic classification of polygonal wheels based on cluster analysis: a) amplitude A1-A2; b) harmonic orders for A1; c) harmonic orders for A2.

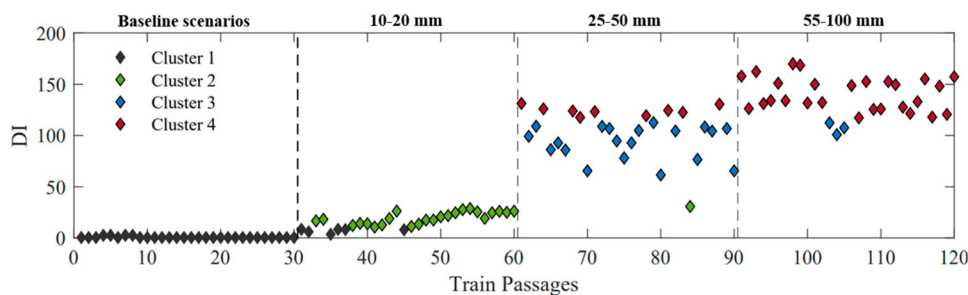


Figure 13. Automatic classification of wheel flats based on cluster analysis.

the calculations, such as non-linear activation functions and sparsity terms. Consequently, it becomes possible to identify relationships in the data independently on the speed of circulation. Particularly in the railway engineering field, the inclusion of this technique showed a good potential to deal with variable train speeds, essentially due to the amplifying effect on the measured track dynamic responses which could induce errors in damage identification.

The application of the methodology to the data set proved to be very successful in detecting and classifying the two types of OOR damages. Damage detection was perfect, without any false detection. In terms of classifying the type of damage, the results are very good, with a success rate of around 85%, but if only two accelerometers are considered the success

rate increases to around 90%. These results are encouraging since the range of speeds considered in the analysis, between 60 and 100 km/h, is the most representative of in-service freight trains. Also, the results show that including the Sparse Autoencoder has a positive effect in the normalization of the train speed in the damage data set.

Regarding the severity classification for each type of damage, the methodology proved equally successful in clustering the different damage severities. For polygonal wheels, the methodology efficiently distinguishes the severity groups in terms of defect amplitude, failing only for the H6-8 harmonic cases, and in terms of harmonic orders for each amplitude, with almost perfect results. For wheel flats, the methodology efficiently distinguishes the three severity groups, though with some misclassifications between the less severe group and baseline cases and between intermediate and severe cases. A success rate of around 86% is achieved in the classification of polygonal wheel severities and 84% in the classification of wheel flat severities. Even so, these high success rates of over 80% show the robustness of this methodology in damage classification regardless of the train speed.

Future works include experimental tests to validate the proposed methodology based on on-site measurements. Additionally, this methodology could be carried out for other types of railway vehicles, for example, passenger vehicles, since the dynamic responses are different from freight vehicles. In that case, exploring other recent types of autoencoders, such as the ensemble autoencoder could also be an option [74].

To increase the complexity of the problem, the objective is to explore the concept of multiple damages (same or different OOR types). For this, it will be necessary to incorporate a new stage in the methodology, the damage location stage, which will make it possible to distinguish isolated damage from multiple damage, as well as their respective location. Thus, damage will be identified using cut signals, which will then be used to classify the data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is a result of project WAY4SafeRail-WAYside monitoring system FOR SAFE RAIL transportation, with reference NORTE-01-0247-FEDER-069595, co-funded by the European Regional Development Fund (ERDF), through the North Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement. This work also received Portuguese National Funds through FCT (Portuguese Foundation for Science and Technology) under project UIDB/00760/2020 and from the Base Funding – UIDB/04708/2020 of the CONSTRUCT – Instituto de I&D em Estruturas e Construções – funded by Portuguese national funds through the FCT/MCTES (PIDDAC).

