

Smart railways: AI-based track-side monitoring for wheel flat identification

Mohammadreza Mohammadi¹ , Araliya Mosleh¹, Cecilia Vale¹ ,
Diogo Ribeiro² , Pedro Montenegro¹  and Andreia Meixedo¹

Proc IMechE Part F:
J Rail and Rapid Transit
2025, Vol. 0(0) 1–18
© IMechE 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09544097251313570
journals.sagepub.com/home/pif



Abstract

The wheel flat detection in trains using Artificial Intelligence (AI) has emerged as a critical advancement in railway maintenance and safety practices. AI systems can effectively identify geometric deformation in wheel rotation patterns, indicative of potential wheel flat damage, resorting to wayside monitoring systems and machine learning algorithms. This study aims to propose an unsupervised learning algorithm to identify and localize railway wheel flats, which considers three stages: (i) wheel flat detection to distinguish a healthy wheel from a damaged one using outlier analysis, achieving 100 percent accuracy; (ii) localizing the damage to pinpoint the location of the defective wheel through the Hidden Markov Model (HMM); (iii) classification of wheel damage based on its severity using *k*-means clustering technique. The unsupervised learning algorithm is validated with artificial data attained from a virtual wayside monitoring system related to freight train passages with healthy wheels and defective wheels with single and multiple defects. The proposed methodology demonstrated efficiency and robustness for wheel flat detection, localization, and damage severity classification regardless of the number of defective wheels and their position.

Keywords

Wheel flat detection, wayside condition monitoring, train-track interaction, damage localization, damage classification, unsupervised learning, artificial intelligence

Date received: 20 September 2024; accepted: 27 December 2024

Introduction

Defective wheels on railway vehicles significantly impact both railway infrastructure and vehicle systems. Continuous train operation with defective wheels raises the risk of track misalignment and potential derailments. Over time, the persistent impact of defective wheels can structurally damage railway vehicles and their components, compromising integrity and longevity. Thus, monitoring the geometrical quality of railway vehicle wheels to detect and classify wheel defects based on severity is essential, particularly for freight trains.^{1–3}

The anticipated shift of goods transport from road to rail in the coming years⁴ heightens the need for effective wheel condition monitoring. Defects in freight trains can cause derailments, affecting railway operations. This necessitates novel solutions for the automatic detection and classification of out-of-roundness (OOR) wheels. OOR defects can induce vibrations, damaging both track and vehicle components.⁵ These defects include flats, eccentricities, polygons, corrugations, missing tread material, and other irregularities.^{6–8} Wheel flats, a common type of wheel damage, can be caused by braking forces that change the wheel's perimeter shape from round to flat.⁹ The General Contract for the Use of Wagons¹⁰ mandates the immediate replacement of wheelsets with flats longer than 60 mm on wheels larger than 840 mm. Therefore, early automatic

detection of defective wheels is crucial for safety, track reliability, and controlled maintenance costs.

Defective wheels can be detected using wayside or onboard monitoring systems. Onboard techniques^{11,12} require all vehicle wheels to have sensors, leading to high costs. To mitigate this, wayside measurement systems are commonly used, monitoring wheels of all vehicle types during train passage.⁹ Various sensors, such as fiber Bragg gratings,^{13,14} acoustic emission sensors,^{15,16} strain gauges,^{17,18} and accelerometers,^{9,17,19} are employed to detect wheel flats.

Innovative automatic detection methods are continually enhancing result reliability. Mosleh et al.³ used the envelope spectrum method to distinguish defective wheels and successfully tested on polygonal wheels.¹⁹ Chen et al.²⁰ developed a two-level adaptive chirp mode decomposition (ACMD) approach for wheel flat detection using vehicle

¹ CONSTRUCT – LESE, Faculty of Engineering, University of Porto, Porto, Portugal

² CONSTRUCT, School of Engineering, Polytechnic of Porto, Porto, Portugal

Corresponding author:

Mohammadreza Mohammadi, CONSTRUCT – LESE, Faculty of Engineering, University of Porto, s/n, R. Dr Roberto Frias, Porto 4200-465, Portugal.

Email: up202202392@edu.fe.up.pt

vibration measurements. Nowakowski et al.²¹ proposed a method using vibration signal processing in both frequency and time domains to recognize wheel flats. Wang et al.²² combined conventional time and frequency domain analysis with a band pass filter for envelope analysis in wheel flat detection.

In recent years, machine learning²³ approaches have been applied to detect vehicle system damage. Dernbakh et al.²⁴ used classifier methods such as support vector machine (SVM) and convolutional neural networks (CNN) to identify flat spots on wheels. Wan et al.²⁵ proposed an unsupervised method for detecting anomalies in passenger train wheels using fiber Bragg grating (FBG) sensors. They implemented four unsupervised learning techniques: non-negative matrix factorization (NMF), one-class support vector machine (OC-SVM), multilayer perceptron autoencoder (MLP-AE), and convolutional neural network autoencoder (CNN-AE) to monitor train wheel conditions. Shaikh et al.²⁶ developed a methodology for wheel defect detection based on hybrid deep learning using onboard acceleration data. Ni and Zhang²⁷ introduced a Bayesian machine learning technique for train wheel damage detection and condition assessment, providing more accurate responses compared to offline methods. Mosleh et al.^{1,2} developed an unsupervised detection methodology using acceleration and shear time histories on rails to distinguish defective wheels, which was later applied by Mohammadi et al.⁹ for wheel flat and polygonization detection.²⁸

However, In the authors' previous studies,^{2,3,9,17,29,30} wheel flat detection was performed considering only one defective wheel per train passage. Therefore, in the first stage of this manuscript, the robustness and accuracy of the proposed methodology for wheel flat detection are evaluated, assuming train passages with both single and multiple defective wheels. Moreover, in none of the prior studies conducted by the authors^{1-3,9,17,29-31} was the position of the damaged wheel localized, even for a single defect. Additionally, the distinction between train passages with a single defect and those with multiple defects has not been studied. Thus, in the second stage of this research, it is intended to localize the axle with defective wheels and distinguish multi defects from single defect scenarios. In addition, in the authors' preceding works,^{1,31} damage classification was conducted only for trains with a single damaged wheel. Hence, in the final stage of the present study, damage classification is performed to categorize wheel flat severities regardless of the position or number of defects. The proposed methodology can classify the severity of wheel flats for each defective axle individually. Therefore, to overcome mentioned limitations, this paper enhances the author's previous works and introduces a novel algorithm based on machine learning techniques. This algorithm aims not only to detect wheel flats but also to localize damage and classify it into three severity levels: low, moderate, and severe. To validate the proposed methodology, an artificial wayside monitoring system is utilized, and numerical simulations are conducted to assess the effectiveness of the proposed methodology. Simulation of train passages, including scenarios with both single and multiple defective wheels, is conducted for freight trains. The defect identification is based on an unsupervised machine learning technique

where a confidence boundary is computed to detect wheel defects. Afterward, an automatic segmentation technique is utilized to cut and segment the signal for wheel flat localization and to distinguish single from multi-damage scenarios. Finally, clustering analysis is performed to classify defect based on its severity regardless of the number of defective wheels.

This research work has made the following significant contributions:

- developing an unsupervised damage identification methodology using an accelerometer-based measurements to distinguish defective wheels either train passages including single or multiple damaged wheels, from healthy ones.
- localizing the defect to define the position of the damage and distinguish train passages with a single defective wheel from multiple ones.
- Proposing cluster analysis using the *k*-mean technique in the classification of wheel flat severity to categorize the intensity of the defect into three classes, namely low, medium, and severe, regardless of the number of defects per train passage and position of the damage.

Numerical modeling

This section presents the numerical simulations of the train, track, and the interaction between the wheel and the rail. Moreover, the proposed virtual wayside monitoring system for wheel flat detection, localization, and damage severity classification is also described in this section.

In the present study, simulations of the numerical dynamic interaction between trains and tracks are performed using the in-house software Vehicle-Structure-Interaction (VSI), allowing the generation of synthetic measurement responses. The Vehicle-structure interaction software is described in detail in the study of Montenegro and Calçada.³² This numerical tool, developed in MATLAB,³³ imports the structural matrices from the structure (in this case, the track) and the vehicle that has previously been modeled in a Finite Element (FE) package. In spite of the fact that these subsystem models are originally created individually in the FE program, the VSI software uses a fully linked technique to couple them. The train is coupled to the track by a 3D wheel-rail contact model using Hertzian theory.³⁴ Normal and tangential contact forces caused by rolling friction creep are calculated by the USETAB routine.³⁵ Moreover, The simulation of the track is conducted using the software ANSYS.³⁶ Beam elements are used to model rails and sleepers, while spring-dashpot components are used to simulate the behavior of the track flexible layers, that is, ballast, fasteners/pad, and mass point components to consider the ballast mass. ANSYS³⁶ has also been used to model the train, which is a freight Laagrss-type vehicle composed of five wagons, each one with two axles. The train modeling uses a multibody formulation in which mass point elements are placed at the center of gravity of each body, specifically the carbody and wheelsets, allowing inertial and mass effects to be simulated. The work of Mosleh et al.^{3,17} thoroughly describes in detail the mechanical characteristics of both the track and train components. The graphical depiction

of the numerical simulation of the wayside monitoring system is shown in Figure 1.

In the authors' previous studies,^{1,2} different sensitivity analyses have been implemented to evaluate the effect of the sensors position on the accuracy of the proposed methodology for wheel flat detection. From the results of the mentioned studies, the proposed methodology can detect wheel flat regardless of the number and position of the accelerometers installed on the rail or sleeper. The location of the sensors in the proposed virtual wayside monitoring system is shown in Figure 1(d). The wheel flat-identification system is composed of eight accelerometers installed on the midspan of the right and left rails. The accelerometers on the right side are identified by measurement points 1 to 4. Additionally, the accelerometers located on the left side of the track are represented by measurement points 5 to 8. The distance between the accelerometers is considered equal to 0.6 m. Moreover, one strain gauge is installed on the midspan near the accelerometer 1 to obtain strain for the damage localization process (stage 2).

AI-based methodology for wheel flat identification

In this research, the proposed methodology to identify wheel flat automatically includes three stages, as presented in Figure 2, according to a machine learning methodology^{1,2,37}:

In the first stage, the rail baseline responses are used to establish a confidence boundary, while the second stage localizes the defective wheels and distinguishes single-damage from multi-damage passages. Finally, the third stage identifies the severity of the damages and classifies them into three categories (low, moderate, and severe damage).

