Proceedings of the Eighteenth International Conference on
Civil, Structural and Environmental Engineering Computing
Edited by: P. Iványi, J. Kruis and B.H.V. Topping
Civil-Comp Conferences, Volume 10, Paper 2.7
Civil-Comp Press, Edinburgh, United Kingdom, 2025
ISSN: 2753-3239, doi: 10.4203/ccc.10.2.7

A Data-Driven Methodology for Damage Detection in a Short-Span Filler-Beam Railway Bridge

A. Silva[1], A. Meixedo[1], P. A. Montenegro[1] and D. Ribeiro[1,2]

[1] CONSTRUCT-iRAIL, Faculty of Engineering, University of Porto, Portugal
[2] iBuilt, ISEP, Polytechnic of Porto, Portugal

Abstract

Ensuring the structural integrity of railway bridges is a vital concern in infrastructure management, particularly for short-span filler-beam bridges that are prone to degradation under repetitive loading. This work proposes a hybrid data-driven methodology to detect early-stage damage in such structures, using the Cascalheira bridge in Portugal as a case study. The approach integrates signal processing techniques (Continuous Wavelet Transform and Principal Component Analysis), deep learning (Sparse Autoencoders), and statistical tools (Mahalanobis distance and outlier analysis) to extract and refine damage-sensitive features from simulated acceleration responses. A comprehensive numerical model that accounts for train-bridge dynamic interactions and realistic track irregularities supports the simulation framework. Results demonstrate that the proposed method achieves reliable damage identification with a low false positive rate, even under significant environmental and operational noise. This robust and scalable strategy offers a promising advancement for indirect Structural Health Monitoring systems in railway infrastructure.

Keywords: structural health monitoring, damage detection, railway bridges, sparse autoencoders, deep learning, filler-beam bridge.

1 Introduction

A critical component of modern infrastructure management is to ensure the structural integrity of railway bridges. Among the various tasks in Structural Health Monitoring (SHM), early damage detection remains one of the most challenging and essential to prevent catastrophic failures and extend the service life of these bridges. This is particularly relevant for short-span filler-beam railway bridges, which are susceptible to degradation mechanisms due to their structural characteristics and frequent loading cycles.

Traditional SHM approaches, typically grounded in model-based methods, often rely on finite element calibration and modal parameter updating. However, these techniques are sensitive to uncertainties in modelling assumptions and material properties, and they are computationally intensive. Consequently, data-driven strategies, particularly those that use Machine Learning (ML) and Deep Learning (DL) algorithms, have emerged as promising alternatives. These methods can process large volumes of structural data and offer scalable and automated tools for damage identification.

Damage detection methods span a wide range of strategies. For example, Autoregressive (AR) modelling uses time-series analysis of strain or acceleration signals to identify changes associated with structural degradation [1]. Dynamic response analysis has been used to analyse traffic-induced dynamic responses using statistical techniques [2]. Vibration-based techniques that employ Convolutional Neural Networks (CNNs) have successfully distinguished damage patterns linked to cross-section losses due to corrosion [3]. Genetic algorithms have optimised residual minimisation between measured and estimated responses, enabling effective detection in complex structures [4]. Recent advances promote indirect, on-board detection through bogie acceleration measurements, enabling efficient localisation under variable operational conditions [5].

Despite these innovations, several challenges persist. One major issue is the masking of damage-sensitive features by operational and environmental variations, such as different train speeds, temperature changes, or track irregularities. These influences can introduce significant noise into raw sensor data, which obscures the subtle signatures of an early-stage damage. Moreover, the lack of labelled damage data in real-world contexts makes it necessary to use unsupervised or semi-supervised methods to distinguish structural anomalies without explicit examples of real events.

2 Damage detection methodology

Damage detection in railway bridges presents a significant challenge due to operational and environmental variability. To address this, a hybrid methodology combining signal processing, statistical analysis, and deep learning was developed to enhance the sensitivity and robustness of damage detection. This methodology is expanded through four sequential stages: feature extraction, feature updating, data fusion, and feature discrimination, as shown in Figure 1.