ORCID

D. Ribeiro  <http://orcid.org/0000-0001-8624-9904>

C. Vale  <http://orcid.org/0000-0003-2470-9834>

A. Mosleh  <http://orcid.org/0000-0001-6751-0388>

P. Montenegro  <http://orcid.org/0000-0001-5699-4428>

References

- [1] Shaikh K, Hussain I, Chowdhry BS. Wheel defect detection using a hybrid deep learning approach. *Sensors*. 2023;23(14):6248. doi:10.3390/s23146248
- [2] Lagnebäck R. Evaluation of wayside condition monitoring technologies for condition-based maintenance of railway vehicles. 2007.
- [3] Shaikh MZ, et al. State-of-the-art wayside condition monitoring systems for railway wheels: a comprehensive review. *IEEE Access*. 2023;11:13257–13279. doi:10.1109/ACCESS.2023.3240167
- [4] Bian J, Gu Y, Murray MH. A dynamic wheel–rail impact analysis of railway track under wheel flat by finite element analysis. *Veh Syst Dyn*. 2013;51(6):784–797. doi:10.1080/00423114.2013.774031
- [5] Zhou C, Gao L, Xiao H, et al. Railway wheel flat recognition and precise positioning method based on multisensor arrays. *Appl Sci*. 2020;10(4):1297. doi:10.3390/app10041297
- [6] Xu J, et al. Effect of wheel flat on dynamic wheel-rail impact in railway turnouts. *Veh Syst Dyn*. 2021;60:1–20.
- [7] Yang J, et al. Influence of wheel flat on railway vehicle helical gear system under Traction/Braking conditions. *Eng Fail Anal*. 2022;134:106022. doi:10.1016/j.engfailanal.2021.106022
- [8] Wu X, et al. Influence of a flexible wheelset on the dynamic responses of a high-speed railway car due to a wheel flat. *Proc Inst Mech Eng F: J Rail Rapid Transit*. 2018;232(4):1033–1048.
- [9] Chen YZ, et al. The influence of wheel flats formed from different braking conditions on rolling contact fatigue of railway wheel. *Eng Fail Anal*. 2018;93:183–199. doi:10.1016/j.engfailanal.2018.07.006
- [10] Balci E, Bezgin N. *Effects of the vehicle speed, wheel diameter, and length of the wheel flat on the dynamic impact forces caused by wheel flats*. 2022.
- [11] Vale C. Wheel flats in the dynamic behavior of ballasted and slab railway tracks. *Appl Sci*. 2021;11:7127. doi:10.3390/app11157127
- [12] Wu TX, Thompson DJ. A Hybrid model for the noise generation due to railway wheel flats. *J Sound Vib*. 2002;251(1):115–139. doi:10.1006/jsvi.2001.3980
- [13] Baasch B, et al. Train wheel condition monitoring via cepstral analysis of axle Box accelerations. *Appl Sci*. 2021;11(4):1432. doi:10.3390/app11041432
- [14] Cui D, et al. Effect of the turning characteristics of underfloor wheel lathes on the evolution of wheel polygonisation. *Proc Inst Mech Eng F: J Rail Rapid Transit*. 2019;233:479–488. doi:10.1177/0954409718795760
- [15] Meinke P, Meinke S. Polygonalization of wheel treads caused by static and dynamic imbalances. *J Sound Vib*. 1999;227(5):979–986. doi:10.1006/jsvi.1999.2590
- [16] Morys B. Enlargement of out-of-round wheel profiles on high speed trains. *J Sound Vib*. 1999;227(5):965–978. doi:10.1006/jsvi.1999.2055
- [17] Wu Y, et al. Experimental analysis of the mechanism of high-order polygonal wear of wheels of a high-speed train. *J Zhejiang Univ-Sci A*. 2017;18:579–592. doi:10.1631/jzus.A1600741
- [18] Nielsen J, Johansson A, Vernersson T. Train-track interaction and mechanisms of irregular wear on wheel and rail surfaces. *Veh Syst Dyn*. 2003;40:3–54. doi:10.1076/vesd.40.1.3.15874
- [19] Tao G, et al. Polygonisation of railway wheels: a critical review. *Railw Eng Sci*. 2020;28(4):317–345. doi:10.1007/s40534-020-00222-x
- [20] Tao G, et al. An investigation into the mechanism of high-order polygonal wear of metro train wheels and its mitigation measures. *Veh Syst Dyn*. 2020;59:1–16.
- [21] Fröhling R, Spangenberg U, Reitmann E. Root cause analysis of locomotive wheel tread polygonisation. *Wear*. 2019;432–433:102911. doi:10.1016/j.wear.2019.05.026
- [22] Zehetbauer F, Edelmann J, Plöchl M. Influences of tram characteristics on wheel polygonal wear evolution. *Eng Fail Anal*. 2023;157:107528.
- [23] Wu X, Chi M, Wu P. Influence of polygonal wear of railway wheels on the wheel set axle stress. *Veh Syst Dyn*. 2015;53(11):1535–1554. doi:10.1080/00423114.2015.1063674