Stage 1: Detection

Damage detection is performed by applying six distinct steps which are shown in Figure 2. First, data acquisition collects dynamic responses from accelerometers. Next, feature extraction reduces the data dimension using Continuous Wavelet Transform (CWT) to convert time-series data into damage-sensitive features. Principal Component Analysis (PCA) then mitigates environmental and operational effects on vibration characteristics. PCA is used to reduce model orders,³⁸ perform modal analyses,³⁹ update nonlinear models,⁴⁰ validate sensors, and locate damage.⁴¹ Afterward, features are merged into a damage index (DI) using Mahalanobis Distance (MD), simplifying data analysis by reducing multivariate data to a single DI. MD measures the similarity between undamaged and damaged features; shorter distances indicate higher similarity. Data fusion enhances damage sensitivity for CWT-PCA-based features. An unsupervised discrimination algorithm defines a confidence boundary (CB) for each accelerometer using

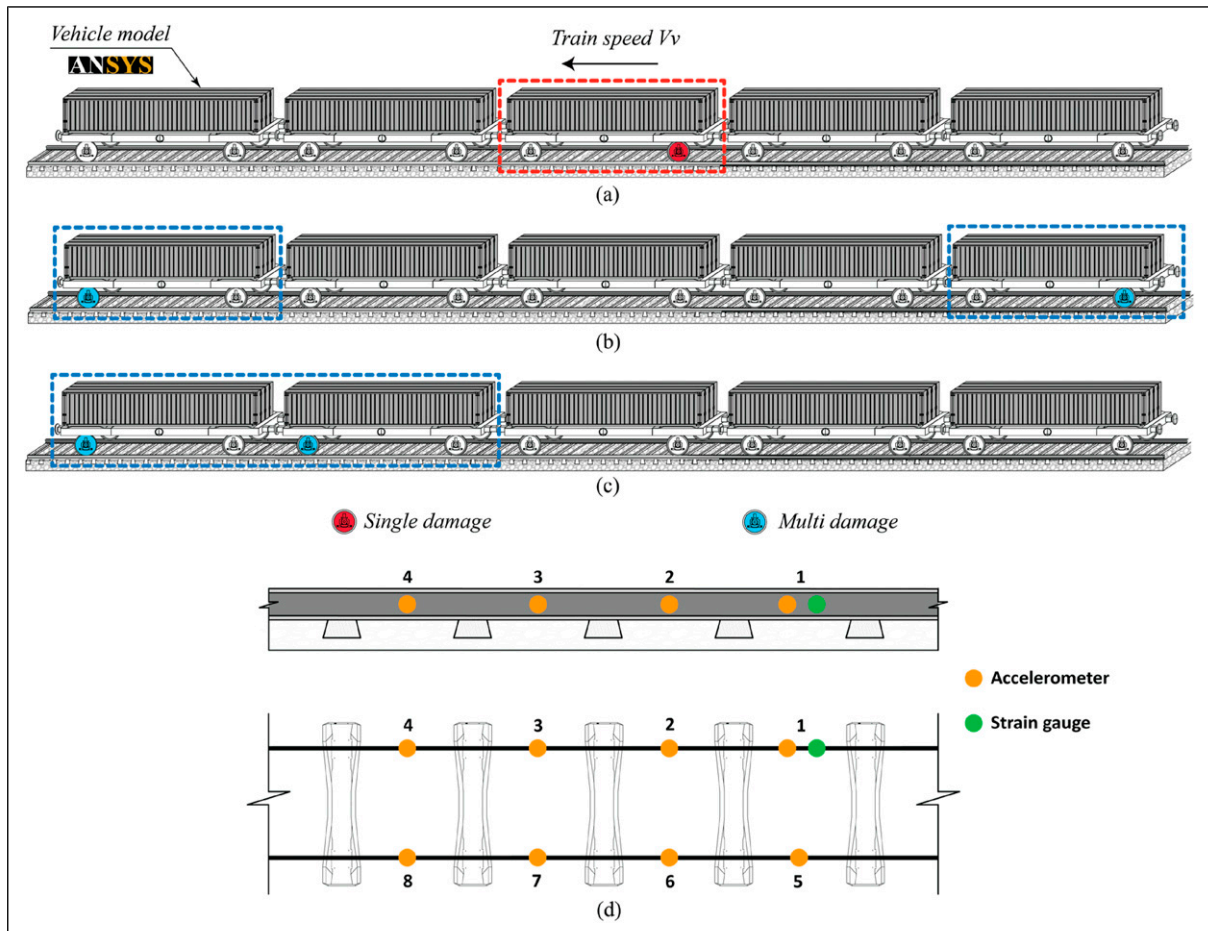


Figure 1. Numerical modeling of the track-side monitoring system.

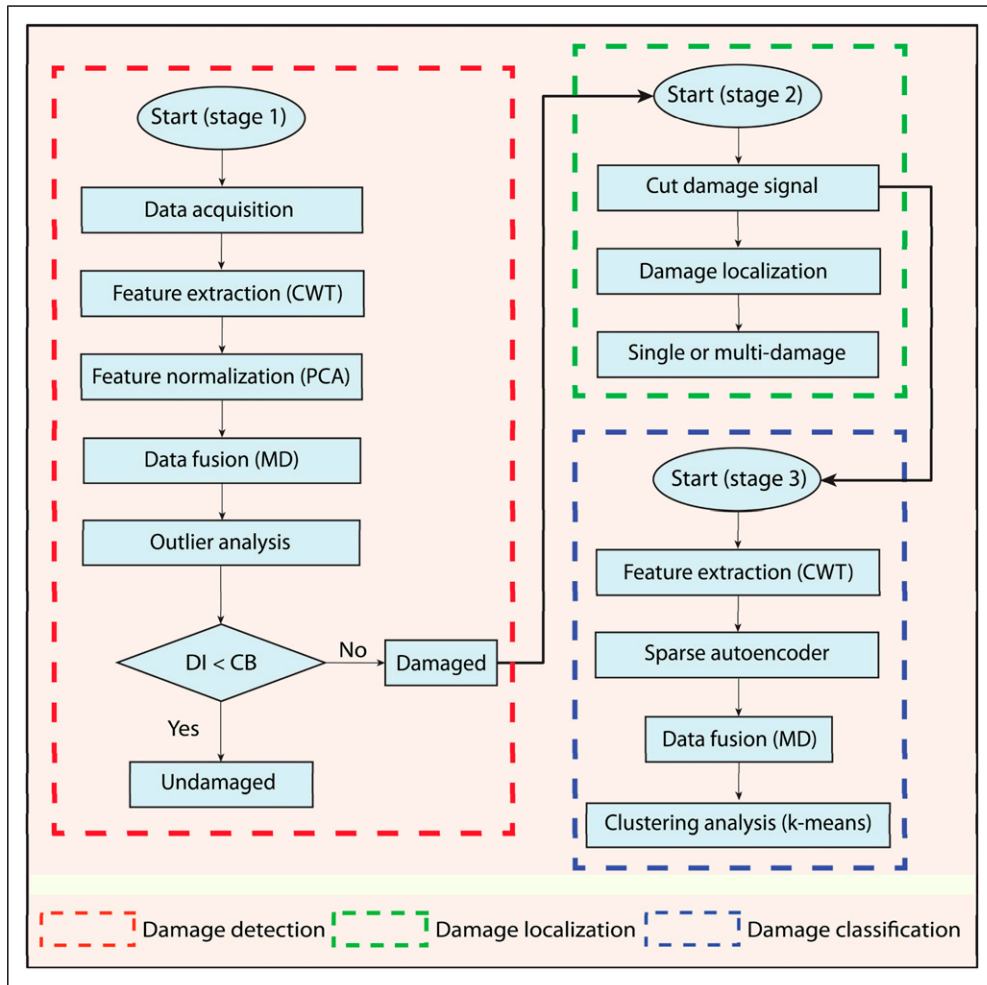


Figure 2. Flowchart of wheel flat detection, localization, and classification of the severity.

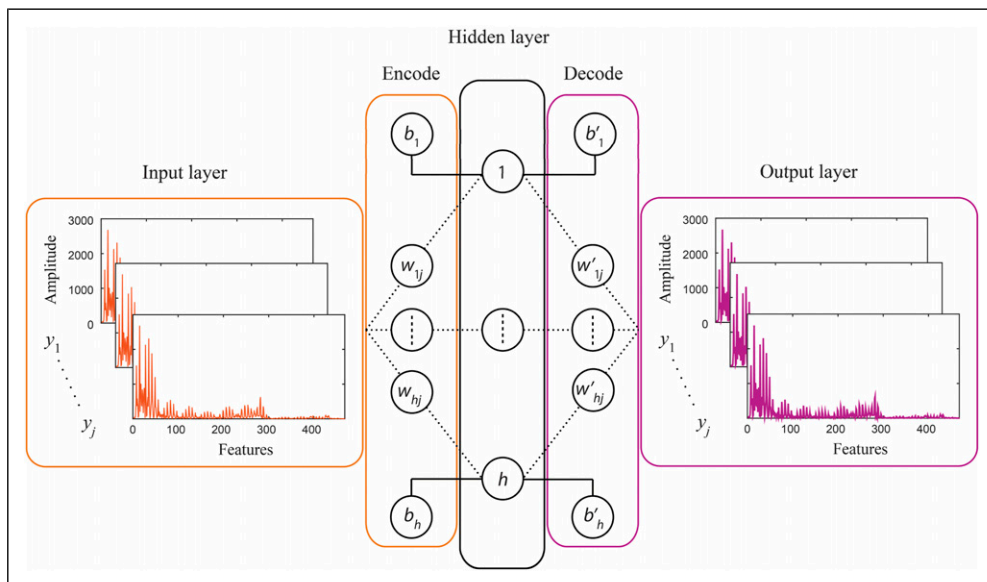


Figure 3. Structure of a sparse autoencoder model.

outlier analysis. Mahalanobis squared distance is approximated by chi-squared distributions, allowing a Gaussian distribution for MD. The Gaussian inverse cumulative distribution function (ICDF) estimates the CB for outlier detection, considering the baseline feature

vector's mean and standard deviation. A feature is an outlier if its damage indicator meets or exceeds the CB, with a 1% threshold based on Sousa Tomé et al.⁴² and other works.^{2,9,37} Finally, defective scenarios are identified and used to localize damaged wheels.^{43,44}

Stage 2: Localization

The method divides the signal to capture detailed vibration patterns to localize a defective wheel, aiding in obtaining specific characteristics in both time and frequency domains. This segmentation counts and labels each wheel's passage, allowing further analysis of wheel-rail interactions. The Hidden Markov Model (HMM) is chosen for its processing speed and ability to link observation sequences to hidden states.⁴⁵ The Baum-Welch algorithm deduces model parameters using an Expectation-Maximization strategy,⁴⁶ while the Viterbi and Forward-Backwards algorithms define probability vectors and grids.⁴⁷ This method links reliable signal sections to the associated wheel, enhancing the examination of wheel-rail interactions and pinpointing specific points for comparing patterns across wheel movements. Gaussian emissions between 2 and 5 states were considered for adjusting hyperparameters, with a 3-state HMM chosen for segmentation as it effectively eliminated segments between wheel movements.

The segmentation captures three phases of a wheel's movement: approach, detection, and departure from the sensor. This detailed segmentation helps identify the effects of a defective wheel on the accelerometer signal. With wheel damage, the accelerometer signal's amplitude varies, while the strain gauge signal remains consistent, facilitating automatic segmentation of significant signal portions for each passage. Following segmentation, axles are individually counted and separated, and defective wheels are localized based on acceleration peak variations in both single and multi-damage scenarios. These scenarios can be distinguished by specific amplitude variation trends, as discussed in the results and discussion section for wheel flat localization.

Stage 3: Classification

Damage classification is performed to categorize defects based on the defect severity. For this proposal, the first

Table 1. Damaged and undamaged scenarios.

	Baseline scenario	Damaged scenario	
		Single damage	Multi-damage
Train	Freight – Laagrss wagon	Freight – Laagrss wagon	Freight – Laagrss wagon
Number of loading schemes	6	1 (full capacity)	1 (full capacity)
Unevenness profiles	4	1	1
Speeds (km/h)	40 – 120	80	80
Noise ratio	5%	5%	5%
Flat lengths (mm)	—	- 10–20 mm (low) - 25–50 mm (moderate) - 55–100 mm (severe)	
Number of numerical analyses	113	30	12

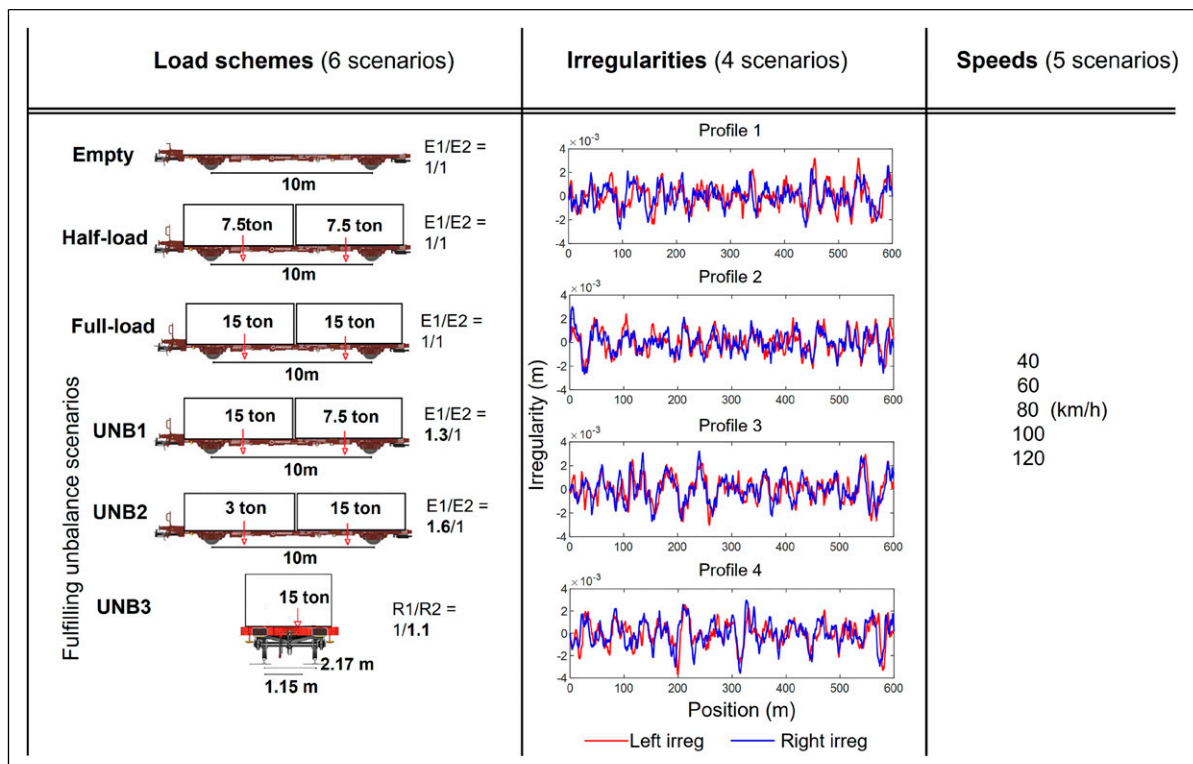


Figure 4. Baseline scenarios.

