Proceedings of the Seventh International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 15, Paper 7.7
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.15.7.7

Data-Driven Models for the Classification and Estimation of Railway Track Irregularities Based on Vehicle Dynamics

M. Nabais, J. Torres, J. Pagaimo, S. Vieira and H. Magalhaes

IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Abstract

Railway track geometry maintenance is crucial for ensuring the comfort, safety, and efficiency of railway operations. Traditional inspection systems, however, rely on costly specialized vehicles, which limits inspection frequency and delays maintenance interventions. To overcome these limitations, this study presents a data-driven methodology that utilizes vehicle dynamic responses to classify and estimate track irregularities. A high-fidelity multibody (MB) model of the EM120 inspection vehicle is developed to simulate realistic operating conditions and map the correlations between vehicle behavior and track defects. A comprehensive dataset, generated from track segments with known irregularities, is used to train and evaluate various Artificial Intelligence (AI) models. The results demonstrate a strong correlation between track irregularities and bogie frame dynamics, enabling highly accurate classification of track sections. This research confirms the feasibility of continuous track condition monitoring using a simple sensor module mounted on the bogie frame, offering a scalable and cost-effective strategy for predictive railway maintenance.

Keywords: railway track geometry, condition monitoring, vehicle dynamics, multibody simulation, machine learning, convolutional neural networks.

1 Introduction

Railway track geometry monitoring is essential for safe, comfortable and reliable railway operations. Track irregularities correspond to deviations of the rails from their intended position. Their assessment is traditionally carried out by dedicated inspection vehicles, such as the EM120 used in the Portuguese railway network, which provide accurate measurements, but are costly and operate periodically, allowing new or evolving defects to remain undetected between inspections, which in worst case scenarios can lead to serious safety risks such as derailments [1, 2]. For this reason, there is increasing interest in transitioning from periodic inspection strategies to continuous track condition monitoring approaches.

Recent research has explored the use of onboard vehicle dynamic measurements as an alternative or complement to traditional track inspection systems [2, 3]. Studies have shown that vehicle accelerations and angular velocities contain information related to track geometry condition [2]. Model-based approaches, such as Kalman filtering, and data-driven methods, such as machine learning algorithms, have been used to classify and estimate irregularities from vehicle response [4-7]. However, for profile reconstruction in particular, further work is needed to extend these approaches to other relevant geometry parameters [5, 6].

This study proposes a data-driven methodology for continuous track condition monitoring based on simulated bogie-frame dynamic responses. An EM120 multibody model is developed and used to generate synthetic datasets for two approaches: SVM classifiers that identify track sections exceeding maintenance thresholds and a multi-scale dilated CNN that reconstructs continuous CL and LL profiles. The CNN combines information from different wavelength ranges, enabling separate analysis of D1, D2, and D3 band reconstruction. By integrating severity classification and profile reconstruction within the same simulation framework, this study evaluates the potential of an AI-based monitoring approach using compact onboard sensors for more frequent and cost-effective track condition assessment.

2 Methods

Track irregularities represent deviations of the rail from its design position and can affect both railway safety and passenger comfort. In regard to these irregularities the wavelength ranges considered in this study were: D1 (3 m $< \lambda \leq$ 25 m); D2 (25 m $< \lambda \leq$ 70 m); and D3 (70 m $< \lambda \leq$ 150 m). Shorter wavelengths are mainly associated with safety-related risks, while longer wavelengths have a stronger influence on ride comfort. Track geometry is described through four parameters represented in Figure 1: alignment level (AL), corresponding to the lateral displacement of the rails; longitudinal level (LL), corresponding to vertical rail displacement; gauge (G), corresponding to deviations in the distance between the two rails; and cross-level (CL), corresponding to the height difference between rails.

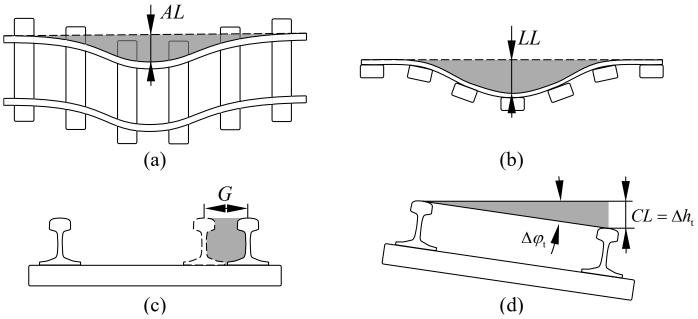

Figure 1: Schematic illustration of (a) Alignment Level, (b) Longitudinal Level, (c) Gauge, (d) Cross Level irregularities.