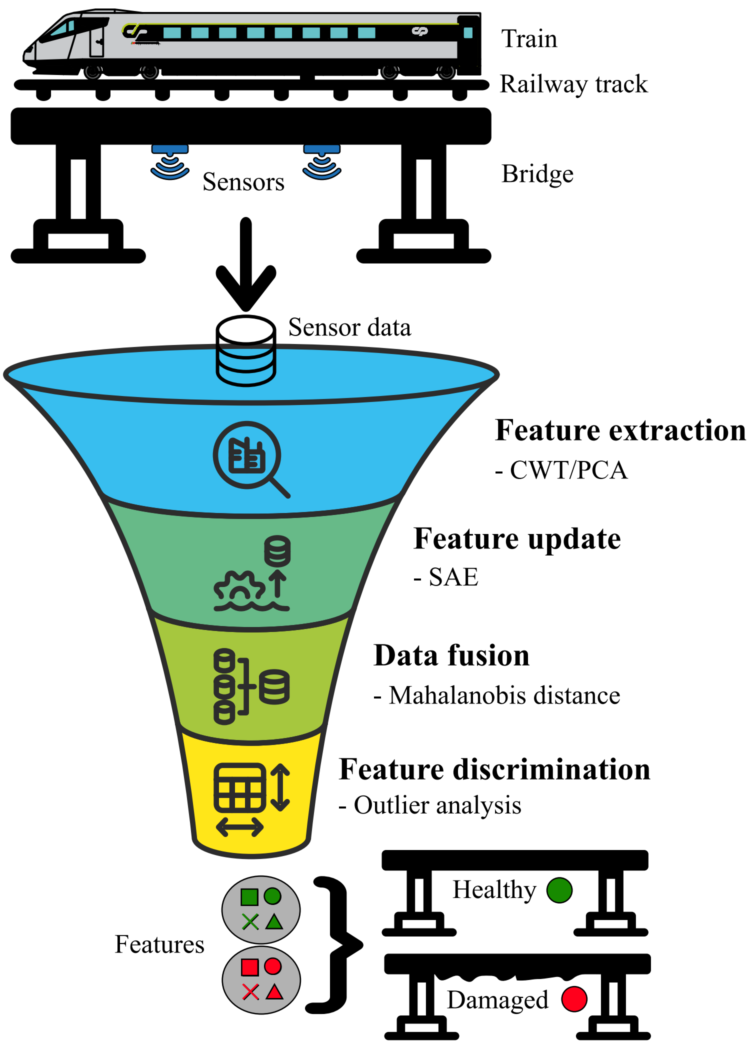


Figure 1: Methodology for damage detection used in this work.

2.1 Feature extraction

Extracting meaningful damage-sensitive features from acceleration signals is fundamental in unsupervised damage detection systems. In this work, the combination of the Continuous Wavelet Transform (CWT) and Principal Component Analysis (PCA) was employed to perform time-frequency decomposition and dimensionality reduction, respectively.

CWT is advantageous for analysing non-stationary signals because it can localise frequency content over time. The wavelet coefficients define it

$$W_{\psi}f\left(a,b\right)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}f\left(t\right)\psi^{*}\left(\frac{t-b}{a}\right)\mathrm{d}t,\quad \text{for } a>0 \tag{1}$$

where $\psi\left(t\right)$ is the mother wavelet, a the scale (inverse of frequency), and b the translation parameter. Morse wavelets in this application enable high adaptability due to their tunable symmetry and bandwidth characteristics, optimising the sensitivity to structural vibrations.

The resulting wavelet coefficients were post-processed using Principal Component Analysis (PCA), which transforms the high-dimensional feature space into a lower-dimensional set of uncorrelated principal components

$$Y=X\cdot T \tag{2}$$

where$X$ is the matrix of wavelet coefficients, T is the transformation matrix, and Y the reduced feature space. The first components retain most of the variance and the most relevant damage-related information. Statistical descriptors such as Root Mean Square (RMS), standard deviation, skewness, and kurtosis were then extracted from these components to represent the behaviour of each sensor compactly.

2.2 Feature updating

A Sparse Autoencoder (SAE) was trained on the PCA-compressed signals to refine the damage-sensitive characteristics of the features further. This deep learning model transforms the input data through a bottleneck layer, learning a sparse and nonlinear representation that amplifies anomalies while reducing the sensitivity to environmental noise.

The encoding-decoding process minimises the reconstruction loss $\mathcal{L}_{rec}\left(\Theta\right)$

$$\mathcal{L}_{rec}\left(\Theta\right)=\frac{1}{n}\sum_{i=1}^{n}\left\|x_i-\hat{x}_i\right\|^2 \tag{3}$$

where x is the input data, $\hat{x}$ is the reconstructed data, and n is the total number of data points. To enforce sparsity and regularisation, the total loss function becomes

$$\mathcal{L}_{total}=\mathcal{L}_{rec}+\beta\mathcal{L}_{sparse}+\alpha\mathcal{L}_{reg} \tag{4}$$

where $\mathcal{L}_{sparse}$ is the Kullback-Leibler divergence and $\mathcal{L}_{reg}$ the L2 regularisation term. The resulting encoded representations are highly compact and informative, which enhances the capacity of the classifier to differentiate subtle damage signatures.

2.3 Data fusion

After generating damage-sensitive features from multiple sensors, it is essential to consolidate this multivariate data into a single damage index. The Mahalanobis distance, MD_i, which accounts for correlations among variables, was employed to measure the deviation of the observed features from a reference healthy state

$$MD_i = (x_i - \bar{x}) \cdot S_x^{-1} \cdot (x_i - \bar{x})^T \tag{5}$$

where x_i is the feature vector of the current condition, $\bar{x}$ the mean of the healthy features, and S_x their covariance matrix. This metric provides a scalar damage indicator per simulation, which enables direct comparison across conditions. The Mahalanobis distance was computed and aggregated for each simulation and sensor to form a comprehensive diagnostic index.