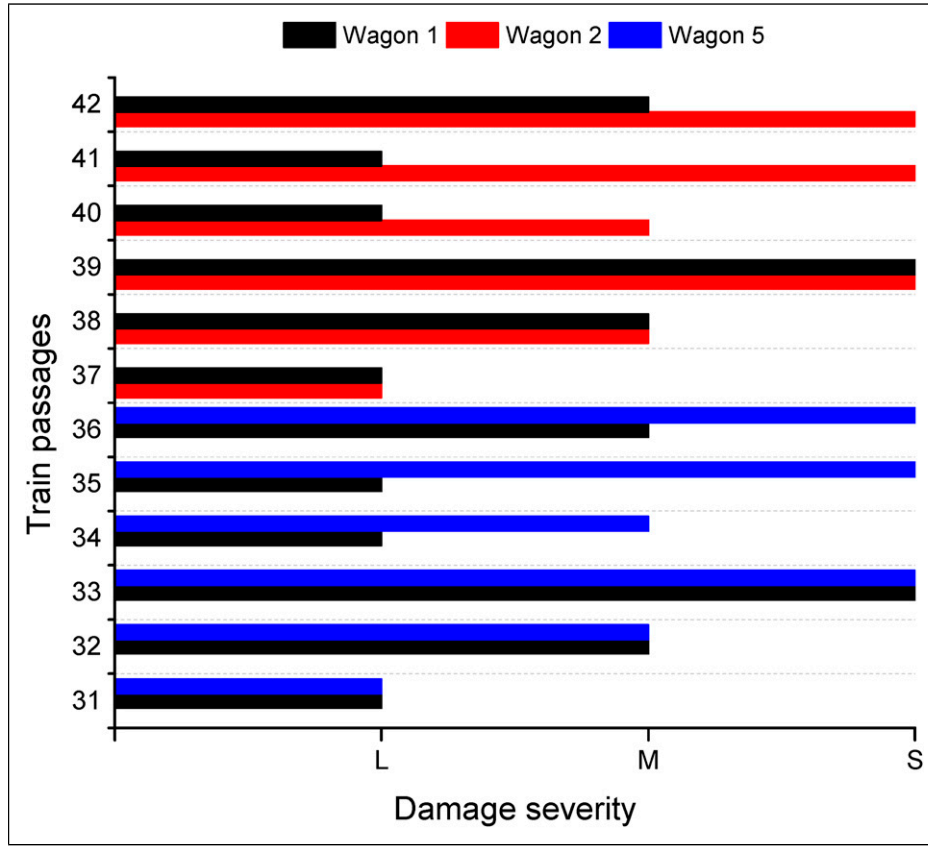


Figure 5. Multi-damage scenario simulations.

output of the second stage, the cut signals corresponding to axles with defective wheels, is used as input data in the third stage. Afterwards, to reduce the size of the dynamic response matrices, feature extraction is conducted. For this intention, the CWT is used to convert time-series data into features sensitive to the damage. Afterward, the sparse autoencoder is implemented to estimate input variables (reconstruction). The encoder component transforms the input data y_j into a lower-dimensional space S and produces an output represented as \hat{y}_j . The autoencoder process for each layer L is described as follows:

$$s_h = \varphi \left(\sum_{j=1}^n w_{hj} \cdot y_j + b_h \right) \quad (1)$$

$$\hat{y}_j = \varphi' \left(\sum_{h=1}^n w'_{hj} \cdot s_h + b'_h \right) \quad (2)$$

In which, n represents the number of elements in the acceleration response vector, y_j denotes the j th element in the original CWT-based damage-sensitive feature response data, \hat{y}_j represents j th element in the reconstructed CWT-based damage-sensitive feature response data. w, w' and b_h, b'_h are the weight matrices and bias vectors for encoder and decoder modules, respectively, while φ, φ' are the activation functions of the encoder and decoder. The cost function utilized for training the sparse autoencoder is a modified version of the mean squared error function, shown as follows:

$$E = \frac{1}{n} \sum_{j=1}^n \sum_{h=1}^h (y_{hj} - \hat{y}_h)^2 + \lambda \cdot \frac{1}{2} \sum_{j=1}^n \sum_{h=1}^h (w_{hj})^2 + \beta \cdot \sum_{h=1}^L \text{KL}(\rho | \hat{\rho}_j) \quad (3)$$

Where, λ represents the coefficient for the regularization term, β represents the coefficient for the sparsity regularization term and ρ denotes the average desired information gain, which reflects the sparsity proportion, while $\hat{\rho}_j$ is the average information gained in the training process. The Kullback-Leibler⁴⁸ divergence is a mathematical function used to quantify the disparity between two distributions. In this case, it considers a zero value when ρ and $\hat{\rho}_j$ are identical and escalate as they move further apart. Minimizing the cost function makes this term small; therefore, ρ and $\hat{\rho}_j$ become close to each other. For the training process of the SAE model, the Scaled Conjugate Gradient algorithm (SCG)⁴⁹ with stopping criteria when the model achieves the maximum number of epochs. Thereafter, Mean Squared Error (MSE) and Mean Absolute Error (MAE) are computed between original CWT-based features y_j and reconstructed ones \hat{y}_j as new features by equations (4) and (5), respectively:

$$\text{MSE} = \frac{1}{n} \cdot \sum_{j=1}^n (y_j - \hat{y}_j)^2 \quad (4)$$

$$\text{MAE} = \frac{1}{n} \cdot \sum_{j=1}^n (y_j - \hat{y}_j) \quad (5)$$

It is possible to differentiate between defect severities through calculated errors (MSE and MAE). Defective wheels with low, medium and, severe damage come with lower, medium and, higher MSE and MAE, respectively. Data fusion is implemented by computing Mahalanobis Distance to increase the sensitivity of the features (MSE and MAE) to the damage. Data fusion can involve two steps. Firstly, the computed MSE and MAE from the autoencoder are merged for each sensor separately (feature fusion). Secondly, the MSE and MAE from all sensors are fused in the subsequent step. A diagram depicting the employed SAE architecture is shown in Figure 3 as follows:

Finally, the k -means clustering technique is applied to classify damage severities, necessitating the predefining number of clusters. To determine the appropriate number of clusters, the global silhouette index (SIL) is used. Through the k -means algorithm, the centroid (CD) of each cluster is then identified. The k -means clustering is implemented through the following equation:

$$d(x, c) = \sum_{j=1}^p |x_j - c_j| \quad (6)$$

Where p is the number of the damage indexes (DI), x , and c represent each DI and centroid of each cluster, respectively.

Damaged and undamaged scenarios

In order to test and validate the automatic wheel flat identification method developed in this study, baseline (undamaged) and damaged wheel scenarios are considered. The baseline scenario describes the train passing with healthy wheels, whereas the damaged scenario represents a train passing with defective wheels. As shown in Figure 1, three defective layout cases are considered for the damage scenarios, including one single damage and two multi-damage scenarios: (i) flat in the left wheel of the rear wheelset of the third wagon corresponding to single-damage scenarios which is presented in red color (Figure 1(a)), (ii) defects on the left wheel of the rear wheelset of the first wagon and the left wheel of the front wheelset of the last wagon which is presented with blue wheels correspond to multi-damage scenarios (Figure 1(b)) and, (iii) flats in the left wheel of the rear wheels of the first and second wagon are shown in blue, related to multi-damage scenarios (Figure 1(c)).

To simulate defective wheels on train, different wheel flat severities are also taken into account in this study. Based on the wheel flat lengths (L), three intervals are defined for defective wheels, namely low, moderate, and severe. In each interval, the uniform distributions $U(10, 20)$, $U(25, 50)$, and $U(55, 100)$ define the lower and upper limits for the wheel's

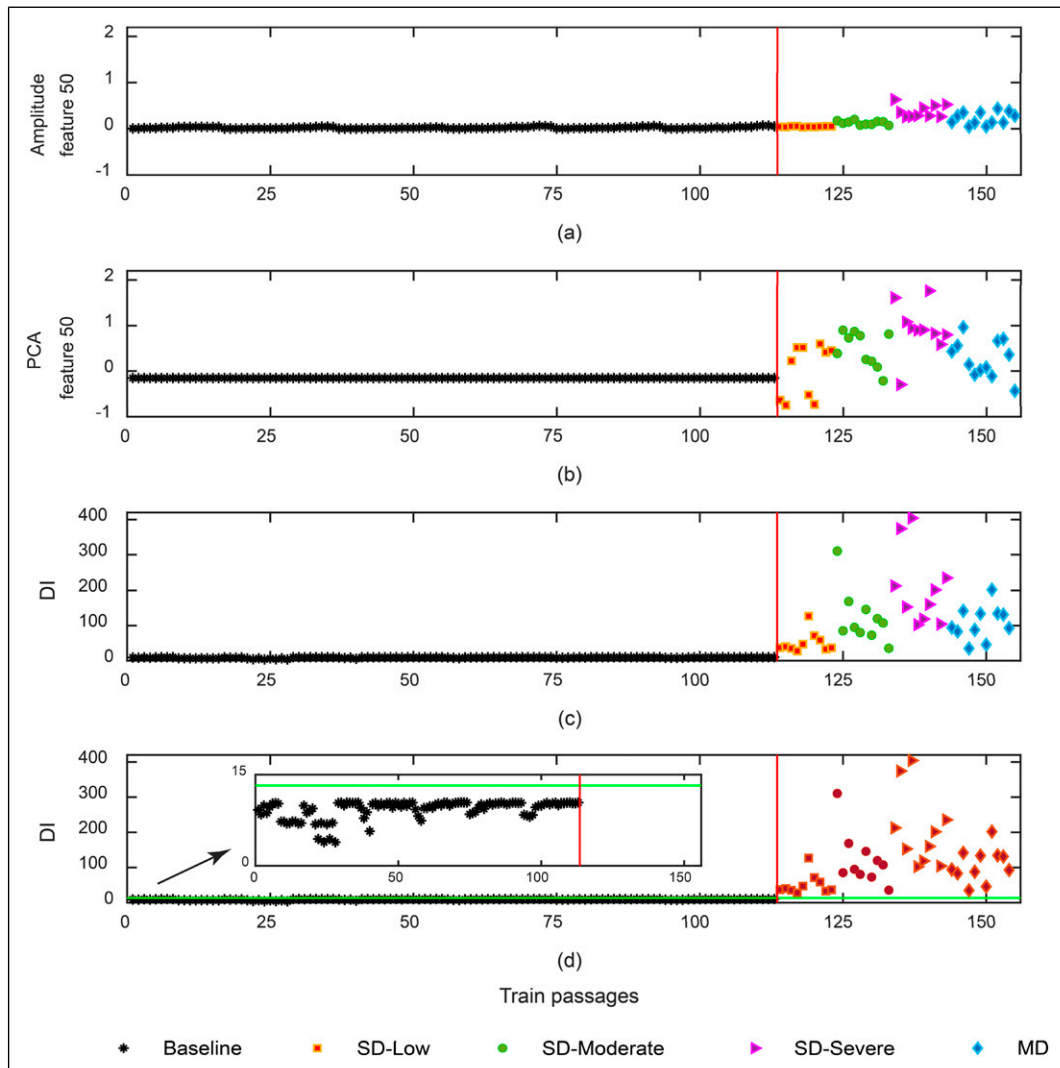


Figure 6. The output of automatic damage detection using measurements from accelerometer on position 5: (a) feature extraction; (b) feature normalisation; (c) data fusion; (d) outlier analysis.

flat length, respectively. The wheel flat depth (D) is computed by the following equation⁵⁰:

$$D = \frac{L^2}{16R_w} \quad (7)$$

in which R_w and L represent the radius of the wheel and the length of the flat length, respectively.