- [24] Mishra S, Sharan P, Saara K. Real time implementation of fiber Bragg grating sensor in monitoring flat wheel detection for railways. *Eng Fail Anal.* **2022**;138:106376. doi:[10.1016/j.engfailanal.2022.106376](https://doi.org/10.1016/j.engfailanal.2022.106376)
- [25] Lourenço A, et al. Adaptive time series representation for out-of-round railway wheels fault diagnosis in wayside monitoring. *Eng Fail Anal.* **2023**;152:107433. doi:[10.1016/j.engfailanal.2023.107433](https://doi.org/10.1016/j.engfailanal.2023.107433)
- [26] Sun Q, et al. An on-board detection framework for polygon wear of railway wheel based on vibration acceleration of axle-box. *Mech Syst Signal Process.* **2021**;153:107540. doi:[10.1016/j.ymssp.2020.107540](https://doi.org/10.1016/j.ymssp.2020.107540)
- [27] Ye Y, et al. Multislice Time-Frequency image Entropy as a feature for railway wheel fault diagnosis. *Measurement (Mahwah N J).* **2023a**;216:112862. doi:[10.1016/j.measurement.2023.112862](https://doi.org/10.1016/j.measurement.2023.112862)
- [28] Ye Y, et al. Shock detection of rotating machinery based on activated time-domain images and deep learning: an application to railway wheel flat detection. *Mech Syst Signal Process.* **2023b**;186:109856. doi:[10.1016/j.ymssp.2022.109856](https://doi.org/10.1016/j.ymssp.2022.109856)
- [29] Krummenacher G, et al. Wheel defect detection With machine learning. *IEEE Trans Intell Transp Syst.* **2017**;19:1–12.
- [30] Worden K, Dulieu-Barton JM. An overview of intelligent fault detection in systems and structures. *Struct Health Monit.* **2004**;3(1):85–98. doi:[10.1177/1475921704041866](https://doi.org/10.1177/1475921704041866)
- [31] Mohammadi M, et al. An unsupervised learning approach for wayside train wheel flat detection. *Sensors.* **2023**;23(4):1910. doi:[10.3390/s23041910](https://doi.org/10.3390/s23041910)
- [32] Mosleh A, et al. Early wheel flat detection: an automatic data-driven wavelet-based approach for railways. *Veh Syst Dyn.* **2023**;61(6):1644–1673. doi:[10.1080/00423114.2022.2103436](https://doi.org/10.1080/00423114.2022.2103436)
- [33] Meixedo A, et al. *Real-Time unsupervised detection of early damage in railway bridges using traffic-induced responses*. Vol. 21. Cham: Springer; **2021**. p. 117–142. doi:[10.1007/978-3-030-81716-9_6](https://doi.org/10.1007/978-3-030-81716-9_6)
- [34] Xu M, et al. Structural damage detection by integrating robust PCA and classical PCA for handling environmental variations and imperfect measurement data. *Adv Struct Eng.* **2022**;25(8):1815–1828. doi:[10.1177/13694332221079090](https://doi.org/10.1177/13694332221079090)
- [35] Bisheh HB, Amiri GG. Damage detection of bridges based on spectral sub-band features and hybrid modeling of PCA and KPCA methods. *Struct Monit Maint.* **2022**;9:179–200.
- [36] Meixedo A, et al. Damage detection in railway bridges using traffic-induced dynamic responses. *Eng Struct.* **2021**;238:112189. doi:[10.1016/j.engstruct.2021.112189](https://doi.org/10.1016/j.engstruct.2021.112189)
- [37] Mosleh A, et al. Automatic clustering-based approach for train wheels condition monitoring. *Int J Rail Transp.* **2022**;11(5):1–26. doi:[10.1080/23248378.2022.2096132](https://doi.org/10.1080/23248378.2022.2096132)
- [38] Guedes A, et al. Detection of wheel polygonization based on wayside monitoring and artificial intelligence. *Sensors.* **2023**;23(4):2188. doi:[10.3390/s23042188](https://doi.org/10.3390/s23042188)
- [39] Javed K, et al. Enabling health monitoring approach based on vibration data for accurate prognostics. *IEEE Trans Ind Electron.* **2015**;62(1):647–656. doi:[10.1109/TIE.2014.2327917](https://doi.org/10.1109/TIE.2014.2327917)
- [40] Cavadas F, Smith IFC, Figueiras J. Damage detection using data-driven methods applied to moving-load responses. *Mech Syst Signal Process.* **2013**;39(1):409–425. doi:[10.1016/j.ymssp.2013.02.019](https://doi.org/10.1016/j.ymssp.2013.02.019)
- [41] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys.* **1943**;5(4):115–133. doi:[10.1007/BF02478259](https://doi.org/10.1007/BF02478259)
- [42] Olshausen BA, Field DJ. Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Res* **1997**;37(23):3311–3325. doi:[10.1016/S0042-6989\(97\)00169-7](https://doi.org/10.1016/S0042-6989(97)00169-7)
- [43] Ng A. Sparse autoencoder. *CS294A Lect Notes.* **2011**;72:1–19.
- [44] Zhang C, et al. Deep sparse autoencoder for feature extraction and diagnosis of locomotive adhesion status. *J Control Sci Eng.* **2018**. 2018: 8676387.
- [45] Amaral R, et al. Structural novelty detection based on sparse autoencoders and control charts. *Struct Eng Mech.* **2022**;81:647–664.
- [46] Amaral R, et al. Numerical and experimental evaluation of structural changes using sparse auto-encoders and SVM applied to dynamic responses. *Appl Sci.* **2021**;11:6127. doi:[10.3390/app11136127](https://doi.org/10.3390/app11136127)

