Operational Wheel Flat Detector in Railway Vehicles

Ibon Erdozain[1], Blas Blanco[2], Luis Baeza[3], and Asier Alonso[4]

[1] CAF, Beasain, Gipuzkoa, Spain.
ierdozain@caf.net
[2]CAF I+D, Beasain, Gipuzkoa, Spain
blas.blanco@caf.net
[3]I2MB Universitat Politècnica de València, Valencia, Spain
baeza@mcm.upv.es
[4] CAF I+D, Beasain, Gipuzkoa, Spain & TECNUN (Universidad de Navarra), Donostia, Gipuzkoa, Spain
asier.alonso@caf.net

ABSTRACT

Maintenance of railway systems is shifting from being based on scheduled interventions to a continuous regime based on the actual status of assets. This change is supported mainly on three pillars: the development of new sensors and signal processing techniques, the capability to store and analyze all the information gathered by this huge amount of new sensors, and the capability of modifying dynamically the maintenance plans. This paper presents a new wayside system for detecting flats whose development has been based on combining physical models with Machine Learning Techniques. Physical models are used to understand the phenomena, define the key indicators to characterize the phenomena and generate synthetic data to train Machine Learning algorithms. Subsequently, regression models are generated to relate the key parameters with the flat severity. The last part of the paper is focused on validating the proposed methodology in a real environment.

1. INTRODUCTION

Maintenance of railway systems is shifting from being based on scheduled interventions to maintenance based on the actual status of assets. This change is supported mainly on three pillars:

- The development of sensors and signal processing techniques that allow measuring variables related to the current status of the elements in a reliable and non-expensive way.
- The capability to transmit, store and analyze the information acquired by the different sensors. The final objective is to have an up-to-date picture of the status of all the monitored elements.
- The dynamic modification of maintenance activities taking into account the health status of the elements and the characteristics of maintenance facilities.

Following this trend, several systems have been developed in the last years to monitor the health status of different railway components (such as brakes, gearboxes, or bearings) to perform Condition-Based Maintenance (CBM).

One of the elements that is more critical from both an economic and safety point of view is the wheelset. Railway wheelsets are responsible for supporting the mass of the vehicle and guiding it across the tracks; also, all the dynamic reactions to the vehicle are transmitted through the wheel-rail contact. The inspection of the wheelset encompasses several features such as wheel profile measurement, crack detections, out-of-roundness, ovalization and flats.

The flats are local wheel tread imperfections with a circumference chord shape. They generally result from excessive braking that locks the wheel and makes it to slide along the rail. The presence of external elements such as leaves, grease or snow reduces the adhesion and increases the likelihood of problems in the wheel sliding protection system. In the sliding process, part of the wheel material, whose hardness is usually lower than that of the rail, is deformed and even torn away, creating the wheel flat. Friction between wheel and rail also raises the temperature at the contact point, although it rapidly cools down when the wheel starts rolling again, leading to martensite formation. The resulting residual stresses and the brittle nature of the martensitic structure, combined with the periodic impacts against the rail, lead to cracks that may evolve into the spalling problem (Jergéus 1998).

However, the problems that wheel defects generate do not just restrict to the wheel condition and maintenance. Flats produce an impact on the rail every wheel turn and can cause serious damage to the track, shortening the life of the rail (Nielsen, Ringsberg, and Baeza 2005). Apart from that, these impacts cause noise and vibration problems that affect the comfort of passengers. Therefore, it is necessary to detect these defects and solve them as promptly as possible.

Traditionally, these defects were detected in thorough workshop inspections, so defective wheels were usually in service for a considerable amount of time before they were corrected. However, as previously stated, technological advances in recent years have enabled the transition towards a CBM approach. Research in this field is mostly focused on wayside systems (Alemi et al. 2019; Gao et al. 2020; Mosleh et al. 2021, 2023; Stratman, Liu, and Mahadevan 2007). Moreover, commercial solutions are also currently available (LBFoster s. f.; Schenck process s. f.; voestalpine s. f.). Both research work and commercial solutions are based on different sorts of sensors. According to (Iwnicki, Nielsen, and Tao 2023), the employed technologies are strain gauges, load cells, optic fibers and accelerometers.