Moreover, a wheel flat's vertical profile is calculated as follows⁵⁰:

$$Z = -\frac{D}{2} \left(1 - \cos \frac{2\pi x}{L} \right) \cdot H(x - (2\pi R_w - L)), 0 \leq x \leq 2\pi R_w \quad (8)$$

where x is the coordinate aligned with the track longitudinal direction, and H represents the Heaviside periodic function. In real-world conditions, where rails are not completely smooth, track irregularities significantly affect wheel-rail contact force values.^{51,52} Thus, it is important to consider these irregularities even though they are very small in the numerical analyses. In accordance with the European Standard EN 13848-2,⁵³ rail unevenness profiles are generated for wavelengths between 1 m and 75 m. Using actual data, PSD curves are developed to create artificial unevenness profiles. An overview of how unevenness profiles are generated can be found in Mosleh et al.⁵⁴ In order to apply the proposed methodology, dynamic responses are measured at sensors (shown in Figure 1(d)) along the rail in both baselines (undamaged) and damaged scenarios whose assumptions are indicated in Table 1.

For the baseline scenario, 113 simulations are performed, as shown in Figure 4, by considering four

unevenness profiles of the rail, five train speeds (vary between 40 and 120 km/h), and six loading schemes: (i) empty train; (ii) half-loaded train; (iii); fully loaded train with longitudinal and transversal unbalanced loads (UNB1, UNB2, and UNB3). The unbalanced loading schemes for the wagon model are defined based on the UIC loading guidelines,¹⁰ where the cargo gravity centre is offset in longitudinal and transversal directions. Additionally, the ratio between reaction forces in the axles is described in Figure 4,¹⁰ in which E1, E2, R1, and R2 represent reaction forces in the front, rear, left, and right axles, respectively.

To simulate damaged cases, 42 scenarios are considered with different flat geometries while the train speed is 80 km/h containing 30 ton load per wagon. 30 simulations include the train passing with one damaged wheel, and the other 12 simulations involve multi-damage cases. In the case of single damage, 30 analyses are conducted for each flat length interval (low, moderate, and severe), as illustrated in Table 1.

For multi-damage scenarios, 12 analyses are performed (Figure 5). Six of these analyses involve two defective wheels with the same severities, namely (low-low, moderate-moderate, and severe-severe) for both the first and second wagons (train passages 31, 32, and 33), identified as far-multi-damaged (Figure 1(c)), as well as for both first and last wagons (simulations 37, 38 and 39), termed near-multi-damaged (Figure 1(b)). Additionally, six analyses examine two defective wheels with different severities, including low and moderate, low and severe, and moderate and severe flat properties (train passages 34, 35, and 36), for both the first and last wagons (Figure 1(b)), as well as for the first and second wagons for simulations 40, 41, and 42 (Figure 1(c)).

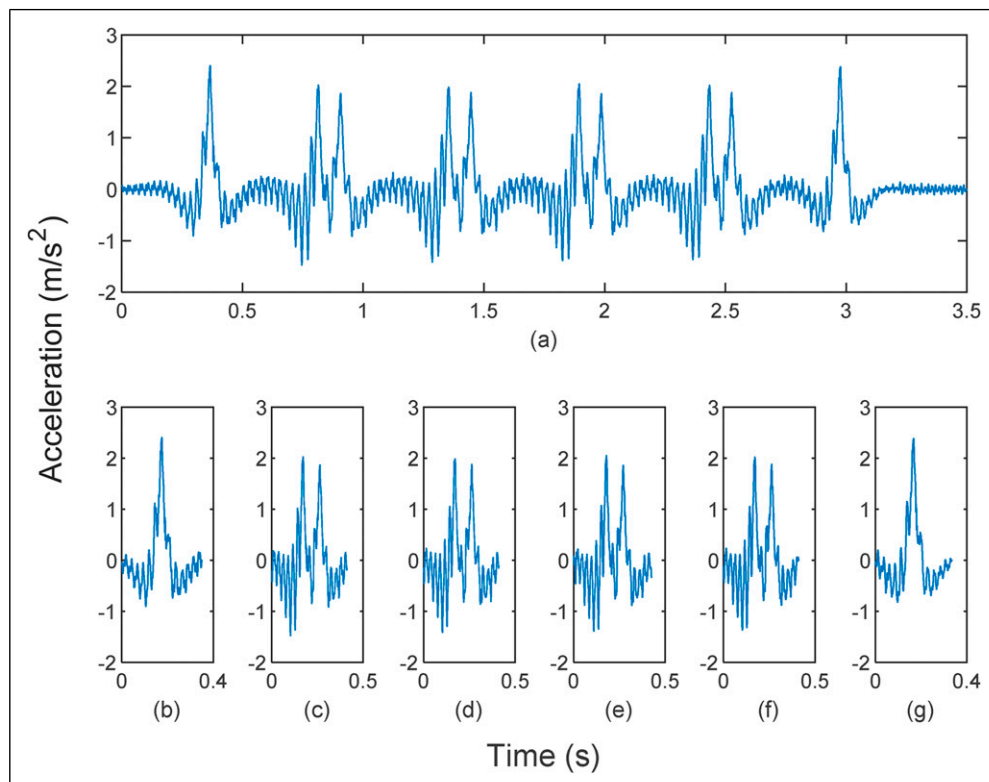


Figure 7. Acceleration signal segmentation considering healthy train passage from accelerometer 1: (a) complete signal; (b) axle 1; (c) axles 2–3; (d) axles 4–5; (e) axles 6–7; (f) axles 8–9 (g) axle 10.

A sampling frequency of 10 kHz is used to evaluate acceleration signals for baseline and damaged scenarios and based on the literature,^{3,17,55} an artificial noise (5% of amplitude) is used to pollute the numerical signal for a more realistic reproduction of the measured rail response. Then, a low-pass Chebyshev type II digital filter with a cut-off frequency of 500 Hz is also implemented to filter all time series.

Results and discussion

The output of the current study is presented in three distinct sections. Initially, defective wheels are distinguished from healthy ones, while damaged wheels are localized through the second stage. Finally, damage classification is performed based on the defect severity.

Wheel flat detection

This section summarizes the results of four techniques for wheel flat detection: feature extraction, feature normalization, data fusion, and outlier analysis (Figure 2, stage 1).

Detailed studies on each technique have been previously published by the authors.^{2,9,30,55}

Figure 6 shows the damage detection output for each technique. Features and damage indexes are classified into baseline (first 113 passages) and damaged scenarios (next 42 passages). Each damage scenario is indicated by symbols representing different wheel flat severity levels. Simulations 114–123 represent low severity (10–20 mm, SD-low), 124–133 moderate severity (25–50 mm, SD-moderate), and 134–143 severe damage (55–100 mm, SD-severe). Simulations 144–155 show multi-damage scenarios detailed in section 4. From time series data, 468 features are extracted per accelerometer as damage-sensitive indicators. The analysis includes 155 simulations for both baseline and damaged scenarios (Table 1). Applying Continuous Wavelet Transform (CWT) produces a three-dimensional matrix ($155 \times 468 \times 8$) for each of the eight measurement points (Figure 1(d)). Figure 6(a) illustrates features extracted by CWT at position 5, showing an ascending pattern in single-damage scenarios with clear amplitude differences. However, no clear distinction between baseline and damage scenarios is observed due to environmental and operational factors, leading to feature normalization.

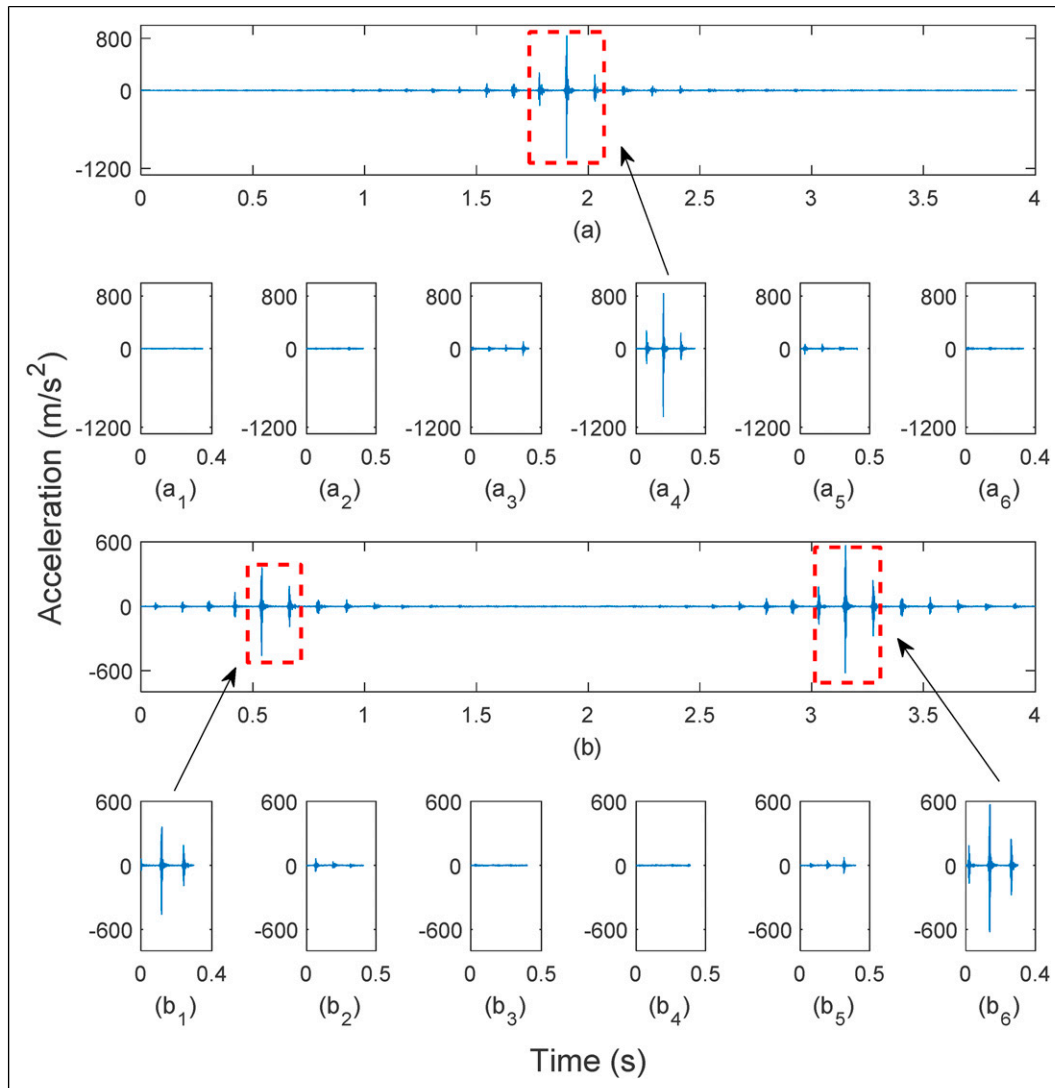


Figure 8. Damage localization considering defective scenarios using measured acceleration from accelerometer I: (a, b) complete signal; (a₁, b₁) axle 1; (a₂, b₂) axles 2–3; (a₃, b₃) axles 4–5; (a₄, b₄) axles 6–7; (a₅, b₅) axles 8–9 (a₆, b₆) axle 10.