- [47] Pathirage C, et al. Development and application of a deep learning-based sparse autoencoder framework for structural damage identification. *Struct Health Monit.* 2019;18(1):103–122. doi:[10.1177/1475921718800363](https://doi.org/10.1177/1475921718800363)
- [48] Nick W, et al. A study of machine learning techniques for detecting and classifying structural damage. *Int J Mach Learn Comput.* 2015;5:313–318. doi:[10.7763/IJMLC.2015.V5.526](https://doi.org/10.7763/IJMLC.2015.V5.526)
- [49] Chen Z, et al. Acoustic emission analysis of crack type identification of corroded concrete columns under eccentric loading: a comparative analysis of RA-AF method and Gaussian mixture model. *Case Stud Constr Mater.* 2023;18:e02021.
- [50] Fiorini L, et al. Unsupervised emotional state classification through physiological parameters for social robotics applications. *Knowl Based Syst.* 2020;190:105217. doi:[10.1016/j.knosys.2019.105217](https://doi.org/10.1016/j.knosys.2019.105217)
- [51] Meixedo A, et al. Online unsupervised detection of structural changes using train-induced dynamic responses. *Mech Syst Signal Process.* 2022;165:108268. doi:[10.1016/j.ymssp.2021.108268](https://doi.org/10.1016/j.ymssp.2021.108268)
- [52] Osman A, et al. A Naïve-Bayes classifier for damage detection in engineering materials. *Mater Des.* 2007;28:2379–2386. doi:[10.1016/j.matdes.2006.07.018](https://doi.org/10.1016/j.matdes.2006.07.018)
- [53] Vitola J, et al. Distributed piezoelectric sensor system for damage identification in structures subjected to temperature changes. *Sensors.* 2017;17(6):1252. doi:[10.3390/s17061252](https://doi.org/10.3390/s17061252)
- [54] Bisheh HB, Amiri GG. Structural damage detection based on variational mode decomposition and kernel PCA-based support vector machine. *Eng Struct.* 2023;278:115565. doi:[10.1016/j.engstruct.2022.115565](https://doi.org/10.1016/j.engstruct.2022.115565)
- [55] Shafique R, et al. A novel approach to railway track faults detection using acoustic analysis. *Sensors.* 2021;21(18):6221. doi:[10.3390/s21186221](https://doi.org/10.3390/s21186221)
- [56] Mosleh A, Costa PA, Calçada R. A new strategy to estimate static loads for the dynamic weighing in motion of railway vehicles. *Proc Inst Mech Eng F: J Rail Rapid Transit.* 2020;234(2):183–200. doi:[10.1177/0954409719838115](https://doi.org/10.1177/0954409719838115)
- [57] Bragança C, et al. Calibration and validation of a freight wagon dynamic model in operating conditions based on limited experimental data. *Veh Syst Dyn.* 2021;60(9):3024–3050. <https://doi.org/10.1080/00423114.2021.1933091>.
- [58] Neves S, Azevedo A, Calçada R. A direct method for analyzing the vertical vehicle–structure interaction. *Eng Struct.* 2012;34:414–420. doi:[10.1016/j.engstruct.2011.10.010](https://doi.org/10.1016/j.engstruct.2011.10.010)