2.4 Feature discrimination

An outlier detection method was implemented using a statistical confidence boundary (CB) to distinguish between healthy and potentially damaged states automatically. This boundary was defined via the Inverse Cumulative Distribution Function (ICDF) of the Gaussian distribution

$$CB = \text{inv}F\,(1 - \alpha) \tag{6}$$

where α is set to 0.005, corresponding to a 99.5% confidence level. Any Mahalanobis distance exceeding this boundary was considered indicative of structural damage.

3 Damage detection in the Cascalheira bridge

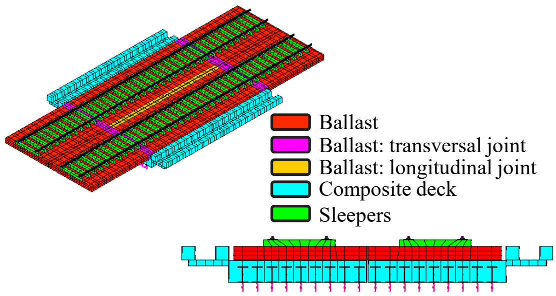
This work focuses exclusively on the damage detection task applied to the Cascalheira bridge, a short-span filler-beam railway bridge in Portugal. Through the implementation of a robust model that accounts for train-bridge dynamic interaction and advanced data-driven techniques, this work aims to identify early signs of damage, despite the variability induced by operational and environmental conditions.

3.1 Bridge and train model

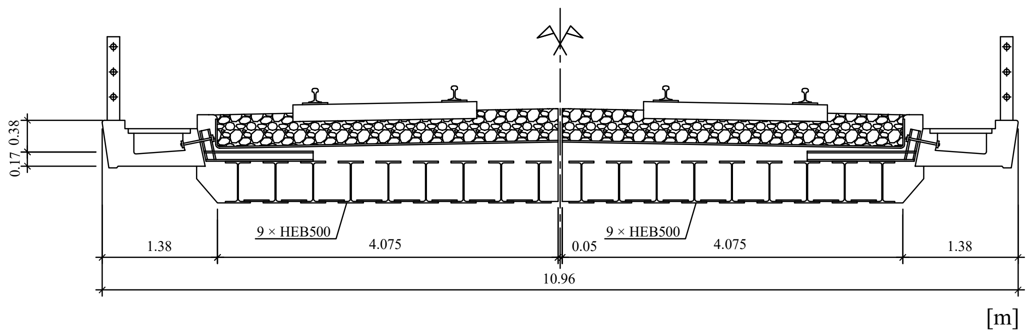
The Cascalheira bridge (Figure 2a) is located on the Northern Line of Portuguese Railways and supports trains travelling up to 160 km/h. It comprises two half-decks, each equipped with nine HEB500 steel beams embedded in a reinforced concrete slab (Figure 2c). A detailed 3D finite element model was developed using ANSYS software, which incorporates the composite deck, ballast, transition zones, and various material nonlinearities (Figure 2b). Shell, beam, solid, mass, spring, and rigid link elements were employed to represent each structural component with high fidelity. More details about the material properties of the model, its calibration and modal configurations can be found in [6].

The Alfa Pendular train (Figure 3a), with six vehicles and 24 axles, was modelled using a finite element approach. The vehicle model captured the dynamic effects through primary and secondary suspension systems and was validated against empirical data. Spring-dashpot elements replicated the directional stiffness and damping of the suspension system. A 3D numerical representation of one of the Alfa Pendular vehicles is shown in Figure 3b. More detailed information about the parameters of the Alfa Pendular vehicle can be found in [6].

(a)



(b)



(c)

Figure 2: Cascalheira bridge [6]: (a) global view; (b) numerical model; (c) cross-section (dimensions in meters).

3.2 Baseline and damage simulation scenarios

To support the detection phase, a finite element model of the bridge was calibrated and validated. A grid of 14 vertical accelerometers, placed along both the upper and lower sides of the deck, was defined to capture the dynamic response with spatial resolution suitable for modal sensitivity analysis (Figure 4).

A total of 96 time-history simulations were performed to represent baseline (undamaged) conditions (Figure 5). These simulations encompassed variations in train

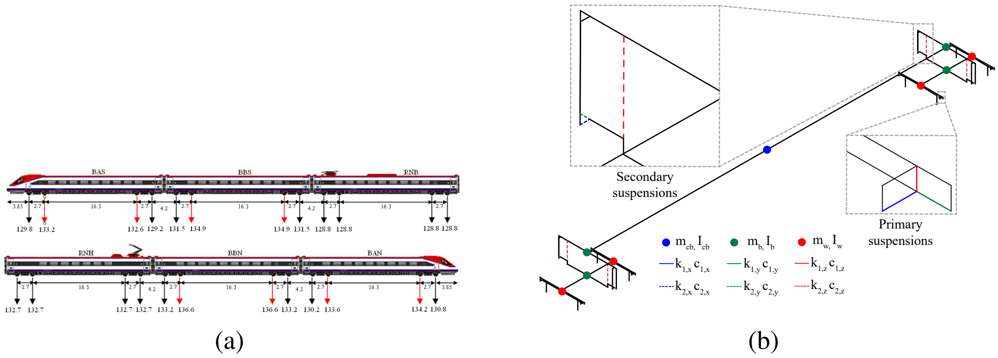


Figure 3: Alfa Pendular train [6]: (a) loading scheme (loads in kN); (b) numerical model.

