Proceedings of the Sixth International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 7, Paper 7.4
Civil-Comp Press, Edinburgh, United Kingdom, 2024
ISSN: 2753-3239, doi: 10.4203/ccc.7.7.4

Dynamic Monitoring in High Speed Railways: A Practical Comparison of Axlebox Accelerations Outliers from Different Train Series

P. Salvador Zuriaga, P. Martínez Fernández,
I. Villalba Sanchis and R. Insa Franco

Institue for Transport and Territory,
Universitat Politecnica de Valencia, Spain

Abstract

The aim of this paper is to find whether there are significant differences in the running performance of different train series when running along the same high-speed line under the same circumstances. Axlebox accelerations have proven to be a useful indicator to monitor train-track dynamic interaction, thus serving as the basis of track maintenance. Nevertheless, there are still some doubts around possible remarkable differences in the running behaviour of different high-speed train series. In order to delve into this issue, three different train series currently operating in the Spanish high-speed network were monitored, and both lateral and vertical axlebox acceleration were gathered. Data from each train series were compared using different statistical significance tests. The 66.7% of tests show a noticeable difference in the running performance among the three train series. Hence, this factor may become relevant in some situations such as choosing the type of train series for track maintenance condition monitoring; or when choosing the type of train for providing commercial service.

Keywords: axlebox accelerations, high-speed rail, train-track interaction, track surveying, nonparametric tests, nonparametric statistics.

1 Introduction

Track quality assessment is a major issue in order to ensure a safe performance of railway tracks. Among other methods, this is carried out by means of dynamic track monitoring. This consists on recording accelerations at different train spots while running along the tracks. The higher the maximum running speed allowed is, the more relevant dynamic assessments are. Indeed, railway administrations in charge of a high-speed rail network usually perform these track inspections monthly. For this purpose, they usually rely on high-speed trains conveniently equipped with a set of accelerometers and other motion monitoring devices, together with a data acquisition system [1]. Furthermore, high-speed trains must provide a safe and comfortable ride when they operate under commercial services.

Focusing on the surveying part of the overall track maintenance process, there are different approaches with their respective advantages and drawbacks [2]. Within these approaches, those based on axlebox accelerations, i.e. accelerations measured at both ends of the wheelsets do not introduce filtering effects, as it happens with chord-based approaches [3].

Axlebox accelerations provide some information on the vehicle-track dynamic interaction in the low frequency range [4], [5] as well as in the high frequency one [6], [7]. They also provide some information about the track geometry under load conditions [8], [9]. This is especially relevant since this track geometry is the actual one the trains pass over.

From the point of view of track maintenance, several authors have based their research on axlebox accelerations. They allow detecting and identifying diverse short wavelength track defects [10] [11]. Other works also applied axlebox accelerations for the detection and analysis of rail corrugation or for more general track surveying purposes, taking into account track defects of several wavelength ranges [12], [13]. Furthermore, accelerometers were used to monitor the working conditions of turnouts. In this case, the accelerometers were not only placed at the wheelsets and the bogie frame [14], [15], but also at the nose of the turnout frog [16].

In spite of the extensive research about axlebox accelerations and dynamic train-track interaction carried on up to these days, very little information can be found about comparing the behaviour of different train types running under similar conditions. Some works delve into that issue by modelling the respective train-track interaction [17], but no studies involving real measurements have been found. Particularly scarce are the studies involving lateral axlebox accelerations. These mainly deal with the relationship with track horizontal alignment [18] and train hunting detection [19].

Based on this literature review, the following question arises: does the type of train series utilised for dynamic track recordings provide non-biased results? In other words: if the monitored train belongs to a different train models, would the track tests vary significantly? This may be of relevance particularly for those lines in which passenger services are offered by high-speed trains models different from that used for monitoring. In order to shed light on this issue, the following experiment was set up: three different high-speed trains of different train series were purposely monitored and run along the same high-speed line to check whether there were any significant differences between the recorded data. In the following sections, first the background and the detailed experiment setup is described, as well as the analysis procedure. Afterwards, the obtained results are discussed in detail and the main conclusions of our research are shown.

2 Experiment background & setup

2.1. Network description

High Speed Rail in Spain has a network of 2,618.5 km with standard track gauge (i.e. 1,435 mm) [20]. Since the first line from Madrid o Seville was opened in 1992, new segments have been added up to its actual extent. This network is currently operated at a maximum speed of 300 km/h, with foresight to be raised up to 350 km/h in the midterm.