As Machine Learning (ML) models require a large amount of data to learn relationships and the available experimental data obtained from the EM120 is not sufficient for this purpose, a synthetic dataset is generated through multibody simulations, following an approach adopted in previous studies [4,8,9]. Irregularity profiles are first generated using the German Track Spectrum and described through their Power Spectral Density (PSD), which represents the distribution of signal energy as a function of frequency.

The resulting irregularity profiles are then used as input in MB simulations, allowing the corresponding vehicle dynamic response to be obtained. The simulations are performed using MUBODyn, which is an in-house MB dynamics code developed to handle a wide range of mechanical systems [10] including railway applications [11], [12]. In an MB formulation, the vehicle is represented by rigid and/or flexible bodies connected by kinematic joints and force elements, and the equations of motion are solved over time considering external forces, prescribed motions, constraints and wheel-rail contact interactions. This allows the computation of the vehicle motion, accelerations, constraint reactions and wheel-rail contact forces.

The developed simulated vehicle is the EM120, as illustrated in Figure 2 (a), the model includes the main components relevant to the vehicle dynamic response, namely the primary suspension, secondary suspension and the bogie-carbody connection. The bogie components and suspension systems are shown in greater detail in Figure 2 (b).

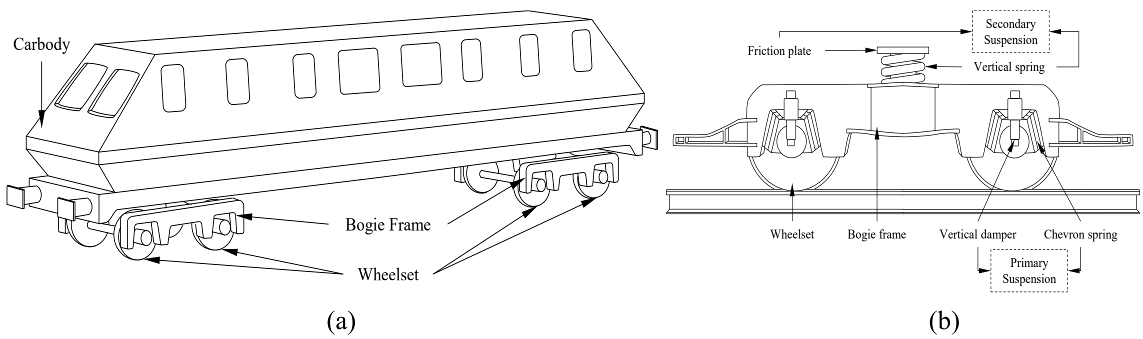

Figure 2: (a) Three-dimensional model of the EM120 inspection vehicle; (b) Bogie components and suspensions

The EM120 MB model is assessed through stability, curving, and experimental comparisons, supporting its suitability for synthetic data generation. The resulting generated datasets contain AL, LL, CL, and G track profiles computed with different standard deviation values and their corresponding vehicle responses, to represent different severity conditions. The EM120 is simulated at 100 km/h over tracks containing six binary irregularity combinations. Moreover, the dynamic signals are extracted at the front bogie frame, representing measurements acquired at the most suitable location identified after conducting a sensor position study.

For the classification task, each 200 m track section is assigned to one of two classes based on the SD of the corresponding track irregularity profile and a predefined threshold: Class 1 represents sections with minor irregularities that do not require maintenance, while Class 2 represents sections exceeding the threshold and requiring maintenance scheduling. The input features are obtained from the SD of the D1-filtered bogie-frame motion, including accelerations and angular velocities.