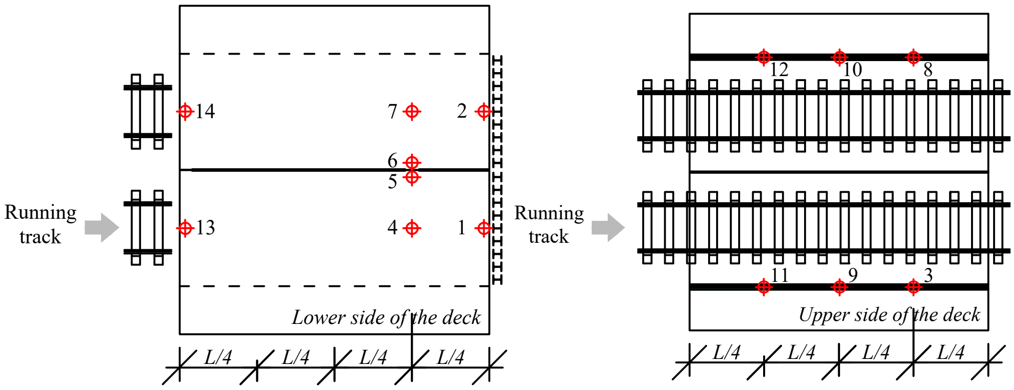


Figure 4: Numerical setup of the Cascalheira bridge.

speeds, train loading configurations, and realistic track irregularities, obtained from inspection data between 2018 and 2019. Three loading schemes and four irregularities profiles were combined to reflect operational diversity (Table 1). The simulations included the train-bridge dynamic interaction using the Vehicle Structure Interaction (VSI) tool developed by [7].

Scenario	**Information**
Loading scheme 1	Train with maximum passenger capacity and full tank levels
Loading scheme 2	Train with half passenger capacity and full tank levels
Loading scheme 3	Train without passengers and half tank levels
Irregularity 1	Simulation without irregularities profile
Irregularity 2	Irregularities profile measured on February 8th, 2018
Irregularity 3	Irregularities profile measured on September 2nd, 2019
Irregularity 4	Irregularities profile measured on October 6th, 2020

Table 1: Loading schemes and irregularities considered in the baseline scenarios.

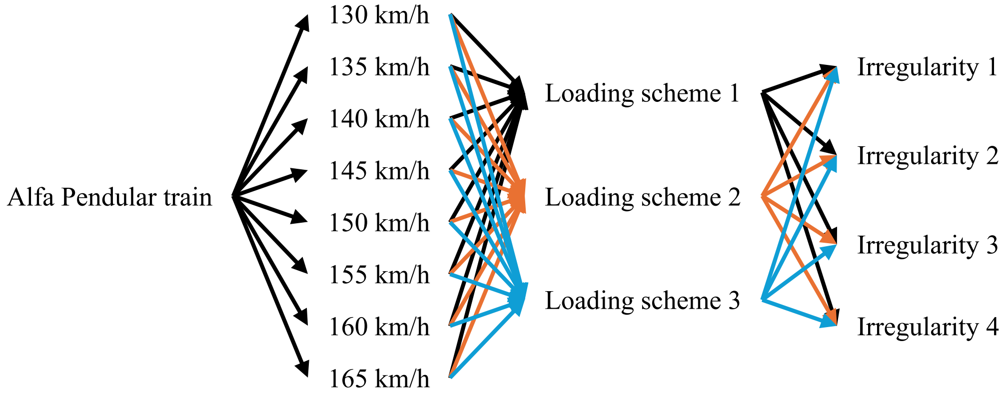


Figure 5: Combination of the 96 baseline simulations.

Damage scenarios were carefully selected to reflect plausible degradation mechanisms in filler-beam railway bridges, which include elastometric support deterioration and concrete cracking. Damage simulations were divided into two main categories (D1 and D2), with parameter variations informed by field observations and literature (Figure 6 and Table 2) [8]. For each category, multiple severity levels were applied. In total, 90 damage simulations were conducted, incorporating combinations of train speed, train loading conditions, and two irregularities profiles.

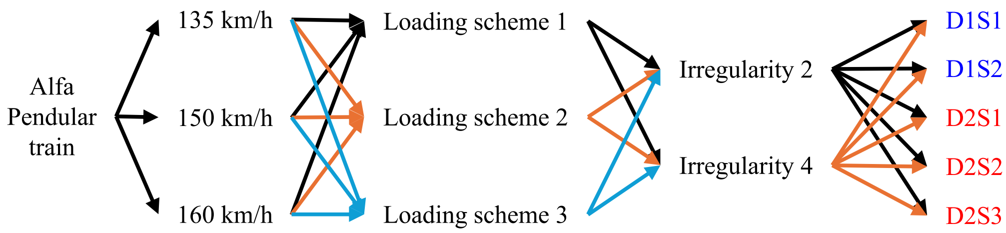


Figure 6: Combination of the 90 damage simulations.