The High Speed Rail fleet used by Renfe Operadora, the main railway operator in Spain, accounts for three different train series, capable of circulating at its maximum speed:
- Series 100, manufactured by Alstom between 1991 and 1995. Its composition consist of two traction power locomotives at both ends of the train unit, plus eight passenger cars. Passenger cars are articulated, so the ends of adjacent cars turn around the centre pin of the same bogie.
- Series 103, manufactured by Siemens between 2005 and 2006. The train composition consists of eight cars, with distributed power along the train unit. Half of the carriages are motorised and each car lies on its own two bogies.
- Series 112, manufactured by Talgo between 2004 and 2010. Its composition also consist of two traction power locomotives at both ends of the train unit, plus 12 passenger cars. Passenger cars are articulated as well. Unlike series 100, the ends of adjacent cars rest on a single wheelset, which is guided by means of a special coupling between both cars and the wheelset.

As all train series are purposely designed for speeds over 300 km/h, all suspension systems and their components (springs, dampers and inertial masses distribution) are assumed to be equivalent. Therefore, a direct comparison of running behaviour among the train series is reasonable.

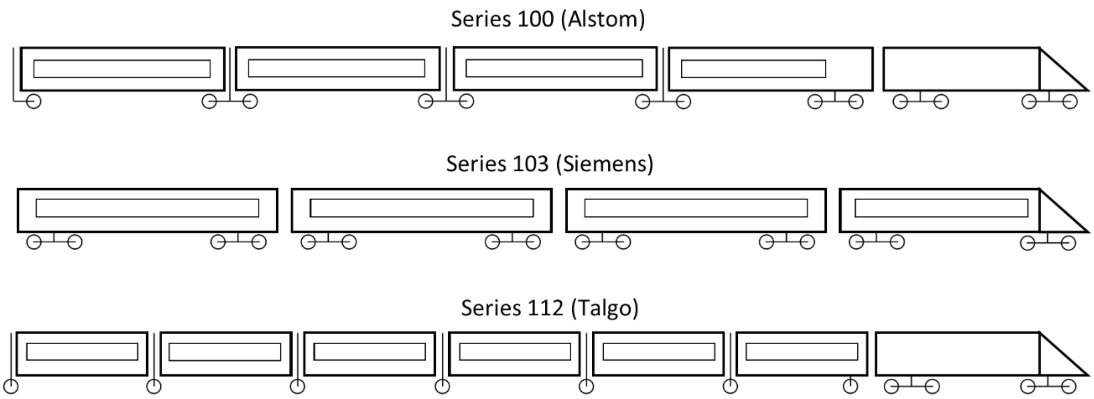

Figure 1. Train composition and wheel arrangement for each train series operating in the Spanish High Speed network. Only half of each train unit is depicted.

Figure. 2. From left to right: train series 100 (Alstom), 103 (Siemens) and 112 (Talgo)

2.2. Line description

The tested line links Madrid Puerta de Atocha with Valencia Joaquín Sorolla stations. It is managed by the Spanish Railway Administrator, Adif. This line is 396 km long and was put into service in December 2010. From Madrid, the line shares the first 29 km with the Madrid-Seville line, and then it turns Eastbound. The line mainly consist of ballasted track with monolithic concrete sleepers. Only tunnels longer than 2 km switch to slab track. It is electrified at 25 kV AC and equipped with Automatic Control Block System. Commercial services are offered in this line by train series 100 and 112. Figure 3 shows the line scheme.

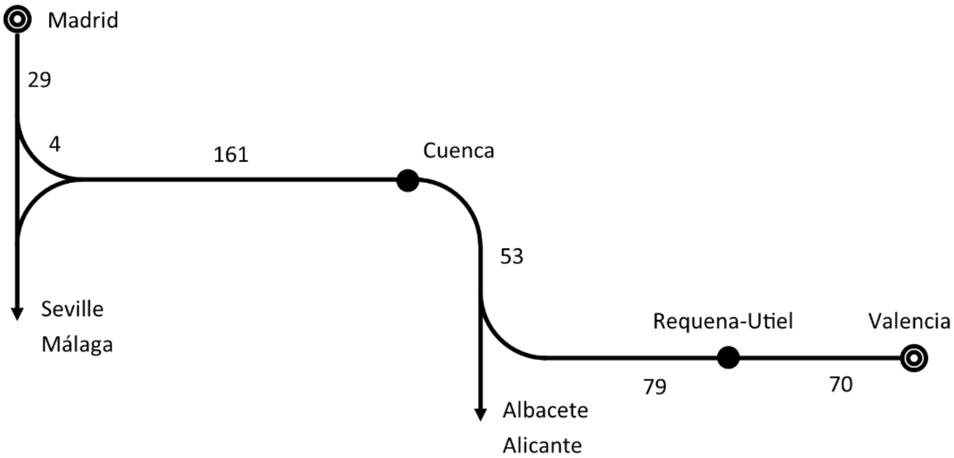

Figure 3. Scheme of the Madrid-Valencia High-Speed Line with kilometric distances.