Systems based on acceleration signals measured in the rail are a non-invasive and low-cost choice. Despite, wheel flats can be detected by rail accelerations, they lack quantitative information about the impact load (Iwnicki, Nielsen, and Tao 2023). This causes the maintenance is usually based on an indirect measurement (acceleration) instead of the impact force that the flat is producing. The relationship between accelerations and forces is not straightforward as it depends on multiple factors (relative position of the impact with respect to the accelerometer, vehicle load, speed flat shape…)

The objective of this paper is to develop a methodology that allows determining the force created by an impact as a function of acceleration measurements. For that, in the first stages a physical model of the phenomena will be created to:

- Analyze in detail the problem and how the different factors affect the measurements.
- Define key parameters that allow characterizing the phenomena and required signal analysis.

With these results, a methodology that allows estimating the impact force will be created using Machine Learning (ML) techniques. As it is expected, very few experimental signals with flats are available to train the model. Therefore, synthetic signals obtained with the physical model will be employed to train the ML model.

Finally, the methodology will be tested on a real track. For that, a prototype of the wheel detector algorithm is installed in a Metro Line and the results of the system are checked analyzing the repeatability and accuracy of the predictions.

2. DIGITAL TWIN DEVELOPMENT

2.1. Vehicle Track Interaction Model

In order to study the dynamic wheel-rail interaction (WRI), a numerical model based on the finite element method is formulated (Figure 1). The model uses the Timoshenko beam theory for the rail modelling, which considers the shear strength and rotatory inertia of the rail section and is able to represent accurately the vertical dynamics of the track in a wider range of frequencies than the Euler-Bernoulli beam theory. As the main variable of interest of the dynamic simulations is the wayside accelerations, it is necessary to avoid the non-physical response of the standard Timoshenko finite element with moving when the contact is transferred from one Timoshenko element to the next one. This is done by implementing the solution proposed by (Blanco et al. 2019). The model also introduces the improvements presented in (Blanco et al. 2022) to represent accurately the finite dimensions of the pads, which is important to calculate accurately the track resonances.

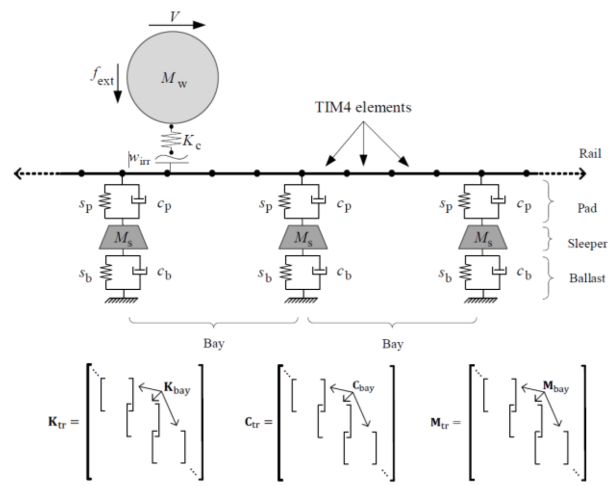

Figure 1 Model diagram

This study is focused on the WRI, so the vehicle has been simplified to its unsprung mass (half of the wheelset mass) with the pre-load of the primary suspension applied to it (considering in this way the mass of the eighth part of the coach). This simplification is justified on the basis that vehicle suspension affects mainly the low-frequency range of the response, whereas the defects on the wheel treads, especially the wheel flats, concern the high-frequency range. This is so because the contact stiffness is some orders of magnitude larger than the primary suspension stiffness.

A Hertzian non-linear contact is used to define the normal WRI and determine the normal contact force. This model allows relating the relative displacement of the bodies in contact with the local deformations and the force transmitted through the wheel rail contact.

2.2. Model Validation

In order to assure that the simulation model represents correctly the train track interaction when flats are simulated two different approaches have been followed: on the one hand the theoretical impact force results provided by the model have been compared with results found in the bibliography; on the other hand, experimental results of wayside accelerations (WA) have been compared with the model predictions.

2.2.1. Impact Force Validation

The results obtained with the model have been compared with the results published by Mazilu in [3]. For that, the WRI in the presence of a 60mm long and 0.35mm deep flat (L=60mm and d=0.35mm) at a rolling speed of 24m/s has been simulated. The contact force evolution for an impact of the flat on the mid-span is shown in Figure 2a and on the mid-support in Figure 2b. These curves seem very similar to the ones shown in (Mazilu 2007) and the maximum values confirm the idea: 197kN and 211.9kN compared to the 196.5kN and 211kN for the mid-span and mid-support impacts, respectively.