To further enhance the fidelity of the simulation data, site-measured noise from the bridge without railway traffic was superimposed on the acceleration responses. This ensured a more realistic simulation of the signals captured by the accelerometers during actual monitoring campaigns.

3.3 Case study results

This section presents the application and performance evaluation of the damage detection methodology, structured in four sequential steps: feature extraction, feature updating, data fusion, and feature discrimination. The method was applied to the finite element model of the Cascalheira bridge subjected to realistic train-bridge dynamic interaction simulations, which includes operational and environmental variabil-

Damage	Parameter	Initial value	Variation (%)	Final value	Representation
Damage 1 Degradation of the supports	k_v (MN/m)	194.90	-21.70	152.60	D1S1
			-37.83	121.20	D1S2
Damage 2 Cracking of deck concrete	E_c (GPa)	36.39	-5	34.57	D2S1
			-10	32.75	D2S2
			-20	29.11	D2S3

Table 2: Details of all damage scenarios and their respective parameter variations.

ity through loading schemes, track irregularities, and train speeds.

Feature extraction

In the first stage, a hybrid signal processing approach which combines the CWT and PCA was employed to extract damage-sensitive features from vertical acceleration responses collected at 14 strategically placed sensors (Figure 4). Each time-series signal comprises 5813 data points and was decomposed into 96 wavelet coefficients using the Morlet wavelet, which was selected for its superior time-frequency localisation. The resulting 5813×96 matrix per sensor was compressed into a 4×96 feature matrix by calculating four statistical descriptors: RMS, standard deviation, skewness, and kurtosis. This resulted in 384 features per simulation per sensor, which captured dominant spectral and statistical characteristics of the structural response.

The effectiveness of the feature extraction process was validated by analysing the evolution of selected parameters across 186 structural scenarios (96 baseline and 90 damaged). Figure 7 presents three specific indices (155, 231, 308) from the total features, for one sensor (sensor 4 in Figure 4). The feature amplitudes showed sensitivity to train speed and track irregularities. However, the separation between damaged and undamaged conditions was not always distinct, emphasising the need for further refinement through nonlinear transformation.

Feature updating

The 384 extracted features were passed through an SAE to enhance damage detectability. This deep learning model was trained using 80% of undamaged simulations and a single sensor to determine optimal hyperparameters. The final configuration utilised a single hidden layer with 372 neurons, a sparsity proportion of 0.01, and regularisation parameters fine-tuned to promote generalisation. The bottleneck layer of the trained SAE generated a 372-dimensional feature vector for each simulation and sensor.

The autoencoder normalised the output features into the $[0, 1]$ interval, which simplifies the comparison and reduces the influence of outliers (Figure 8). The SAE thus