2.3. Railway track surveying

The process followed by Adif for assessing the maintenance condition of railway tracks by means of train accelerations is as follows: Raw data is gathered at a sampling frequency of 400 Hz with a cut-off frequency of 100 Hz. Lateral and vertical accelerations are further low-pass filtered at frequencies of 10 Hz and 20 Hz, respectively, and subsampled at 100 Hz. After this, a peak analysis in the time domain is carried out. For this, accelerations are classified into three categories according to their absolute value. These categories are: Alert Limit (AL), Intervention Limit (IL), and Immediate Action Limit (IAL). Thresholds for each direction and category are shown in Table 1. Such values are proposed based upon previous research [21].

Category	Lateral (m/s^2)	Vertical (m/s^2)
AL	15	30
IL	25	50
IAL	35	70

Table 1. Threshold values of axlebox accelerations for lateral and vertical directions

2.4. Setup

From this line, three 10 km long track stretches have been analysed. These are located between Kilometric Posts (KPs) 250-260, 320-330 and 360-370, respectively. In these stretches, ballasted track was set up and all train series were running at speeds between 295 and 300 km/h, i.e. the maximum authorised speed. Therefore, some variabilities such as the train speed and the track system are avoided, which would blur the lines when comparing each train performance.

Tests took place during a short timespan so the track maintenance conditions did not present significant variations. During the tests, lateral and vertical axlebox accelerations were recorded with a frequency of 1000 Hz. In all train series, the monitored wheelset was the closest to the train centre. According to the procedure utilised by Adif, data were low-pass filtered by 10 Hz for lateral accelerations and 20 Hz for vertical accelerations. After this, data were down-sampled to 100 Hz. This downsampling procedure only uses 1 out of each 10 data, which means that the remaining 9 items are not used. In order to take advantage of all registered data, systematic downsamplig is carried out. In this way, 10 different independent samples sampled at 100 Hz from the same track stretch are obtained by choosing the 1st, 2nd, 3rd, and so on from the original data recorded at 1000 Hz. In other words, the *j*-th sub-sample takes the *j*-th datum out of each 10. Figure 4 shows this procedure.

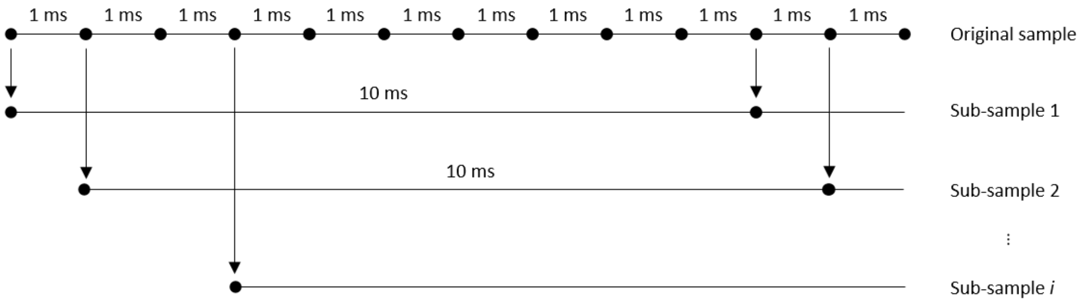

Figure 4. Systematic downsampling for obtaining 10 subsamples equal-spaced 10 ms from a single sample equal-spaced 1 ms.

3 Analysis procedure

The purpose of the analysis is to find whether the train performance along the tracks is similar for all train series, or there are one or more train series that behave significantly differently. Data analysis is based on different statistical confidence tests that prove whether the accelerations recorded by each train series may have the same probability distribution, i.e. come from the same population.

Each sample is around 12,000 items. Such large amount of data would cause the least significant difference (LSD) intervals to be very small in ANOVA-type tests, thus leading to erroneous evidence of dissimilarity. In order to overcome this handicap, only outliers are considered. For this study, acceleration values exceeding their respective IL have been considered as outliers. This yields a more reasonable number of data to which statistical analyses can be applied.