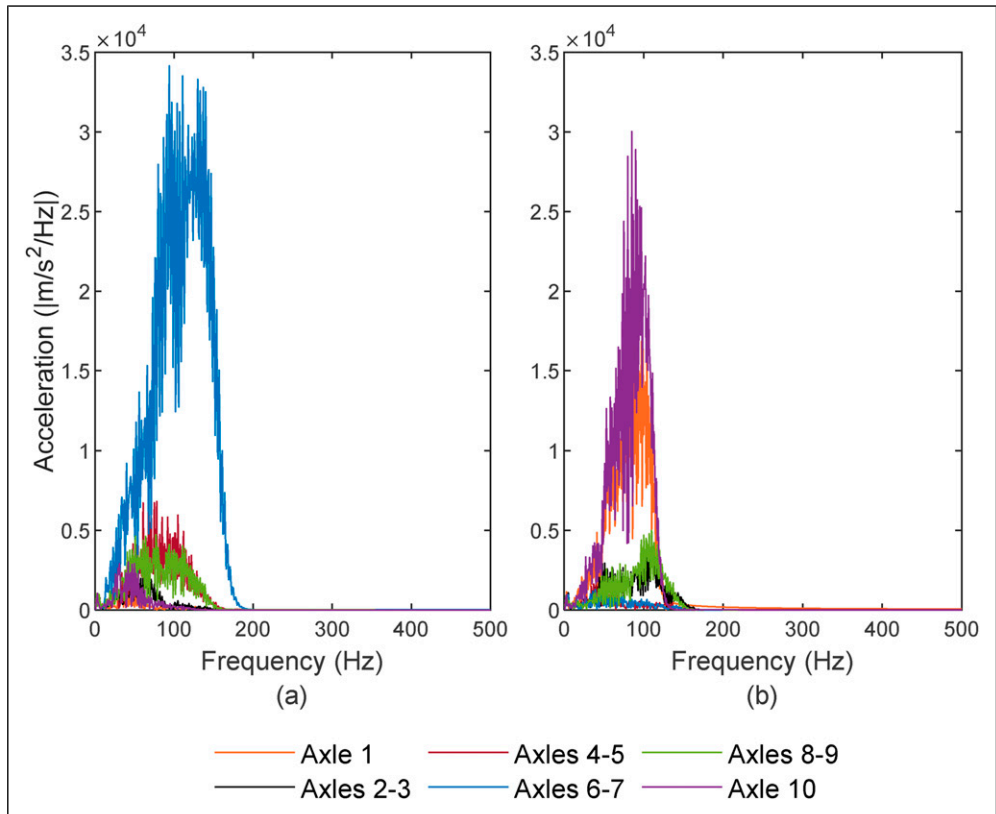


Figure 9. Damage localization considering defective scenarios using measured acceleration in the frequency domain from accelerometer 1: (a) single damage; (b) multi-damage.

Afterward, PCA is applied to create an 8×468 feature matrix based on CWT features for each passage. The first 15 components are discarded, as they account for over 80% of variance. Figure 6(b) shows feature normalization, suppressing changes from environmental and operational variations (EOVs). Despite clear sensitivity patterns for single-damage scenarios, distinguishing between undamaged and damaged wheels after PCA is not possible. Therefore, data fusion is implemented to increase damage sensitivity.

Data fusion is implemented to compute Mahalanobis Distance (MD) obtaining a damage index (DI) to transform each sensor and passage into a damage-sensitive feature based on 453 PCA-CWT parameters. This results in 155×1 vectors of MD for each of the 8 accelerometers. Figure 6(c) shows improved damage sensitivity at accelerometer 5, with varying sensitivities across the sensor, resulting in different damage indexes.

Figure 6(d) shows the methodology's effectiveness in distinguishing undamaged from damaged scenarios by comparing confidence boundaries (113 baseline passages) and damage indices. Healthy wheels come with DI_s below the confidence boundary (CB) for the first 113 passages while for defective scenarios, DI_s exceeds CB. The robustness of the confidence boundary (CB) improves as more baseline data is added daily, enabling more accurate detection of defective wheels in real-time. Once the confidence boundary is established, a single passage with defective wheel is sufficient to test the accuracy of the proposed methodology.² This methodology effectively detects all damage scenarios without false positives or negatives, indicating that a single sensor is

Table 2. SAE model characteristics.

Number of hidden layers	Maximum number of epoch	λ	B	α	Activation function
10	1000	10^{-3}	1	0.7	Sigmoid

sufficient for damage detection regardless of the number and position of the defect.

Wheel flat localization

The second stage of the current study represents the outputs of the damage localization to diagnose the axles, including the defective wheels. As described in section 3, damage localization is performed based on the segmentation technique explained in detail in the work of Lourenço et al.⁵⁶ Initially, the HMM was tailored to the contour of the gauge signal, taking into account 2 to 5 states of Gaussian emissions. All states managed to distinguish between successive wheels even when wagons were closely linked. Yet, only when using 2 and 3 hidden states was a consistent division for every wheel movement. Although other segmentations effectively conformed to the shape of the monitoring signal, they proved unreliable in cases with high levels of noise and defective conditions. As a result, the HMM with 3 states was selected for automatic segmentation since it effectively eliminated the wheel signal segments between wheel movements. Additionally, it can be noted that the proposed methodology for signal segmentation is unable to segment the signal for each

wheel individually. The segmentation method currently employed divides the signal into segments, each representing two wheels, except for the defective wheel on the first and the last wagons. To implement the mentioned segmentation method, it is necessary to use both strains and accelerations measured by sensors while the train is passing with the defective wheel.

Figure 7 illustrates the acquired acceleration from accelerometer 1 (Figure 1(d)) for train passage without defective wheel. Figure 7(a) represents the complete signal before the implementation of the segmentation technique. Figure 7(b) and (g) show the segmented acceleration signal for the first and last axles individually, while Figure 7(c)–(f) represent the signal for the other axles as a pair of front and rear wheels of wagons in succession after segmentation. As shown in these figures, when the train is passing considering all wheels in healthy condition, the measured acceleration does not represent any impact and significant peak.

Figure 8 represents damage localization for both single and multi-damage scenarios using measured acceleration on position 1 (Figure 1(d)). Figure 8(a) illustrates the measured acceleration for a train passage with a single damaged wheel located on the third wagon ($L = 97$ mm), showing the complete signal before segmentation. Figure 8(a)_{1,6} display the segmented acceleration signal for the first and last axles, respectively, while Figure 8(a)_{2–5} represent the signal for the

other axles in pairs (front and rear wheels of successive wagons) post-segmentation. As seen in Figure 8(a)₄, the maximum acceleration peak occurs on axles 6–7, indicating significant discrepancies. The dynamic response in Figure 8(a)₂ shows a higher acceleration peak compared to the healthy condition (Figure 11(e)), suggesting the defective wheel is at axles 6–7. Additionally, the acceleration peaks increase from axle 1 to axles 6–7 and decrease from axles 6–7 onward, a trend not observed in healthy train passages (Figure 8(a)_{1–6}). Sensor 1 captured a single impact from the defective wheel, notifying the presence of only one damaged wheel. Thus, when responses from different axles come with a lag, it can be indicated as a wheel flat.

Figure 8(b) presents the damage localization for multi-damage scenarios with defective wheels on the first and last wagons ($L = 68, 72$ mm), showing the complete signal before segmentation. Figure 8(b)_{1, 6} illustrate the segmented signals for the first and last axles, respectively, while Figure 8(b)_{2–5} display the segmented signals for the other axles in pairs. As shown in Figure 8(b)_{1,6}, accelerometer 1 detects two significant peaks, the first at the first axle and the second at the last axle. The acceleration peak decreases from axle 1 to axles 6–7, then increases from axles 8–9 to axle 10, with the second peak at the last axle. This pattern indicates two impacts recorded by accelerometer 1, confirming a multi-damaged wheel scenario.

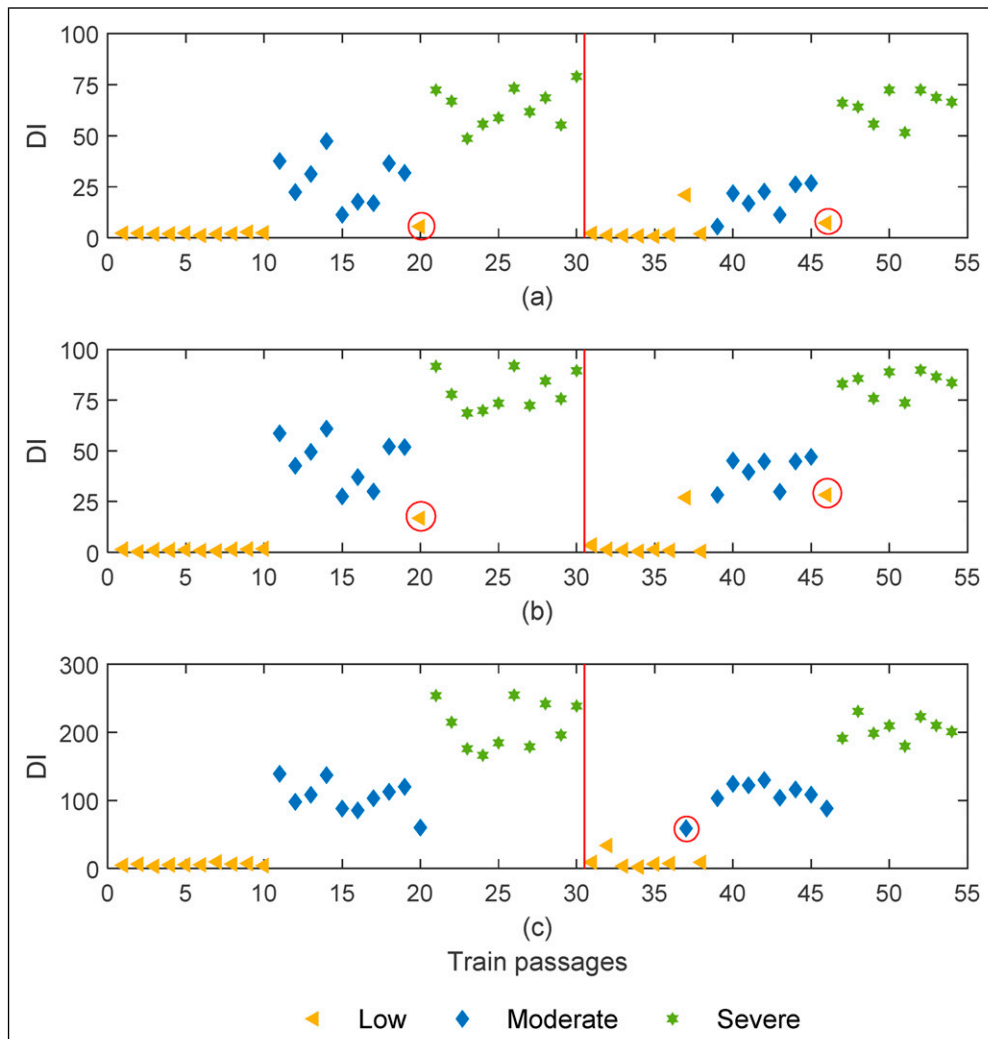


Figure 10. Automatic wheel flat damage classification considering both single and multi-damage scenarios using dynamic responses from accelerometers: (a) accelerometer 1, (b) accelerometer 5, (c) all accelerometers.

Comparing Figure 8(b)_{1,6} with Figure 7(b) and (g), it is evident that a wheel flat causes a significant lag in acceleration amplitude across different axles, whereas healthy wheels show insignificant lag. The difference in acceleration peak amplitude between healthy and defective scenarios further helps localize damaged wheels. Hence, train passages with defective wheels exhibit higher acceleration peaks than healthy scenarios (Figure 8(b)_{1,6} compared to Figure 7(b) and (g)).

Figure 9(a) and (b) illustrate the measured acceleration from accelerometer 1 in the frequency domain, considering train passages with single and multiple damaged wheels, respectively. For single damage scenarios (Figure 9(a)), axles 6–7 exhibit a higher frequency peak compared to other axles, indicating the presence of a damaged wheel at these axles. By comparing the frequency data from individual axles (Figure 8(a)_{1,6}) and paired axles (Figure 8(a)₂₋₅), a significant frequency lag between axles 6–7 and the others is evident. This suggests that accelerometer 1 detects a substantial impact caused by the wheel defect, confirming a single defective wheel scenario.