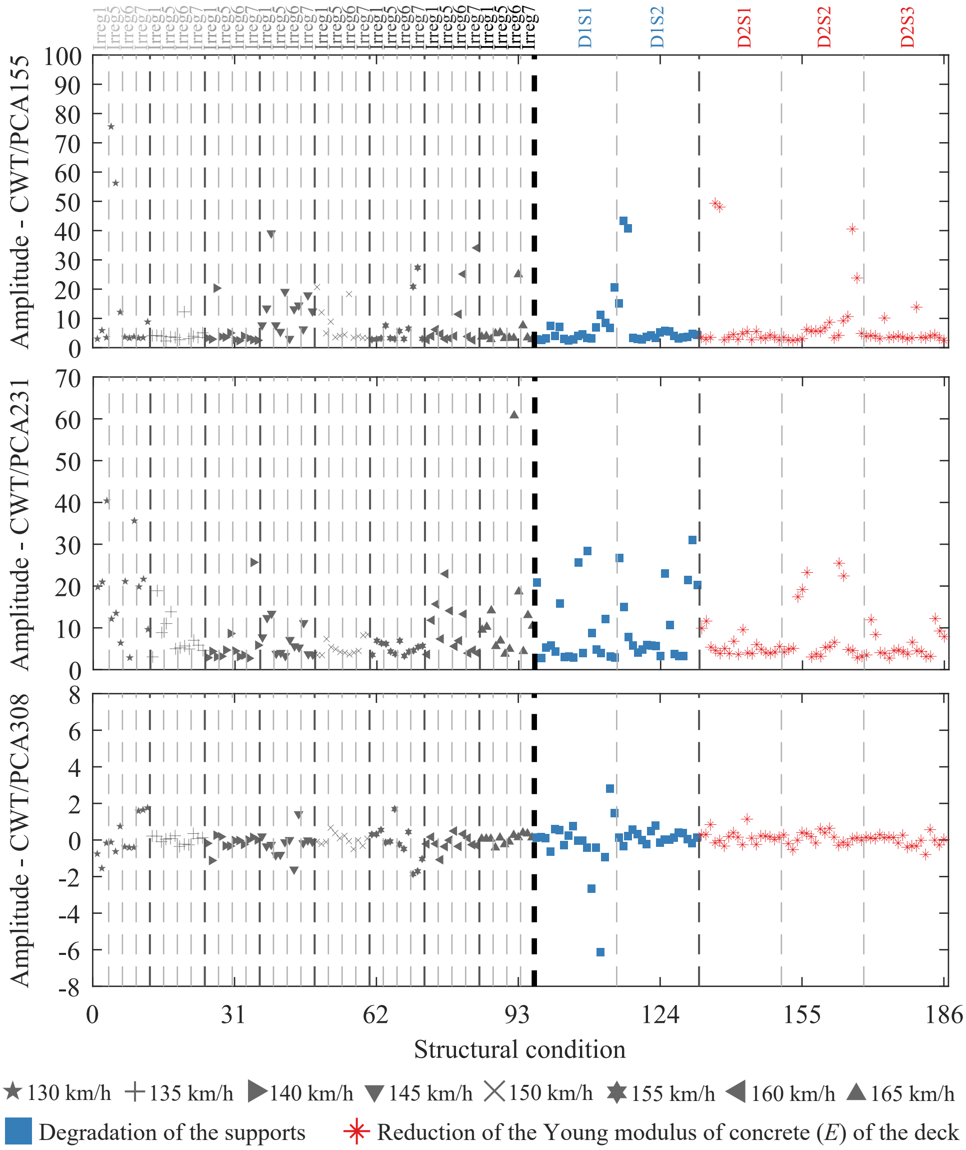


Figure 7: Amplitudes of three of the 384 CWT/PCA parameters, considering all the 186 structural conditions for the sensor 4.

acted as a nonlinear filter, removing redundant information and enhancing damage-sensitive representations of the structural state.

Data fusion

A multivariate data fusion strategy was implemented by computing the Mahalanobis distance between each simulation and the baseline reference. This computation was performed at the sensor level first, reducing the 372-dimensional SAE output into a single distance-based damage indicator (DI). A second fusion step integrated these indicators across the 14 sensors, which produced a unified DI vector for all simulations.

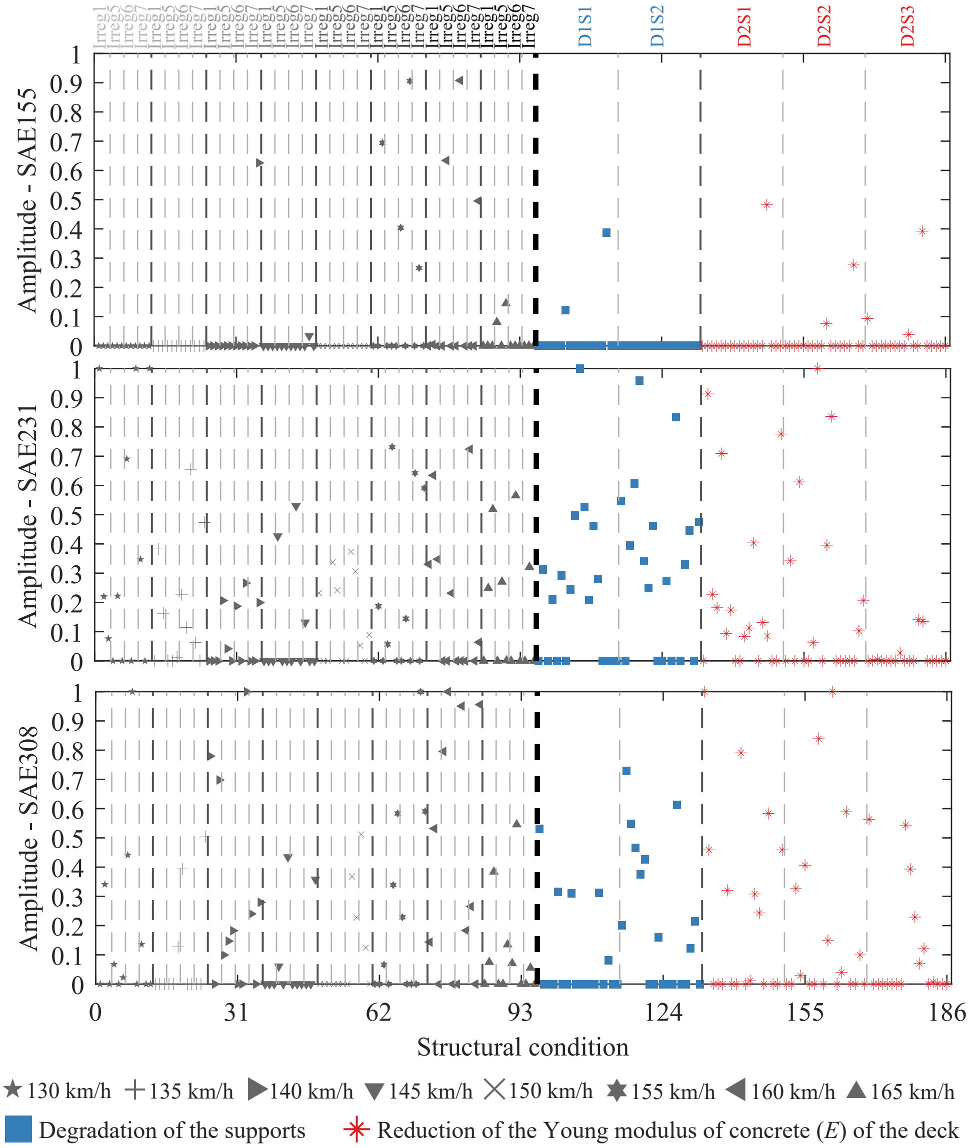


Figure 8: Amplitudes of three of the 372 SAE parameters, considering all the 186 structural conditions for the sensor 4.