3.1. Preliminary data appearance

Prior to choosing the statistical tests to be applied, Figure 5 shows the box plots of the acceleration outliers for the three track stretches, the three train series and both directions (i.e. lateral and vertical). For each stretch, the first subsample out of 10 is selected as representative of the whole stretch. In all cases, a simple visual analysis of data shows a strong non-normal distribution. This is also valid for the remaining 9 subsamples and was an expected result due to the nature of outliers. Therefore, only a non-parametric analysis is performed. Something worth noting is that train series 112 does not yield lateral acceleration values greater than 25 m/s^2. For this reason, only lateral accelerations for series 100 and 103 are shown.

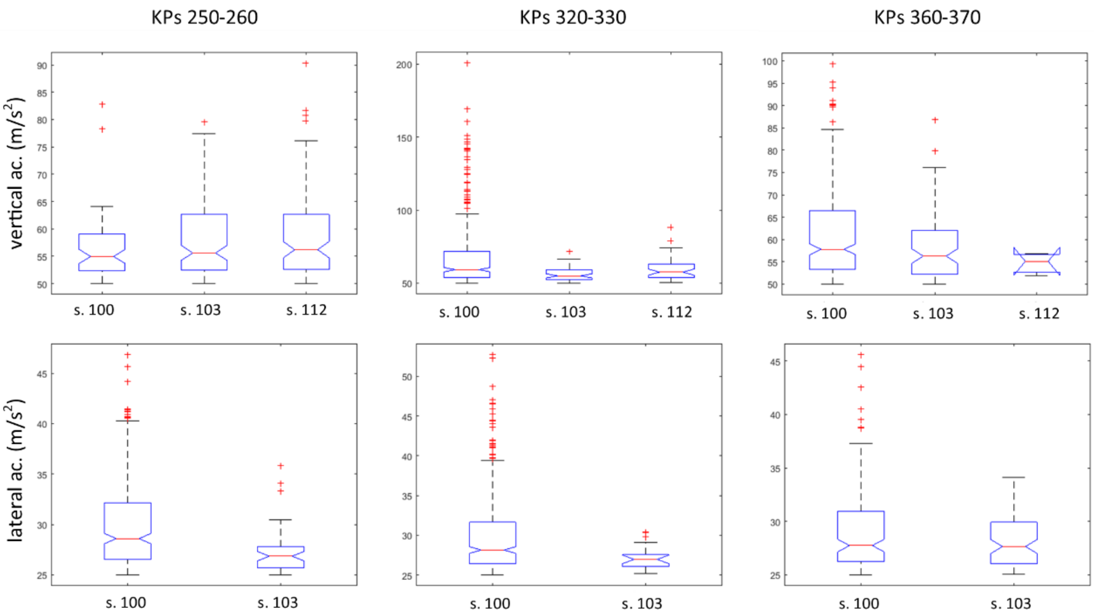

Figure 5. Box plots for lateral and vertical accelerations outliers of different train series at the three tested stretches

A preliminary look at this data lets us see some different behaviours among train series and track stretches. In KPs 320-330, vertical accelerations from series 100 are twice as high as the other two train series in the same stretch, and also twice as high as the same train series in the other two track stretches. In contrast, for lateral accelerations, although series 100 presents higher values than series 103, maximum values among the stretches for both train series are consistent.

Other important remarkable findings are the absence of outliers in the box plots, marked as isolated crosses, for lateral accelerations of series 103 and vertical accelerations of series 112, both in the stretch between KPs 360-370. In addition, no values above the IAL are found for lateral accelerations of train series 103 in KPs 320-330 and KPs 360-370, nor for vertical accelerations of train series in KPs 360-370. These findings are important to keep in mind when discussing the results given by the different statistical tests.

In the following subsection, a set of statistical tests are selected for comparing samples from different train series. Each test makes the comparison from a different approach, thus strengthening the evidences and the validity of results.

3.2. Analysis with all train series together

The first set of tests considers all train series at once, so that they provide a holistic point of view. Let *i* denote the track stretch, *j* the data subsample from track stretch *i*, *k* the direction of axlebox accelerations, and *l* the train series. Each variable {*i*, *j*, *k*, *l*} has the following range:

- $i = 1, \ldots, 3$
- $j = 1, \ldots, 10$
- $k = 1, 2$
- $l = 1, \ldots, 3$

In order to give a stronger evidence in the similarity/dissimilarity of results among the train series, three different well-known statistical tests are considered. For each track stretch *i* and each direction *k*, each subsample *j* from train series *l* is compared against the rest of subsamples from the remaining two train series. This yields a total number of 1000 comparisons for each direction, track stretch and type of test. This allows having not an “all-nothing” decision about the similarity/dissimilarity of trains behaviour, but a degree of validity expressed as the percentage of comparisons among samples that are significantly different.