- [59] Montenegro PA, et al. Wheel–rail contact formulation for analyzing the lateral train–structure dynamic interaction. *Comput Struct.* 2015;152:200–214. doi:[10.1016/j.compstruc.2015.01.004](https://doi.org/10.1016/j.compstruc.2015.01.004)
- [60] Montenegro PA, Calçada R. Wheel–rail contact model for railway vehicle–structure interaction applications: development and validation. *Rail Eng Sci.* 2023;31(3):181–206. doi:[10.1007/s40534-023-00306-4](https://doi.org/10.1007/s40534-023-00306-4)
- [61] Silva R, et al. Early identification of unbalanced freight traffic loads based on wayside monitoring and artificial intelligence. *Sensors.* 2023;23(3):1544. doi:[10.3390/s23031544](https://doi.org/10.3390/s23031544)
- [62] ANSYS®, *Academic Research.* 2018.
- [63] MATLAB®, *Version 2018.a.* Available from: <https://www.mathworks.com>.
- [64] Kalker JJ. Mathematics, and informatics, *book of tables for the hertzian creep-force Law.* Faculty of Technical Mathematics and Informatics, Delft University of Technology; 1996.
- [65] EN 13848-2. (2006). *Railway applications—track—track geometry quality—Part 2: measuring systems—track recording vehicles.*
- [66] Mosleh A, et al. Railway vehicle wheel flat detection with multiple records using spectral kurtosis analysis. *Appl Sci.* 2021;11(9):4002. doi:[10.3390/app11094002](https://doi.org/10.3390/app11094002)
- [67] Iwnicki S, Nielsen JCO, Tao G. Out-of-round railway wheels and polygonisation. *Veh Syst Dyn.* 2023;61(7):1787–1830. doi:[10.1080/00423114.2023.2194544](https://doi.org/10.1080/00423114.2023.2194544)
- [68] Peng B. Mechanisms of railway wheel polygonization. University of Huddersfield; 2020. <https://eprints.hud.ac.uk/id/eprint/35359/>.
- [69] Cai W, et al. Experimental and numerical investigation into formation of metro wheel polygonalization. *Shock Vib.* 2019;2019(1):1–18. doi:[10.1155/2019/1538273](https://doi.org/10.1155/2019/1538273)
- [70] Wu X, et al. A study of formation of high order wheel polygonalization. *Wear.* 2019;424–425:1–14. doi:[10.1016/j.wear.2019.01.099](https://doi.org/10.1016/j.wear.2019.01.099)

- [71] Yan B, Han G. Effective feature extraction via stacked sparse autoencoder to improve intrusion detection system. *IEEE Access*. 2018;6:41238–41248. doi:[10.1109/ACCESS.2018.2858277](https://doi.org/10.1109/ACCESS.2018.2858277)
- [72] Meixedo A, et al. Structural health monitoring strategy for damage detection in railway bridges using traffic induced dynamic responses. *Rail Infrastructure Resilience*; 2022. p. 389–408. doi:[10.1016/B978-0-12-821042-0.00011-3](https://doi.org/10.1016/B978-0-12-821042-0.00011-3)
- [73] Chopra P, Yadav SK. Fault detection and classification by unsupervised feature extraction and dimensionality reduction. *Complex Intell Sys*. 2015;1(1):25–33. doi:[10.1007/s40747-015-0004-2](https://doi.org/10.1007/s40747-015-0004-2)
- [74] Kansal K, Subramanyam AV. *Autoencoder ensemble for person Re-identification*. in 2019 *IEEE Fifth International Conference on Multimedia Big Data (BigMM)*. 2019.