This hierarchical data fusion successfully differentiated between healthy and damaged conditions, with damaged scenarios consistently showing higher DI values (Figure 9). The approach proved robust to operational variability, which captured structural anomalies even when signal variations due to speed or track irregularities were present.

Feature discrimination

A CB was established using the Gaussian ICDF at a significance level of 0.5% to automate damage detection (Figure 10). Any simulation with a DI that exceeds the CB was classified as damaged. Applying this threshold to the 186 simulations resulted in a clear binary separation between baseline and damage cases, with only 4.17% Type

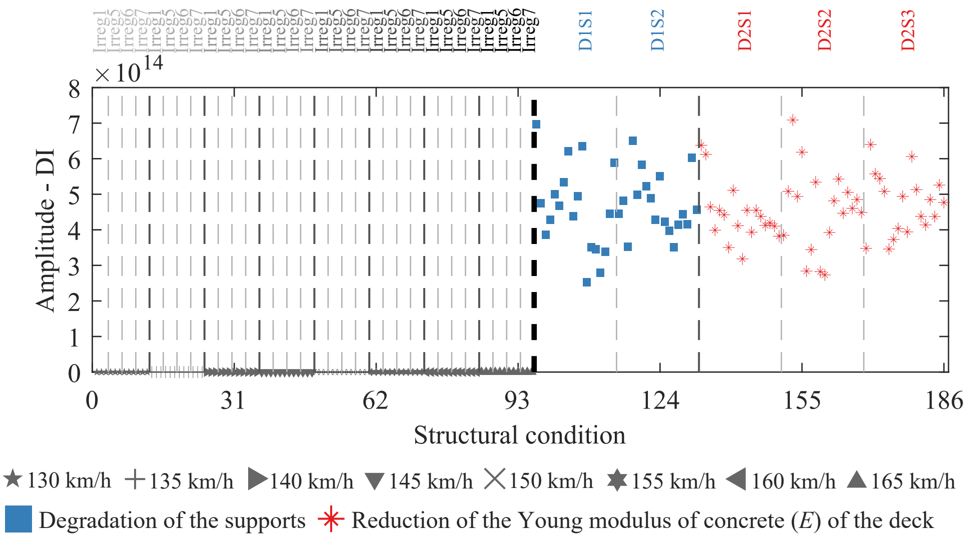


Figure 9: DI values obtained from SAE-based features, considering all the 186 structural conditions for all the 14 sensors.

I errors (false positives).

This outlier detection approach, grounded in statistical novelty detection, demonstrated strong performance under environmental and operational noise. It effectively leveraged the robustness of the Mahalanobis distance and the expressiveness of the SAE-generated features to detect early-stage damage scenarios.

4 Concluding remarks

This work presents a robust hybrid methodology for unsupervised damage detection in short-span railway bridges, validated through a detailed case study of the Cascalheira bridge. By combining advanced signal processing, deep learning, and statistical decision-making, the proposed approach effectively overcomes the challenges posed by operational and environmental variability. The integration of CWT, PCA, and SAEs enables the extraction and enhancement of damage-sensitive features from sensor data. The subsequent use of Mahalanobis distance and a confidence-based outlier analysis facilitates reliable differentiation between healthy and damaged states, with only 4.17% false positives reported. These results underscore the potential of the methodology for early damage detection in real-world monitoring systems, contributing to safer and more efficient infrastructure management. Future research may extend this framework to experimental datasets and explore its application to other bridge typologies or structural systems.