The first kind are contingency tables [22]. This test classifies data from each sample into two or more categories. Then, it compares the proportion of data in each class among all samples. For this test, the following classes are defined:

- For lateral accelerations ($k = 1$): [25, 35[, [35, ∞[
- For vertical accelerations ($k = 2$): [50, 70[, [70, ∞[

These classes purposely coincide with the IL and IAL. Indeed, this test allows comparing samples taking into account the outliers distribution into the limits set up for track maintenance.

The statistic is given by Eq. (1)

$$T_{tc} = \sum_{l=1}^{3} \sum_{m=1}^{2} \frac{O_{lm}^2}{E_{lm}} - N \tag{1}$$

where

$E_{lm} = \frac{n_l C_m}{N}$

n_l is the number of outliers for train series *l*

C_m is the number of outliers in each class *m* (i.e. IL or IAL)

O_{lm} is the number of outliers for each train series *l* and class *m*

For this test to work properly, it is recommended that each $O_{lm} > 5$.
The second test is the Kruskal-Wallis test [23]. It is the non-parametric analogue ANOVA test for more than two samples. In this test, data from all samples are ranked in ascending order, and the sum of ranks of each sample is compared with the total set of data. Hence, this test compares the different samples according to the rank of each element inside their own sample with the rank in the overall set of data.
The statistic is given by Eq. (2)

$$T_{KW} = \frac{12}{N(N+1)} \sum_{l=1}^{3} \frac{{R_l}^2}{n_l} - 3(N+1) \tag{2}$$

where $R_l = \sum_{p=1}^{n_l} R(X_{lp})$ is the sum of the ranks assigned to the l-th sample.
The third test is the Bell-Doksum test [24]. This test measures correlation among samples, i.e. difference among subsamples is determined in terms of whether they are weakly correlated. For this, the test uses order statistics from normally distributed random samples instead of ranks as the Kruskal-Wallis test does.
The test statistic is given by Eq. (3)

$$T_{BD} = \sum_{l=1}^{3} n_l (Z_l - \bar{Z})^2 \tag{3}$$

where
$Z_l = \frac{1}{n_l} \sum_{p=1}^{n_l} Z[R(X_{lp})]$ is the average of the outliers for each train series
$Z[R(X_{lp})]$ is the r-th smallest random normal deviate, assigned to X_{lp}
$\bar{Z} = \frac{1}{N} \sum_{r=1}^{N} Z(r)$ is the average Z value

3.3. Analysis of paired samples

In order to have more specific information, samples from each train series are compared 2 by 2. This yields a number of comparisons given by $\binom{3}{2} = 3$. Again, for each direction k, each subsample j from train series l is compared with the rest of subsamples from the other train series. This yields a total number of 100 comparisons of each type of test per direction, track stretch and paired train series. For this comparison, two well-known nonparametric statistical tests, Mann-Whitney and Cramer-von Mises, are chosen as well.
The Mann-Whitney test [25] is the nonparametric analogue ANOVA test for two samples. Like Kruskal-Wallis test, data from both samples are ranked in ascending order, and the sum of ranks of each sample is compared with the total set of data.
The statistic test is given by Eq. (4)

$$T_{MW} = S - \frac{n(n+1)}{2} \tag{4}$$

where
n is the size of sample 1 (provided $n > 0$)
$S = \sum_{p=1}^{n} R(X_p)$ is the sum of the ranks assigned to the observations from sample 1.

The Cramer-von Mises test [26], [27] belongs to the group of statistics of the Kolmogorov-Smirnov type [28]. The statistic value computes the sum of the differences between the empirical distributions of both samples, rather than comparing data ranks as it does the Mann-Whitney test, thus providing a different insight. It is given by Eq. (5).

$$T_{CvM} = \frac{n_x n_y}{\left(n_x + n_y\right)^2}\left[\sum_{i=1}^{n_x}\left(\frac{R\left(X^{(i)}\right)}{n_y} - i\,\frac{n_y + n_x}{n_x n_y}\right)^2 + \sum_{j=1}^{m}\left(\frac{R\left(Y^{(j)}\right)}{n_x} - j\,\frac{n_y + n_x}{n_x n_y}\right)^2\right] \quad (5)$$

where *X* and *Y* are the tested samples of length n_x and n_y, respectively. $R\left(X^{(i)}\right)$ and $R\left(Y^{(j)}\right)$ are the ranks, in the combined ordered sample, of the *i*th smallest of the *X*'s denoted by $X^{(i)}$, and the *j*th smallest of the *Y*'s denoted by $Y^{(j)}$, respectively.