Figure 9(b) shows the damage localization results for multi-damage cases, where two damaged wheels are located on the first and last wagons. This is identified using segmented measured acceleration in the frequency domain. The figure highlights two prominent acceleration peaks,

corresponding to axle 1 and axle 10, represented by orange and purple colors, respectively. This indicates defective wheels on the first and last axles. Additionally, there is a noticeable acceleration lag between these axles and the other paired axles, confirming the presence of two defective wheels. The higher acceleration peak for the last axle compared to the first axle is due to the longer flat length of 72 mm on the last axle's defective wheel, which causes a more intense impact than the 62 mm flat length of the first axle's defective wheel.

Wheel flat classification

Based on the methodology presented in section 3, it is possible to detect a wheel flat automatically in the first stage, utilizing the complete signal. Subsequently, in the second stage, defect positions are determined based on the segmentation technique corresponding to each axle individually. Following that, the severity of wheel flat is classified in 42 passages with damaged wheels, including both single and multi-damage scenarios. It's important to note that only truncated signals (not complete signals) are considered in this third stage. Following this, the methodology involves the extraction of features using Continuous Wavelet Transform, application of autoencoder, and data fusion techniques to determine a damage index. The

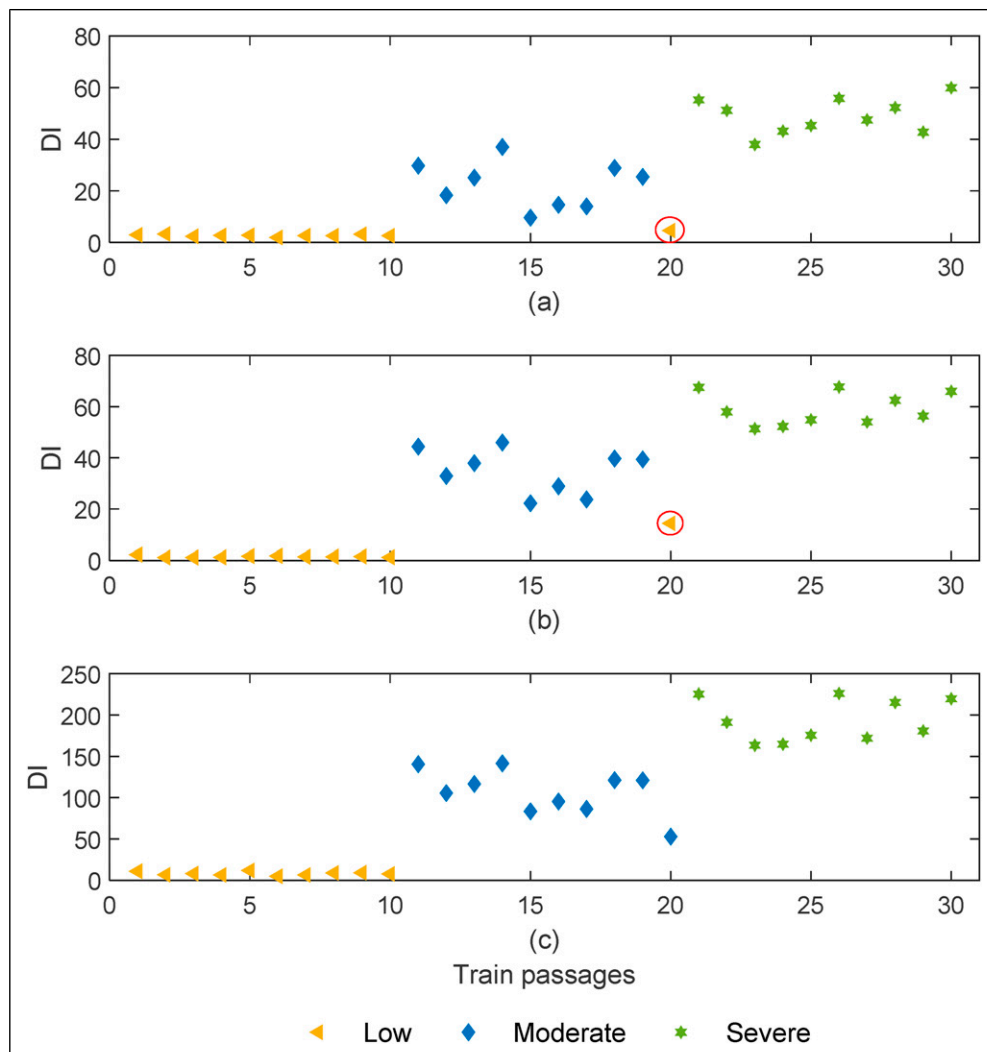


Figure 11. Automatic classification of wheel flat damage using accelerometer data for the second wheel of the 3rd wagon: (a) accelerometer 1, (b) accelerometer 5, (c) all accelerometers.

severity of the defect is classified using the k-means technique based on the defect severity. The model training and numerical computations are performed on PC with processors including Intel (R) Core i7 – 10700 CPU @ double 2.90 GHz and 40 GB RAM. The design of the SAE structure is based on the train Autoencoder algorithm in MATLAB software.³³ To train the Sparse Autoencoder model, the baseline cases are used. Moreover, all damaged scenarios are utilized for the testing phase. The implemented SAE is designed considering a coefficient for the regularization term (λ), the coefficient for the sparsity regularization term (β), and the percentage of activation of the hidden unit (α). Table 2 represents the SAE model as follows:

Figure 10 shows automatic wheel flat classification considering segmented responses obtained from each accelerometer. As shown in this figure, the damaged scenarios are divided into two sections by red line. The initial 30 trains refer to passages including single defective wheels, while the remaining 24 indicators correspond to multi-damage scenarios involving two defective wheels. As indicated in Table 1, multi-damage scenarios include 12 damaged train passages. However, as detailed in stage 2, each acceleration signal is segmented separately for all 2 single and 4 pair axles. Given that each multi-passage involves two damaged wheels, therefore, multi-defective scenarios consist of 24 signals. It's worth noting, as depicted in Figure 1, that defective wheels are observed in both the far wagons (Figure 1(b)) and the near wagons (Figure 1(c)).

From the results illustrated in Figure 10(a) and (b), it is possible to conclude that the proposed methodology can classify damaged wheels based on their severity (low, moderate, and severe) with only one sensor, regardless of

whether the train has single or multiple damaged wheels. The damage index for each defective wheel is computed by implementing only the first part of the data fusion, which merges the features for each sensor separately. As depicted in Figure 10(a) and (b), wheel flat classification using only one sensor exhibits only two misclassifications.

One of these damage indexes corresponds to single damage scenarios, where the length of the flat measures 27.5 mm. This value closely approaches the boundary of low damage severity (10–20 mm), leading to this misclassification. Conversely, the second misclassification occurs in near-multi-damage scenarios, where the defective wheel is located on the second wagon with a flat length of 40 mm. Therefore, it can be mentioned that the proposed unsupervised methodology is capable of classifying damage severities into three classes namely, low, moderate, and severe defects, regardless of the number of defective wheels in each train passage with acceptable accuracy using the fused features for each accelerometer separately.

In order to modify the clustering analysis and clarify the mentioned misclassifications which are train passages 20 and 46, including defective wheels with flat lengths = 27.5 and 40 mm respectively, the second step of the data fusion which is fusing the sensors is implemented. As shown in Figure 10(c), the damaged classification is performed without any misclassification for single-damage scenarios. Regarding multi-damage scenarios, clustering analyses come with one misclassification belonging to the near-multi-damage scenario when the defective wheel is located on the first wagon with a flat length equal to 15 mm. This defective wheel belongs to the train passage including 2 defective wheels, 1 low damage on the first and 1 severe damage on the second wagons.

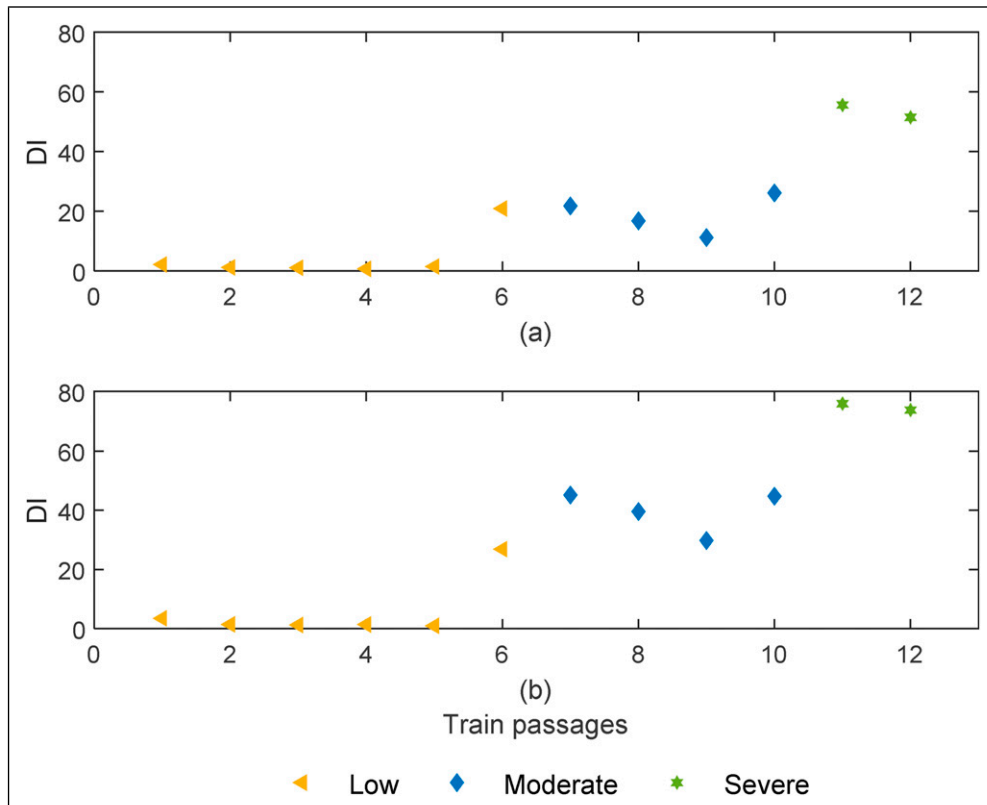


Figure 12. Automatic classification of wheel flat damage using accelerometer data for the first wheel of the first wagon: (a) accelerometer 1, (b) accelerometer 5.

Therefore, it can be mentioned that the dynamic effect of the defective wheel located on the second wagon ($L = 75$ mm) increased the peak of the acceleration of the damaged wheel on the first vehicle. It is possible to conclude that this defect is classified as moderate damage when it occurs near the severe defect. As described in Figure 10(c), the proposed algorithm is able to classify damaged wheels with low, moderate and severe intensity by yellow, blue and green indicators respectively, considering both single and multi-damage scenarios with admissible accuracy.

Figure 11(a) and (b) represents automatic wheel flat damage classification using the dynamic responses from accelerometers 1 and 5 considering 30 simulated damaged wheels on the third wagon, as shown in Figure 11(a). The characteristics of the wheel flat are detailed in Table 1. The passages from 1 to 10 represent defects with low damages and indicators from 11 to 20 describe moderate defects, and the remaining train passages correspond to severe defects.

As shown in this figure, damage classification comes with only one misclassification with flat length equal to 27.5 mm, which belongs to the last train passage with moderate damage. As explained in the previous section, this value closely approaches the threshold for low damage severity (10–20 mm), resulting in misclassification. To adjust the clustering analysis and address the noted

misclassification (from train passage 20 including damaged wheel with flat length = 27.5 mm), the second phase of data fusion is implemented, which involves integrating features from all sensors. As depicted in Figure 11(c), the damage classification was executed without any misclassification and the result of the clustering analysis is robust.

Figure 12 represents the classification of wheel flat damage for the first wheel of the first wagon considering accelerometers 1 and 5, which includes 6 low damage severity and 4 moderate defects, and 2 severe damages. As shown in this figure, the damage classification is performed accurately without any misclassification, and the first phase of the data fusion, merging the features for each sensor separately, is enough to increase the sensitivity to the damage.