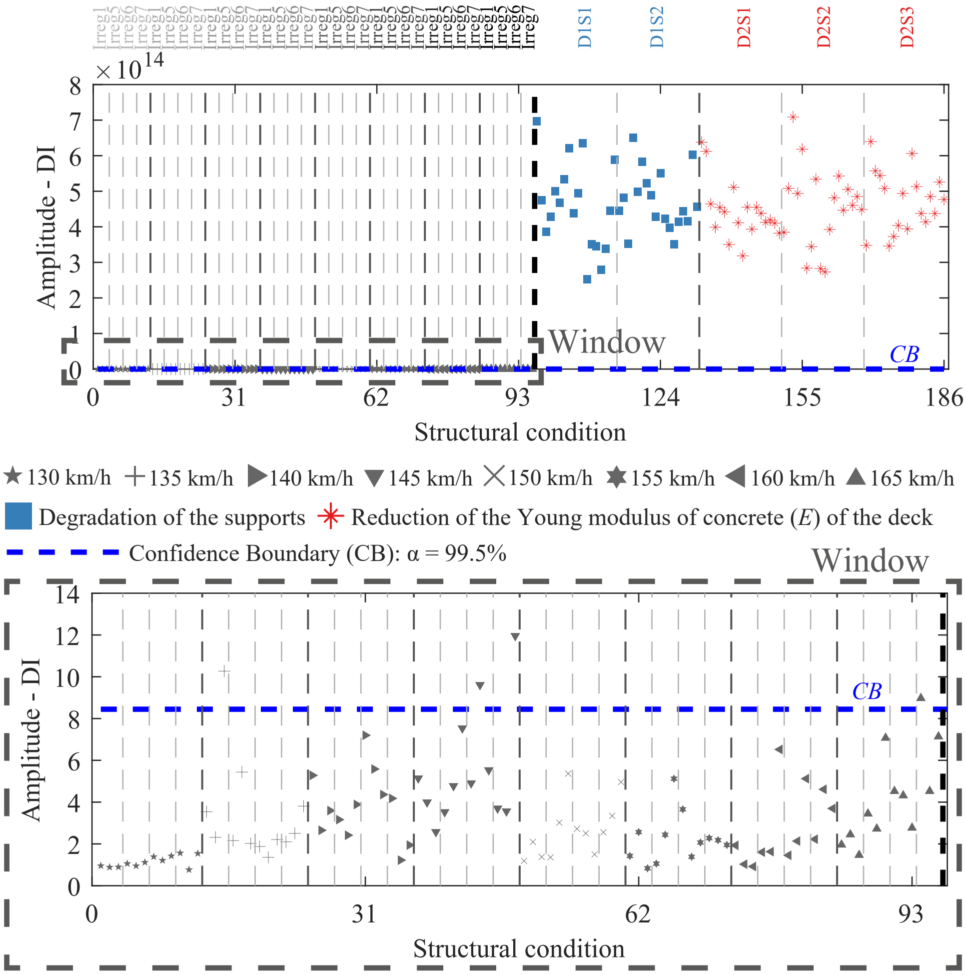


Figure 10: Automatic damage detection using DI values and a CB determined for a significance level of 0.5%, considering all the 186 structural conditions for all the 14 sensors.

Acknowledgements

This work was financially supported by Funding - UID/04708 of the CONSTRUCT - Instituto de I&D em Estruturas e Construções - funded by Fundação para a Ciência e a Tecnologia, I.P./ MCTES through the national funds. The first author acknowledges the project Operation PRR-C644940527-00000048 (SMART WAGONS) supported by Agência para a Competitividade e Inovação, I.P., under the PRR - Plano de Recuperação e Resiliência - Agendas Mobilizadoras, and the project Operation NORTE-08-5369-FSE-000027 supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Social Fund (ESF).

References

[1] S. Anastasia, E. García-Macías, F. Ubertini, V. Gattulli, and S. Ivorra, "Damage Identification of Railway Bridges through Temporal Autoregressive Modeling," *Sensors*, vol. 23, no. 21, p. 8830, Jan. 2023, number: 21 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/1424-8220/23/21/8830

[2] A. Meixedo, J. Santos, D. Ribeiro, R. Calçada, and M. Todd, "Damage detection in railway bridges using traffic-induced dynamic responses," *Engineering Structures*, vol. 238, p. 112189, Jul. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141029621003394

[3] A. Ghiasi, M. K. Moghaddam, C.-T. Ng, A. H. Sheikh, and J. Q. Shi, "Damage classification of in-service steel railway bridges using a novel vibration-based convolutional neural network," *Engineering Structures*, vol. 264, p. 114474, Aug. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S014102962200582X

[4] V. N. Alves, M. M. de Oliveira, D. Ribeiro, R. Calçada, and A. Cury, "Model-based damage identification of railway bridges using genetic algorithms," *Engineering Failure Analysis*, vol. 118, p. 104845, Dec. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1350630720306580

[5] J. Shi, H. Shi, Z. Wu, and J. Li, "An On-Board Damage Detection Method for Heavy-Haul Railway Bridge Based on Sensitivity Analysis of Bogie Responses," *IEEE Sensors Journal*, vol. 24, no. 4, pp. 4642–4655, Feb. 2024, conference Name: IEEE Sensors Journal. [Online]. Available: https://ieeexplore.ieee.org/document/10388063

[6] G. Saramago, P. A. Montenegro, D. Ribeiro, A. Silva, S. Santos, and R. Calçada, "Experimental Validation of a Double-Deck Track-Bridge System under Railway Traffic," *Sustainability*, vol. 14, no. 10, p. 5794, Jan. 2022, number: 10 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2071-1050/14/10/5794

[7] P. A. Montenegro, S. G. M. Neves, R. Calçada, M. Tanabe, and M. Sogabe, "Wheel–rail contact formulation for analyzing the lateral train–structure dynamic interaction," *Computers & Structures*, vol. 152, pp. 200–214, May 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045794915000139

[8] J. Manterola, *Puentes: apuntes para su diseño, cálculo y construcción*. Madrid, Spain: Colegio de Ingenieros de Caminos, Canales y Puertos, 2006, google-Books-ID: ASxQPwAACAAJ.