4 Results and discussion

Tables 2 and 3 show the percentage of tests in each track stretch in which the null hypothesis H_0 is rejected in favour of the alternative hypothesis H_1 for lateral and vertical accelerations, respectively. The results compare each *j*-th subsample for all train series. Tests results with p-values lower than 0.05 indicate that H_0 is rejected in favour of H_1. This means that, in these cases, there are significant differences among the train series in their running behaviour.

	lateral accelerations								
	S.100 - S.103		S.100 - S.112		S.103 - S.112		S.100 - S.103 - S.112		
KPs	M-W	C-vM	M-W	C-vM	M-W	C-vM	Cnt. Tbl.	K-W	B-D
250-260	100%	100%	100%	100%	100%	100%	2,0%	100,0%	100,0%
320-330	100%	100%	100%	100%	100%	100%	0,0%	100,0%	88,7%
360-370	1%	0%	100%	100%	100%	100%	100,0%	1,0%	9,9%
Average	67%	67%	100%	100%	100%	100%	34,0%	67,0%	66,2%

Table 2. Percentage of tests in which H_0 is rejected in favour of H_1 for lateral accelerations

	vertical accelerations								
	S.100 - S.103		S.100 - S.112		S.103 - S.112		S.100 - S.103 - S.112		
KPs	M-W	C-vM	M-W	C-vM	M-W	C-vM	Cnt Tbl.	K-W	B-D
250-260	68%	70%	74%	76%	0%	0%	34,7%	66,0%	59,6%
320-330	100%	100%	53%	70%	61%	57%	100,0%	100,0%	100,0%
360-370	16%	8%	38%	48%	24%	28%	23,5%	38,6%	40,3%
Average	61%	59%	55%	65%	28%	28%	52,7%	68,2%	66,6%

Table 3. Percentage of tests in which H_0 is rejected in favour of H_1 for vertical accelerations

The most evident result, already shown in Figure 5 is the total absence of lateral accelerations outliers for train series 112. This means that all accelerations values in this direction remain under the IL, thus making a significant difference with the other train series. Therefore, for lateral accelerations, tests comparing samples from all train series at once are actually comparing series 100 and series 103 only. For the same reason, Mann-Whitney and Cramer-von Mises tests comparing series 100 and 103 with series 112 have their p-values set to zero and, hence, H_0 is rejected in all cases.
In overall terms, there is consistency in the results yielded by different tests. The only exception are contingency tables. For lateral accelerations, while the other statistical tests yield significant differences among train series in KPs 250-260 and PKs 320-330, contingency tables do not reject the null hypothesis. In contrast, in KPs 360-370, the opposite occurs. The reasons for this apparently paradoxical result are the following: Firstly, some of the O_{lm} in some contingency tables equal 0 as shown in section 3.1, and this alters the test. In the second place, from Figure 5, a different distribution of lateral accelerations among the three track stretches is seen. In KPs 250-260 and KPs 320-330, where both samples present extreme values in their respective box plots, the test tends to be rather conservative, not yielding significant differences. However, in PKs 360-370, series 103 does not present extreme values, so the test interprets this as a significant difference, despite acceleration values for series 103 being higher in PKs 360-370 than in PKs 320-330.

In spite of this fact, results mostly show significant differences among train series. In stretches between KPs 250-260 and 320-330, only vertical accelerations in KPs 250-260 for train series 103 and 112 are not conclusive. In contrast, KPs 360-370 show less significant results. For lateral accelerations, only 1% of Mann-Whitney tests and none of Cramer-von Mises tests show conclusive differences between series 100 and 103. For vertical accelerations, there are no significant differences between series 100 and 103. Indeed, only the Mann-Whitney test shows in Table A2 a p-value lower than 0.05 in subsample #9 from KPs 360-370, but this value is 0.047, very close to the significance threshold. The rest of the comparisons show significance among train series only in the 20%-50% of the tested subsamples. This fact may be explained by the small number of acceleration values registered in this stretch for train series 112 compared to the other train series. Indeed, whereas series 100 and 103 have more than 280 and 80 items per sample, respectively, series 112 only accounts for 6 items at most. Therefore, conclusions are weaker than in the other track stretches, where a more homogeneous number of items per train series is available.