Figure 13 illustrates the classification of wheel flat damage for the second wheel of the second wagon considering accelerometers 1 and 5, which includes 1 low damage severity and 2 moderate defects and 3 severe damage. As depicted in Figure 13(a) and (b), damage classification comes with 1 misclassification corresponding to flat length equal to 40 mm. To address this misclassification, sensor fusion is implemented to merge all accelerometers. As shown in Figure 13(c), sensor fusion increases the sensitivity to the damage and the clustering analysis comes with reliable performance while the defect is located on the second wagon.

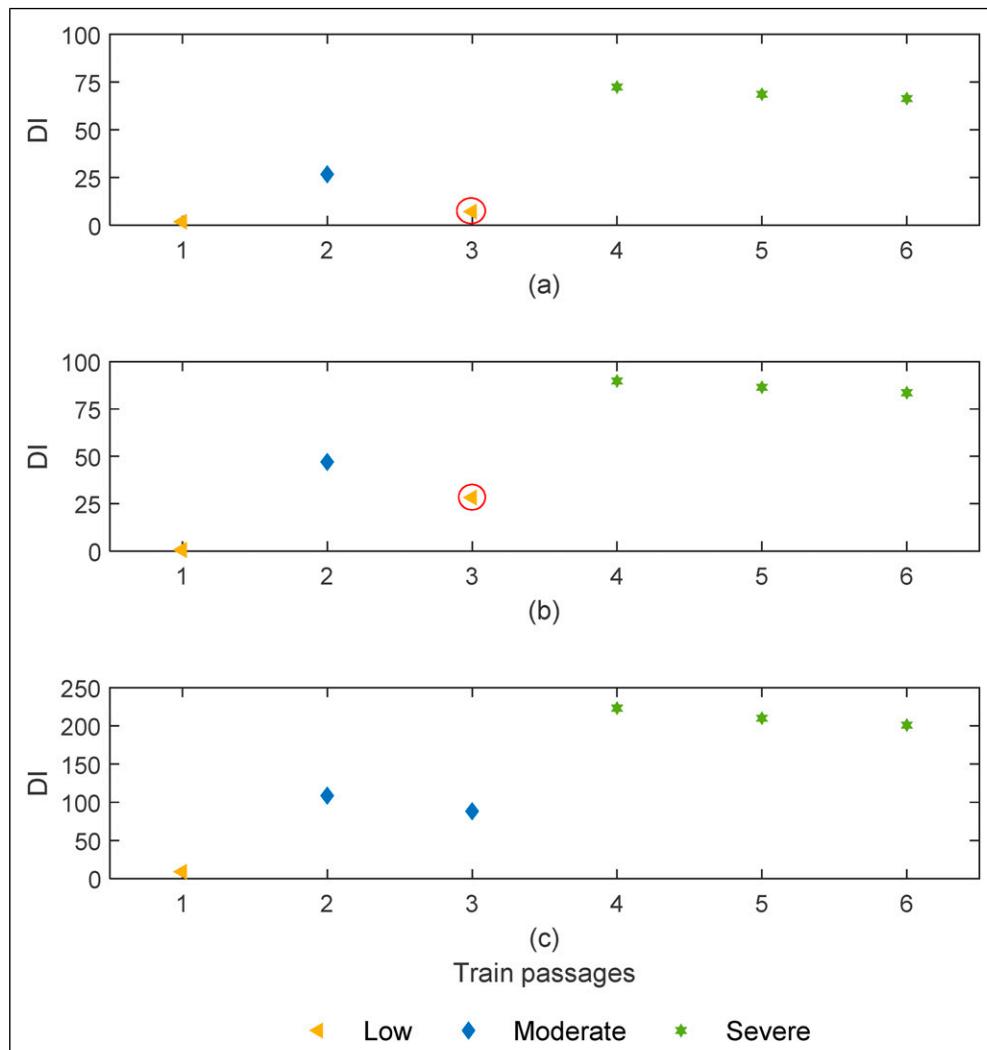


Figure 13. Automatic classification of wheel flat damage using accelerometer data for the second wheel of the second wagon: (a) accelerometer 1, (b) accelerometer 5, (c) all accelerometers.

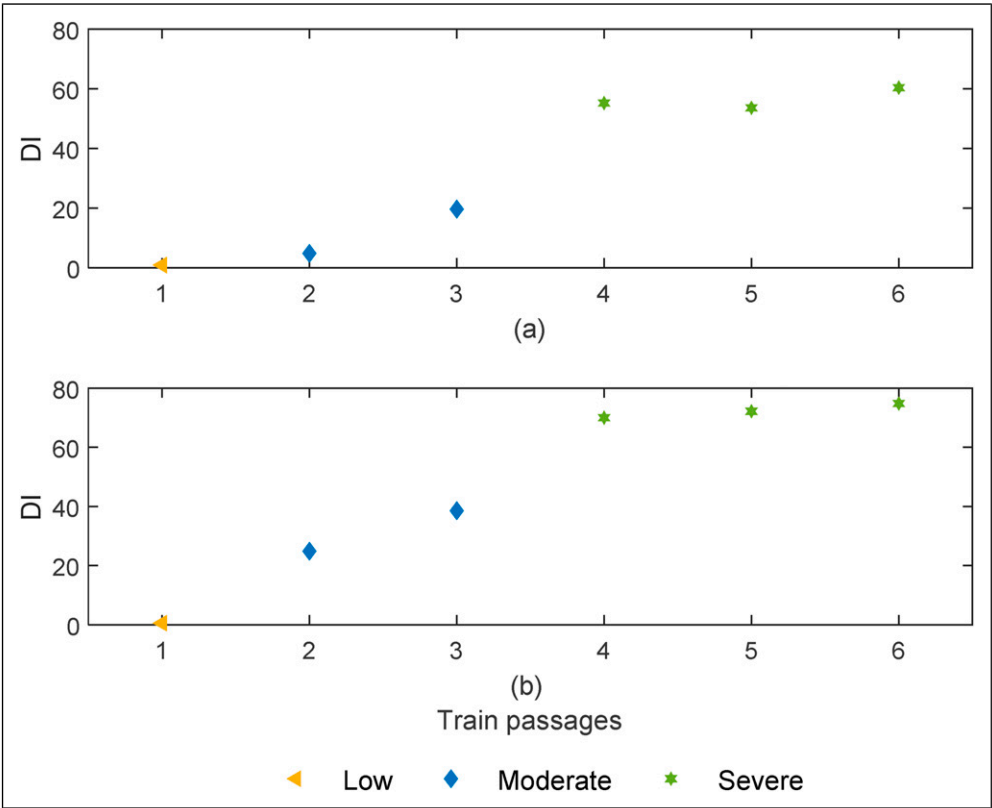


Figure 14. Automatic classification of wheel flat damage using accelerometer data for the second wheel of the 5th wagon: (a) accelerometer 1, (b) accelerometer 5.

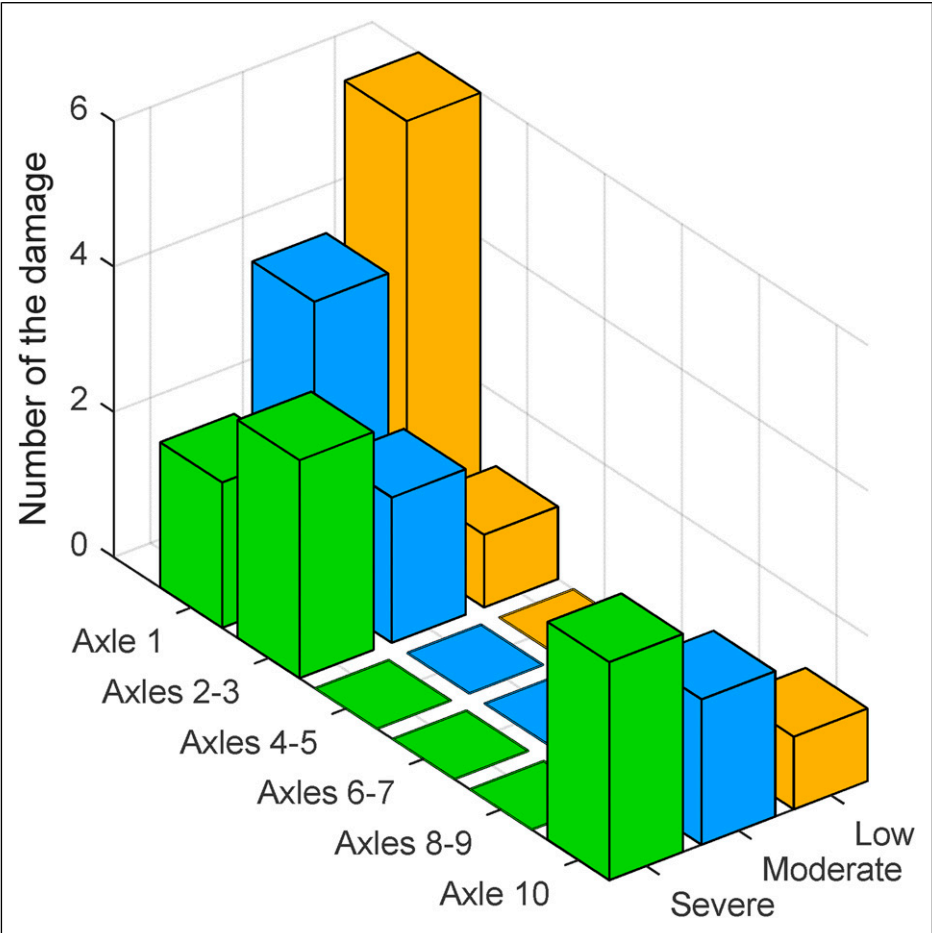


Figure 15. The output of the current study using proposed unsupervised methodology to identify and localize wheel flat considering train passages with multi-defective wheels.

Figure 14 depicts the classification of wheel flat damage for the second wheel of the 5th wagon, incorporating data from accelerometers 1 and 5. This includes one low damage severity, two moderate defects, and three cases of severe damage. As shown in this figure, all defective scenarios that are defined for the damaged wheel on the last wheel are classified appropriately, considering only merging the features for each sensor individually.

Figure 15 summarizes the output of the current study for wheel flat detection, damage localization, and defect severity classification considering single and multi-damage scenarios. As represented in Figure 1(b) and (c), two layouts are considered to simulate scenarios with multi-damaged wheels. Additionally, different damage severities are considered to generate wheel flats for these cases (Table 1). As depicted in Figure 15, the proposed unsupervised methodology identifies 12 defects on the first axle (6 low, 4 moderate, and 2 severe damages), while more 6 damages are identified on both axles 2–3 and axle 10 (1 low, 2 moderate, and 3 severe damages). Therefore, it can be inferred that the proposed methodology is capable of detecting, localizing, and classifying the wheel flat regardless of the number of defective wheels and the position of the damages.

Conclusions

This research study presents an unsupervised discrimination algorithm for detecting, localizing and classifying train wheel flat based on damage severity, in three stages: (i) wheel flat detection using the establishment of confidence boundary; (ii) damage localization based on an automatic segmentation technique; (iii) classification of wheel damage based on its severity. The unsupervised learning algorithm is validated with artificial data attained from a virtual wayside monitoring system related to freight train passages with healthy wheels and defective wheels with single and multiple defects. The proposed methodology provides efficiency and robustness for damage identification and localization regardless of the number of defective wheels and their position. As a result of the research study presented herein, the following conclusions can be drawn:

- the proposed methodology is capable of wheel flat identification based on the damage severity;
- defective and undamaged wheels can be distinguished by outlier analysis based on the confidence boundary at an early stage;
- defective wheels can be localized by segmented acceleration signals corresponding to single axles (first and last ones) and pair (combination of front and rear wheels of wagons in succession) axles individually.
- the feature set can be analyzed and separated according to the wheel conditions in a completely automated manner by using *k*-means;
- the developed methodology is effective in classifying the wheel flat severity for each axle individually;
- the proposed methodology distinguishes train passages including multiple from single-defective wheels and defines the position of the defect in the damage localization stage;

- damage detection comes with appropriate results using only one sensor even when train passages include multi-defective wheels;
- the proposed methodology can classify damaged wheels based on their severity (low, moderate, and severe) regardless of whether the train has single or multiple damaged wheels, even when the defective wheels are situated in either distant or adjacent wagons;
- it can be concluded that when there is a lag between the responses evaluated from different axles, it indicates the detection of a flat. Conversely, when the responses evaluated from different axles are very similar (no lag), it indicates the absence of flats;
- further development is intended to be done to assess the performance of proposed methodology to identify multiple wheel flats on single defective wheel.