Support Vector Machine (SVM) models are developed to perform this classification task. SVMs are selected due to their suitability for classification problems and their ability to define decision boundaries using a subset of relevant training samples, known as support vectors [13]. First, a general binary classification experiment is conducted to compare different SVM kernels: linear, quadratic, cubic, and Gaussian. The best-performing kernel is then used to train four specialized binary classifiers, one for each irregularity type: AL, LL, CL, and G. Each specialized classifier estimates whether its corresponding irregularity belongs to Class 1 or Class 2. Before training, the input features are standardized to zero mean and unit variance, and Bayesian optimization is used for hyperparameter tuning. Model training is performed using 10-fold cross-validation. Performance was assessed using accuracy, which measures the overall proportion of correct classifications, sensitivity, which measures the ability to identify sections requiring maintenance, the F1 score, which balances precision and sensitivity, and Cohen's Kappa, which measures agreement between predicted and true classes while accounting for chance agreement.

For the regression task, a multi-scale dilated model based on Convolutional Neural Networks (CNN) is developed to reconstruct continuous track geometry profiles from the simulated vehicle dynamic response, performing a full signal-to-signal reconstruction. CNNs are used because they can extract waveform patterns from spatially ordered signals, and previous studies have demonstrated their potential for track geometry reconstruction [5, 6].

Both the dynamic signals and the target irregularity profiles are filtered over the combined D1, D2, and D3 wavelength ranges considered in this study. The regression task focused on the reconstruction of CL and LL profiles. The model receives 200 m windows, corresponding to 1440 samples, with five dynamic inputs: lateral and vertical bogie-frame accelerations and roll, pitch, and yaw angular velocities. The output is the corresponding CL or LL profile for the same window. Before training, the inputs and targets were standardized using the training-set mean and standard deviation to avoid data leakage.

As shown in Figure 3, the model's architecture is designed to extract spatial features at different scales, reflecting the different wavelength ranges of track irregularities. The input signals are processed by three parallel branches with short, medium, and long-range receptive fields. Each branch first projects the dynamic signals into feature maps using a Conv1D, batch normalization, and ReLU block. Dilated residual blocks then progressively add spatial context without reducing the signal resolution. The resulting receptive fields are approximately 25 m, 58 m, and 121 m, providing spatial contexts associated with the D1, D2, and D3 ranges. Finally, the branch feature maps are concatenated and passed through a fusion block, which combines the multi-scale information and generates the predicted CL or LL profile.

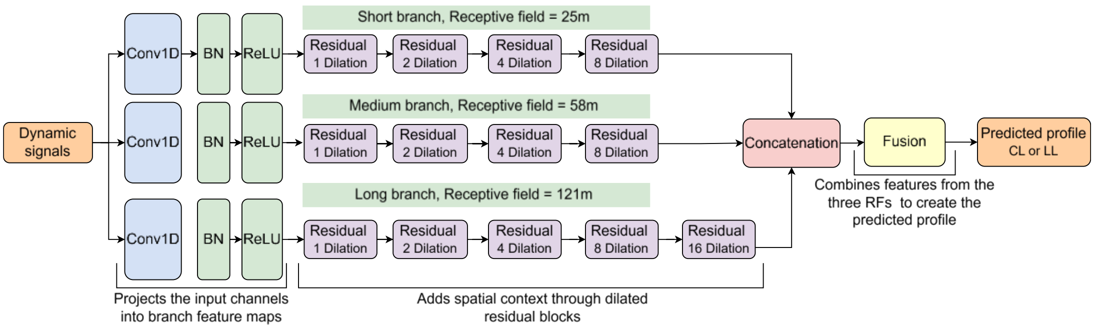

Figure 3: Multi-scale Dilation CNN model architecture

The model is trained using mean squared error loss and the Adam optimizer. Separate models are trained for each binary irregularity combination containing the target profile. For CL reconstruction, the CL-G, CL-LL, and AL-CL combinations were considered, while for LL reconstruction, the CL-LL, AL-LL, and LL-G combinations were used. Performance is averaged over the three combinations associated with each target. Grouped cross-validation by synthetic defect severity is adopted. Predictions are assessed over the complete 3-150 m range and separately for D1, D2, and D3 by applying the same band-pass filtering procedure to the predicted and reference profiles. Performance is evaluated using R^2 and root mean squared error (RMSE), while the true energy fraction was computed to quantify the contribution of each wavelength band to the target profile.