As a main conclusion, recorded values shown in Figure 5 together with results from Tables 2 and 3 show dissimilarity in the running performance among the analysed train series. The p-values given by different statistical tests, in most cases under 0.05, support this finding. Those cases where p-values are higher than 0.05 are mainly due to a low number of available items per sample rather than a true similarity. Hence, this remarkable difference in the number of items of each sample in some of the track stretches for some train series also support the aforementioned conclusion.

5 Conclusions

The running performance of different high-speed train series along a railway line has been analysed and compared. For this purpose, axlebox accelerations were gathered in lateral and vertical directions. Three different train series, with different wheel arrangements, which usually serve high-speed commercial services in Spain, were monitored. The analysis was carried out over the acceleration outliers exceeding the IL levels previously defined. From the whole line, three 10 km long track stretches where all train series were running at maximum speed were chosen.

A preliminary representation of these outliers by means of box plots confirmed their non-normal statistical distribution. Under this circumstance, a set of non-parametric tests were selected for the analysis. All statistical tests were consistent in showing significant differences in the running performance among the train series for both lateral and vertical directions. Those cases in which tests have not yielded concluding results are mainly due to a lack of data from certain train series in some track stretches. Therefore, dynamic behaviour of lab trains, in charge of monitoring the ride comfort and track quality, may substantially differ from the trains actually utilised for commercial services.

Funding

Project funded by Spanish Ministry of Science and Innovation. Ref. TED2021-131560A-I00

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

[1] I. La Paglia, M. Carnevale, R. Corradi, E. Di Gialleonardo, A. Facchinetti, and S. Lisi, “Condition monitoring of vertical track alignment by bogie acceleration measurements on commercial high-speed vehicles,” *Mech Syst Signal Process*, vol. 186, p. 109869, Mar. 2023, doi: 10.1016/J.YMSSP.2022.109869.

[2] S. L. Grassie, “Measurement of railhead longitudinal profiles: a comparison of different techniques,” *Wear*, vol. 191, no. 1–2, pp. 245–251, Jan. 1996, doi: 10.1016/0043-1648(95)06732-9.

[3] A. Falamarzi, S. Moridpour, and M. Nazem, “A Review on Existing Sensors and Devices for Inspecting Railway Infrastructure,” *Jurnal Kejuruteraan*, vol. 31, no. 1, pp. 1–10, 2019, doi: 10.17576/jkukm-2019-31(1)-01.

[4] L. Xu and W. Zhai, “A three-dimensional dynamic model for train-track interactions,” *Appl Math Model*, vol. 76, pp. 443–465, Dec. 2019, doi: 10.1016/j.apm.2019.04.037.

[5] C. J. Yang, Y. Xu, W. D. Zhu, W. Fan, W. H. Zhang, and G. M. Mei, “A three-dimensional modal theory-based Timoshenko finite length beam model for train-track dynamic analysis,” *J Sound Vib*, vol. 479, p. 115363, Aug. 2020, doi: 10.1016/j.jsv.2020.115363.

[6] L. Baeza, J. Giner-Navarro, D. J. Thompson, and J. Monterde, "Eulerian models of the rotating flexible wheelset for high frequency railway dynamics," *J Sound Vib*, vol. 449, pp. 300–314, Jun. 2019, doi: 10.1016/j.jsv.2019.03.002.

[7] L. Baeza, J. Giner-Navarro, and D. J. Thompson, "Reply to 'Discussion on "Eulerian models of the rotating flexible wheelset for high frequency railway dynamics" [J. Sound Vib. 449 (2019) 300-314],'" *J Sound Vib*, p. 115665, Aug. 2020, doi: 10.1016/j.jsv.2020.115665.

[8] M. Uranjek, A. Štrukelj, S. Lenart, and I. Peruš, "Analysis of influential parameters for accelerated degradation of ballast railway track," *Constr Build Mater*, vol. 261, p. 119938, Nov. 2020, doi: 10.1016/j.conbuildmat.2020.119938.

[9] I. Soleimanmeigouni, A. Ahmadi, and U. Kumar, "Track geometry degradation and maintenance modelling: A review," *Proc Inst Mech Eng F J Rail Rapid Transit*, vol. 232, no. 1, pp. 73–102, Jan. 2018, doi: 10.1177/0954409716657849.

[10] P. Haji Abdulrazagh, M. T. Hendry, A. Roghani, and E. Toma, "Evaluating Rail Surface Roughness from Axle-Box Acceleration Measurements: Computational Metrology Approach," *J Transp Eng A Syst*, vol. 147, no. 12, p. 04021087, Dec. 2021, doi: 10.1061/JTEPBS.0000601/ASSET/15F21249-D75E-4212-9DCF-DE095EB1003A/ASSETS/IMAGES/LARGE/FIGURE11.JPG.