Results such as these clearly demonstrate that this innovative application of data mining in the railway industry has great potential, especially for infrastructure managers. Using on-site measurements, the proposed methodology will be validated in a future field trial. Additionally, it is considered to improve the segmentation technique to obtain the signal for each wheel individually. Moreover, proposing an indicator to perform damage localization automatically is one of the innovative phases for future work.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was financially supported by Base Funding-UIDB/04708/2020 and Programmatic Funding-UIDP/04708/2020 of the CONSTRUCT - Instituto de Estruturas e Construções, funded by national funds through the FCT/ MCTES (PIDDAC). The second author acknowledges Grant no. 2021.04272.CEECIND from the Stimulus of Scientific Employment, Individual Support (CEE-CIND) - 4rd Edition provided by "FCT – Fundação para a Ciência, DOI: 10.54499/2021.04272.CEECIND/CP1679/CT0003". This work is a result of Agenda "SMART WAGONS – Development of Production Capacity in Portugal of Smart Wagons for Freight", nr. C644940527-00000048, investment project nr. 27, financed by the Recovery and Resilience Plan (PRR) and by European Union - NextGeneration EU".

ORCID iDs

Mohammadreza Mohammadi  <https://orcid.org/0000-0003-1079-4687>
 Cecilia Vale  <https://orcid.org/0000-0003-2470-9834>
 Diogo Ribeiro  <https://orcid.org/0000-0001-8624-9904>
 Pedro Montenegro  <https://orcid.org/0000-0001-5699-4428>

References

1. Mosleh A, Meixedo A, Ribeiro D, et al. Automatic clustering-based approach for train wheels condition monitoring. *Int J Real Ther* 2023; 11(5): 639–664.

2. Mosleh A, Meixedo A, Ribeiro D, et al. Early wheel flat detection: an automatic data-driven wavelet-based approach for railways. *Veh Syst Dyn* 2023; 61(6): 1644–1673.
3. Mosleh A, Montenegro P, Alves Costa P, et al. An approach for wheel flat detection of railway train wheels using envelope spectrum analysis. *Structure and Infrastructure Engineering* 2021; 17(12): 1710–1729.
4. European Commission. Directorate General for Mobility and Transport. *White paper on transport: roadmap to a single European transport area: towards a competitive and resource-efficient transport system*. Luxembourg City, Luxembourg: Publications Office of the European Union, 2011.
5. Vale C. Wheel flats in the dynamic behavior of ballasted and slab railway tracks. *Appl Sci* 2021; 11(15): 7127.
6. Rail Industry Safety and Standards Board. *Code of practice for wheel defects*. Spring Hill, QLD: Rail Industry Safety and Standards Board, 2020.
7. Barke DW and Chiu WK. A review of the effects of out-of-round wheels on track and vehicle components. *Proc Inst Mech Eng F J Rail Rapid Transit* 2005; 219(3): 151–175.
8. Johansson A and Nielsen JC. Out-of-round railway wheels—wheel-rail contact forces and track response derived from field tests and numerical simulations. *Proc Inst Mech Eng F J Rail Rapid Transit* 2003; 217(2): 135–146.
9. Mohammadi M, Mosleh A, Vale C, et al. An unsupervised learning approach for wayside train wheel flat detection. *Sensors* 2023; 23(4): 1910.
10. UIC. *Code of practice for the loading and securing of goods on railway wagons*. Paris, France: UIC, 2022.
11. Alemi A, Corman F and Lodewijks G. Condition monitoring approaches for the detection of railway wheel defects. *Proc Inst Mech Eng F J Rail Rapid Transit* 2017; 231(8): 961–981.
12. Bosso N, Gugliotta A and Zampieri N. Wheel flat detection algorithm for onboard diagnostic. *Measurement* 2018; 123: 193–202.
13. Filograno ML, Corredera P, Rodriguez-Plaza M, et al. Wheel flat detection in high-speed railway systems using fiber Bragg gratings. *IEEE Sensor J* 2013; 13(12): 4808–4816.
14. Mishra S, Sharan P and Saara K. Real time implementation of fiber Bragg grating sensor in monitoring flat wheel detection for railways. *Eng Fail Anal* 2022; 138: 106376.
15. Aktas M, Gunel EH, Yilmazer P, et al. Detection of wheel flatten defect on the moving train with acoustic emission sensor. In: 2018 IEEE 29th annual international symposium on personal, indoor and mobile radio communications (PIMRC), Bologna, Italy, 9–12 September 2018. IEEE.
16. Amini A, Entezami M, Huang Z, et al. Wayside detection of faults in railway axle bearings using time spectral kurtosis analysis on high-frequency acoustic emission signals. *Adv Mech Eng* 2016; 8(11): 1687814016676000.
17. Mosleh A, Montenegro PA, Costa PA, et al. Railway vehicle wheel flat detection with multiple records using spectral kurtosis analysis. *Appl Sci* 2021; 11(9): 4002.
18. Sun Z, Ye X-W and Lu J. Estimating stay cable vibration under typhoon with an explainable ensemble learning model. *Struc and Infrastructure Eng* 2023; 20: 1912–1924.
19. Gonçalves V, Mosleh A, Vale C, et al. Wheel out-of-roundness detection using an envelope spectrum analysis. *Sensors* 2023; 23(4): 2138.
20. Chen S, Wang K, Chang C, et al. A two-level adaptive chirp mode decomposition method for the railway wheel flat detection under variable-speed conditions. *J Sound Vib* 2021; 498: 115963.
21. Nowakowski T, Komorski P, Szymański GM, et al. Wheel-flat detection on trams using envelope analysis with Hilbert transform. *Lat Am J Solid Struct* 2019; 16: 78255010.
22. Wang R, Crosbee D, Beven A, et al. Vibration-based detection of wheel flat on a high-speed train. In: *Advances in asset management and condition monitoring: comadem 2019*. Berlin, Germany: Springer, 2020, pp. 159–169.
23. Azimi M, Eslamlou AD and Pekcan G. Data-driven structural health monitoring and damage detection through deep learning: state-of-the-art review. *Sensors* 2020; 20(10): 2778.
24. Dernbach G, Lykartsis A, Sievers L, et al. Acoustic identification of flat spots on wheels using different machine learning techniques. *Fortschritte der Akustik - DAGA. 2020: 46. Deutsche Jahrestagung für Akustik*. Berlin, Germany: Deutsche Gesellschaft für Akustik, 2020, pp. 367–370.
25. Wan TH, Tsang CW, Hui K, et al. Anomaly detection of train wheels utilizing short-time Fourier transform and unsupervised learning algorithms. *Eng Appl Artif Intell* 2023; 122: 106037.
26. Shaikh K, Hussain I and Chowdhry BS. Wheel defect detection using a hybrid deep learning approach. *Sensors* 2023; 23(14): 6248.
27. Ni Y-Q and Zhang Q-H. A Bayesian machine learning approach for online detection of railway wheel defects using track-side monitoring. *Struct Health Monit* 2021; 20(4): 1536–1550.
28. Guedes A, Silva R, Ribeiro D, et al. Detection of wheel polygonization based on wayside monitoring and artificial intelligence. *Sensors* 2023; 23(4): 2188.
29. Mosleh A, Meixedo A, Ribeiro D, et al. Machine learning approach for wheel flat detection of railway train wheels. *Transport Res Procedia* 2023; 72: 4199–4206.
30. Mosleh A, Mohammadi M, Vale C, et al. Smart detection of wheel defects using artificial intelligence and wayside monitoring system. *International J of Railway Res* 2023; 10(2): 9–18.
31. Mosleh A, Meixedo A, Ribeiro D, et al. Intelligent clustering-based approach for railway wheel flat detection. In: *Proceedings of the fifth international conference on railway technology: research, development and maintenance*, Montpellier, France, 22–25 August 2022. Civil-Comp Press.
32. Montenegro PA and Calçada R. Wheel–rail contact model for railway vehicle–structure interaction applications: development and validation. *Rail Eng Science* 2023; 31(3): 181–206.
33. MATLAB®. *Version R2022a*. Natick, MA: The MathWorks Inc, 2022.
34. Hertz H. Ueber die Berührung fester elastischer Körper. *J Für Die Reine Und Angew Math* 1882; 92: 156–171.
35. Kalker JJ. *Book of tables for the herzian creep-force law*. Delft, The Netherlands: Delft University of Technology, Faculty of Technical Mathematics and Informatics, 1996.
36. ANSYS®. *ANSYS release 19.2*. Canonsburg, PA: Academic Research, 2018.
37. Meixedo A, Santos J, Ribeiro D, et al. Online unsupervised detection of structural changes using train-induced dynamic responses. *Mech Syst Signal Process* 2022; 165: 108268.
38. Kerschen G, Feeny BF and Golinval JC. On the exploitation of chaos to build reduced-order models. *Comput Methods Appl Mech Eng* 2003; 192(13): 1785–1795.
39. Han S and Feeny BF. Enhanced proper orthogonal decomposition for the modal analysis of homogeneous structures. *J Vib Control* 2002; 8: 19–40.

40. Lenaerts V, Kerschen G and Golinval JC. Identification of a continuous structure with a geometrical non-linearity. Part II: proper orthogonal decomposition. *J Sound Vib* 2003; 262(4): 907–919.
41. De Boe P and Golinval JC. Principal component analysis of a piezosensor array for damage localization. *Struct Health Monit* 2003; 2(2): 137–144.
42. Sousa Tomé E, Pimentel M and Figueiras J. Damage detection under environmental and operational effects using cointegration analysis – application to experimental data from a cable-stayed bridge. *Mech Syst Signal Process* 2020; 135: 106386.
43. Bull LA, Worden K, Fuentes R, et al. Outlier ensembles: a robust method for damage detection and unsupervised feature extraction from high-dimensional data. *J Sound Vib* 2019; 453: 126–150.
44. Figueiredo E, Park G, Farrar CR, et al. Machine learning algorithms for damage detection under operational and environmental variability. *Struct Health Monit* 2010; 10(6): 559–572.
45. Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE* 1989; 77(2): 257–286.
46. Bilmes JA. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *Intern Comp Sci Insti* 1998; 4(510): 126.
47. Stamp M. *A revealing introduction to hidden Markov models*. San Jose, CA: Department of Computer Science San Jose State University, 2004, pp. 26–56.
48. Skarlatos D, Karakasis K and Trochidis A. Railway wheel fault diagnosis using a fuzzy-logic method. *Appl Acoust* 2004; 65(10): 951–966.
49. Möller MF. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Network* 1993; 6(4): 525–533.
50. Zhai WM, Wang QC, Lu ZW, et al. Dynamic effects of vehicles on tracks in the case of raising train speeds. *Proc Inst Mech Eng F J Rail Rapid Transit* 2001; 215(2): 125–135.
51. Vale C and Calçada R. Dynamic response of a coupled vehicle-track system to real longitudinal rail profiles. In: Proceedings of the tenth international conference on computational structures technology (CST 2010), Valencia, Spain, 14–17 September 2010.
52. Vale C and Calçada R. A dynamic vehicle-track interaction model for predicting the track degradation process. *J Infrastruct Syst* 2014; 20(3): 04014016.
53. European Committee for Standardization. *Railway applications - track - Track geometry quality - part 2: measuring systems - track recording vehicles (EN 13848-2)*. Brussels, Belgium: European Committee for Standardization, 2006.
54. Mosleh A, Costa PA and 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* 2019; 234(2): 183–200.
55. Mosleh A, Meixedo A, Ribeiro D, et al. Feature extraction and normalization for wheel flat detection on railway vehicles. In: International conference on structural health monitoring of intelligent infrastructure: transferring research into practice, SHMII, Porto, Portugal, 30 June 2021–2 July 2021.
56. Lourenço A, Ferraz C, Ribeiro D, et al. Adaptive time series representation for out-of-round railway wheels fault diagnosis in wayside monitoring. *Eng Fail Anal* 2023; 152: 107433.