3 Results

Among the SVM models evaluated for the general classification of track sections, the Gaussian SVM achieves the best overall performance. It obtained an accuracy of 88.49%, an F1 score of 89.83%, and a Cohen's Kappa of 0.77. It is therefore selected for the irregularity-specific classification tasks. Table 1 contains the main results obtained for the specialized models for which LL and CL obtained the highest accuracies, reaching 97.96% and 97.70%, respectively, with F1 scores of around 93% and Cohen's Kappa values above 0.90. G also presented strong results, with an accuracy of 96.65% and an F1 score of 89.86%. AL is the most challenging parameter: despite an overall accuracy of 92.80%, its lower sensitivity of 72.0% indicates that some sections with significant AL irregularities are not identified. Overall, the results obtained indicate that bogie-frame dynamic responses contain relevant information for assessing track irregularity severity, especially for LL and CL.

Irreg.	Accuracy	Precision	Sensitivity	F1 score	Cohen's kappa
AL	92.77%	81.19%	71.95%	76.29%	0.72
LL	97.96%	100.00%	88.52%	93.91%	0.93
CL	97.70%	88.01%	100.00%	93.62%	0.92
G	96.65%	91.87%	87.94%	89.86%	0.88

Table 1: Classification results for the specialized models

The multi-scale CNN is evaluated for the reconstruction of CL and LL profiles. Since separate models are trained for each binary irregularity combination, the reported results correspond to the average performance over the three combinations associated with each target. The coefficient of determination, R^2, indicates how well the predicted profiles explain the variability of the reference profiles: values closer to 1 indicate better reconstruction, while negative values indicate poor predictive performance. Overall, the model reconstructed CL profiles more accurately than LL profiles. For the complete 3-150 m range, CL achieved an average R^2 of 0.938 and an RMSE of 0.376 mm, while LL achieved an R^2 of 0.584 and an RMSE of 1.573 mm.

The band-separated results are presented in Table 2. The true energy fraction indicates the relative contribution of each wavelength band to the reference profile. For CL, most of the energy is concentrated in D1 and D2, which are reconstructed accurately, with R^2 values of 0.977 and 0.928, respectively. Although D3 presented a negative R^2 of -0.871, it represents only 4.0% of the CL target energy, limiting its influence on the overall result. For LL, D1 is reconstructed reasonably well, with R^2 = 0.819, while performance decreased in D2, with R^2 = 0.606, and became poor in D3, with R^2 = -2.456. Since D3 represents 15.2% of the LL target energy, its poor reconstruction has a greater effect on the overall LL performance.

The results in Table 2 correspond to the multi-scale configuration, for which all CL and LL combinations are evaluated using a small hyperparameter grid to limit computational cost. To assess the effect of further tuning, an additional LL-focused optimization is performed using wider capacity. For the 3-150 m range, this increased the LL R^2 from 0.584 to 0.656 and reduced the RMSE from 1.573 mm to 1.440 mm. Improvements are mainly observed in D1 and D2, with R^2 increasing from 0.819 to 0.861 and from 0.606 to 0.696, respectively. However, D3 remained poorly reconstructed, indicating that long-wavelength estimation is the main limitation of the current model and cannot be resolved through hyperparameter optimization alone.

Several factors may contribute to the lower D3 performance. D3 components contain fewer oscillations within each 200 m window, potentially making their reconstruction more difficult than that of shorter wavelength components. Moreover, the simulations are performed at a constant speed of 100 km/h. Further investigation is required to determine whether, at this speed, long-wavelength irregularities excite the vehicle response sufficiently for accurate D3 reconstruction. Finally, an MSE only loss may favour smoother predictions. These hypotheses motivate future studies using data generated at higher vehicle speeds and alternative loss functions designed to preserve long-wavelength content.