[11] S. Rapp, U. Martin, M. Strähle, and M. Scheffbuch, "Track-vehicle scale model for evaluating local track defects detection methods," *Transportation Geotechnics*, vol. 19. Elsevier Ltd, pp. 9–18, Jun. 01, 2019. doi: 10.1016/j.trgeo.2019.01.001.

[12] P. Salvador, V. Naranjo, R. Insa, and P. Teixeira, "Axlebox accelerations: Their acquisition and time-frequency characterisation for railway track monitoring purposes," *Measurement (Lond)*, vol. 82, 2016, doi: 10.1016/j.measurement.2016.01.012.

[13] J. Morais *et al.*, "Continuous monitoring and evaluation of railway tracks: System description and assessment," in *Procedia Structural Integrity*, Elsevier B.V., Jan. 2019, pp. 448–455. doi: 10.1016/j.prostr.2019.08.059.

[14] J. Silva-Rodríguez, P. Salvador, V. Naranjo, and R. Insa, "Supervised contrastive learning-guided prototypes on axle-box accelerations for railway crossing inspections," *Expert Syst Appl*, vol. 207, p. 117946, Nov. 2022, doi: 10.1016/J.ESWA.2022.117946.

[15] M. Sysyn, O. Nabochenko, U. Gerber, V. Kovalchuk, and O. Petrenko, "Common crossing condition monitoring with on-board inertial measurements," *Acta Polytechnica*, vol. 59, no. 4, pp. 423–434, 2019, doi: 10.14311/AP.2019.59.0423.

[16] S. Bruni, I. Anastasopoulos, S. Alfi, A. Van Leuven, and G. Gazetas, "Effects of train impacts on urban turnouts: Modelling and validation through measurements," *J Sound Vib*, vol. 324, no. 3–5, pp. 666–689, 2009, doi: 10.1016/j.jsv.2009.02.016.

[17] A. Rhodes, G. Hunt, N. Harwood, N. Kuka, L. Baron, and J. Borrell, "Track damage comparison between conventional and articulated trains operating on a Great Britain railways 'classic' mainline route and a high-speed route," *Proc Inst Mech Eng F J Rail Rapid Transit*, vol. 233, no. 7, pp. 743–761, Aug. 2019, doi: 10.1177/0954409718805256.

[18] A. De Rosa, S. Alfi, and S. Bruni, "Estimation of lateral and cross alignment in a railway track based on vehicle dynamics measurements," *Mech Syst Signal Process*, vol. 116, pp. 606–623, Feb. 2019, doi: 10.1016/j.ymssp.2018.06.041.

[19] R. Kulkarni, A. Qazizadeh, M. Berg, U. Carlsson, and S. Stichel, “Vehicle running instability detection algorithm (VRIDA): A signal based onboard diagnostic method for detecting hunting instability of rail vehicles,” *Proc Inst Mech Eng F J Rail Rapid Transit*, vol. 236, no. 3, pp. 262–274, May 2021, doi: 10.1177/09544097211020578.
[20] Adif Alta Velocidad, “Network Declaration 2020,” *Network Declaration 2020*. Adif Alta Velocidad, Madrid, pp. 1–249, Aug. 05, 2020.
[21] P. Salvador Zuriaga, “Desarrollo de técnicas avanzadas en la auscultación dinámica para la optimización del mantenimiento de vías férreas,” Editorial Universitat Politècnica de València, Valencia (Spain), 2014. doi: 10.4995/Thesis/10251/39003.
[22] W. J. Conover, *Practical Nonparametric Statistics*, First. New York, 1971.
[23] W. H. Kruskal and W. A. Wallis, “A Nonparametric test for the Several Sample Problem,” *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 525–540, Dec. 1952.
[24] C. B. Bell and K. A. Doksum, “Distribution-Free Tests of Independence,” *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 429–446, Apr. 1967, doi: 10.1214/aoms/1177698959.
[25] H. B. Mann and D. R. Whitney, “On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other,” *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, Mar. 1947, doi: 10.1214/aoms/1177730491.
[26] H. Cramer, *Mathematical methods of statistics*, Princeton. Princeton: Princeton University Press, 1946.
[27] R. von Mises, *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik* , First. Leipzig, 1931.
[28] E. J. Burr, “Small-Sample Distributions of the Two-sample Cramer-Von Mises' W^2 and Watson's U^2,” *The Annals of Mathematical Statistics*, vol. 35, no. 3, pp. 1091–1098, Sep. 1964, doi: 10.1214/aoms/1177703267.