Target	Band	Band R^2	Band RMSE [mm]	True energy fraction
CL	D1	0.977	0.184	0.610
	D2	0.928	0.238	0.335
	D3	-0.871	0.194	0.040
	ALL	0.938	0.376	0.985
LL	D1	0.819	0.444	0.226
	D2	0.606	1.191	0.602
	D3	-2.456	0.846	0.152
	ALL	0.584	1.573	0.980

Table 2: Average band-separated regression results

4 Conclusions and Contributions

This work investigated the potential use of bogie-frame dynamic responses for continuous railway track geometry monitoring through ML approaches, based on synthetic data generated from the EM120 multibody model. The Gaussian SVM showed strong performance in classifying track-section severity, particularly for LL and CL irregularities. For continuous profile reconstruction, the proposed multi-scale dilated CNN achieved accurate CL estimation and encouraging results for the D1 and D2 wavelength components of LL. However, D3 reconstruction remained challenging for both targets, with a greater effect on the overall LL results. These results support the potential of onboard sensors for scalable track condition assessment while identifying long-wavelength reconstruction as a key direction for future work, including the evaluation of different vehicle speeds and alternative loss functions.

Acknowledgements

The authors acknowledge Fundação para a Ciência e a Tecnologia (FCT) for its financial support via LAETA (project https://doi.org/10.54499/UID/50022/2025) and via the Twin-Rail project (DOI: https://doi.org/10.54499/2024.13264.PEX).

References

[1] P. F. Weston, C. S. Ling, C. J. Goodman, C. Roberts, P. Li, and R. M. Goodall, "Monitoring lateral track irregularity from in-service railway vehicles," *Proc. Inst. Mech. Eng. F J. Rail Rapid Transit*, vol. 221, no. 1, pp. 89–100, 2007.

[2] P. Weston, C. Roberts, G. Yeo, and E. Stewart, "Perspectives on railway track geometry condition monitoring from in-service railway vehicles," *Vehicle System Dynamics*, vol. 53, no. 7, pp. 1063–1091, 2015.

[3] European Rail Research Advisory Council (ERRAC), "RAIL Strategic Research and Innovation Agenda," 2020.

[4] A. De Rosa *et al.*, "Monitoring of lateral and cross level track geometry irregularities through onboard vehicle dynamics measurements using machine learning classification algorithms," *Proc. Inst. Mech. Eng. F J. Rail Rapid Transit*, vol. 235, no. 1, pp. 107–120, Jan. 2021, doi: 10.1177/0954409720906649.

[5] C. Li, Q. He, and P. Wang, "Estimation of railway track longitudinal irregularity using vehicle response with information compression and Bayesian deep learning," *Computer-Aided Civil and Infrastructure Engineering*, vol. 37, no. 10, pp. 1260–1276, Aug. 2022, doi: 10.1111/mice.12802.

[6] X. Hao, J. Yang, F. Yang, X. Sun, Y. Hou, and J. Wang, "Track geometry estimation from vehicle–body acceleration for high-speed railway using deep learning technique," *Vehicle System Dynamics*, vol. 61, no. 1, pp. 239–259, Jan. 2023, doi: 10.1080/00423114.2022.2037669.

[7] S. Munoz, J. Ros, P. Urda, and J. L. Escalona, "Estimation of Lateral Track Irregularity Through Kalman Filtering Techniques," *IEEE Access*, vol. 9, pp. 60010–60025, 2021, doi: 10.1109/ACCESS.2021.3073606.

[8] H. Tsunashima, "Condition Monitoring of Railway Tracks from Car-Body Vibration Using a Machine Learning Technique," *Applied Sciences*, vol. 9, no. 13, p. 2734, Jul. 2019, doi: 10.3390/app9132734.

[9] A. C. Pires, M. C. A. Viana, L. M. Scaramussa, G. F. M. Santos, P. G. Ramos, and A. A. Santos, "Measuring vertical track irregularities from instrumented heavy haul railway vehicle data using machine learning," *Eng. Appl. Artif. Intell.*, vol. 127, p. 107191, Jan. 2024, doi: 10.1016/j.engappai.2023.107191.

[10] J. Ambrósio and J. Pombo, "MUltiBOdy DYnamics Analysis Program – MUBODyn: User's Manual," Lisbon, 2006.

[11] P. Antunes, H. Magalhães, and J. Ambrósio, "A co-simulation approach to the wheel–rail contact with flexible railway track," *Multibody Syst. Dyn.*, vol. 45, no. 2, pp. 245–272, 2019.

[12] J. Pombo and J. Ambrósio, "An alternative method to include track irregularities in railway vehicle dynamic analyses," *Nonlinear Dyn.*, vol. 68, no. 1–2, pp. 161–176, 2012.

[13] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. New York: Springer, 2006.