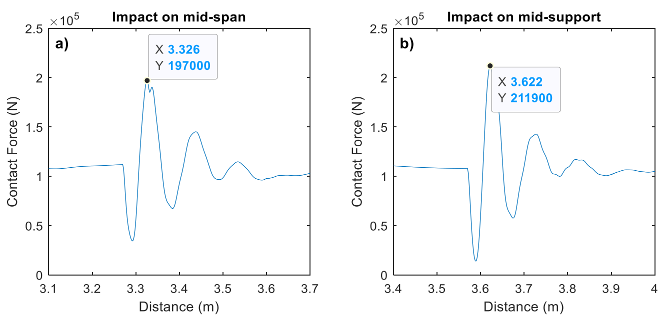

Figure 2. Contact forces when a $\boldsymbol{L = 60}$ mm and $\boldsymbol{d = 0.35}$ mm flat impacts on a) mid-span, and b) mid-support.

2.2.2. Wayside acceleration validation

The accuracy of the model in predicting rail accelerations caused by flats is validated in relation to measured rail accelerations. Figure 3 shows the measured response due to a flat and the simulated response. The flat dimensions were adjusted to match the obtained response, leading to a slightly worn flat with a length of 35 mm and 7 μm of depth, which lies within the common flat sizes (Maglio et al. 2021). There is a high agreement between both responses, in both peak values and frequency content. Nevertheless, dissipation seems to be higher in the simulated response.

2.2.3. Discussion

In conclusion, it can be said that the model allows reproducing accurately the behavior of the WRI when the wheel has a flat. Due to that, simulated responses can be used to find the dominant features affecting the response due to flat impact, determine the signal analysis, the important parameters that comprise the information related to a flat impact and to create labeled simulations to train Machine Learning Models.

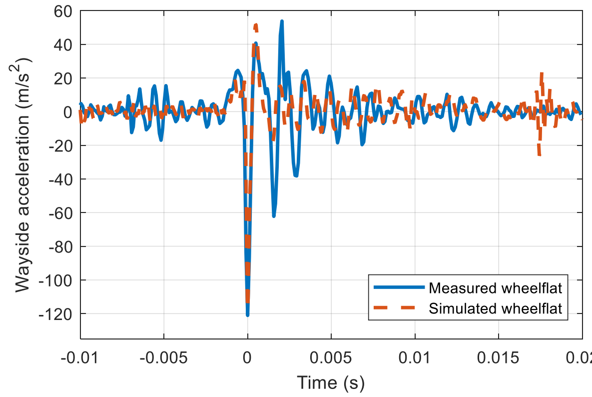

Figure 3. Comparison of measured and simulated WA due to a flat impact.

3. Analysis of simulated flat impact signals

3.1. Simulation of WA due to different flat profiles

The shape of a flat evolves with time from its original shape when it originates. Originally, the flat has a circumference chord shape with sharp transitions from a circular profile to a straight flat profile. With time, wear smoothens the flat shape leading to a rounded profile, despite not being circular. Figure 4**¡Error! No se encuentra el origen de la referencia.** shows the simulated rail acceleration at a stationary point due to an impact at three different stages of its evolution. Newly formed flats cause high peak values of the rail acceleration, which progressively diminishes as the flat gets worn. However, the frequency content does not seem to change significantly.

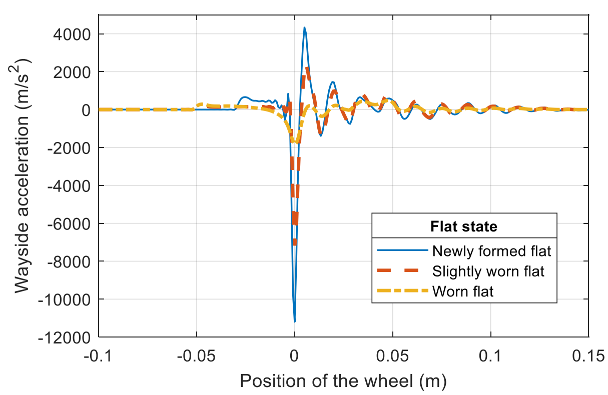

Figure 4. WA due to flats at different stages of its evolution.

3.2. Influence of vehicle speed and impact position in simulated WA

The role of different operational conditions (vehicle mass, speed and position of the impact) in the resulting WA was studied Figure 5 and Figure 6 show the impact of vehicle speed and impact position, respectively (it has been checked that the influence of the vehicle mass is much lower). The higher the speed the higher the peak magnitude due to the impact is. The impact position also influences the peak magnitude, being the highest value associated with impacts occurring at the position where the accelerometer is located. In general, peak magnitude is higher when impact occurs between supports, and lower when it takes place above a support. The influence of these operational conditions should be considered when defining the features for the detection and analysis of flats.

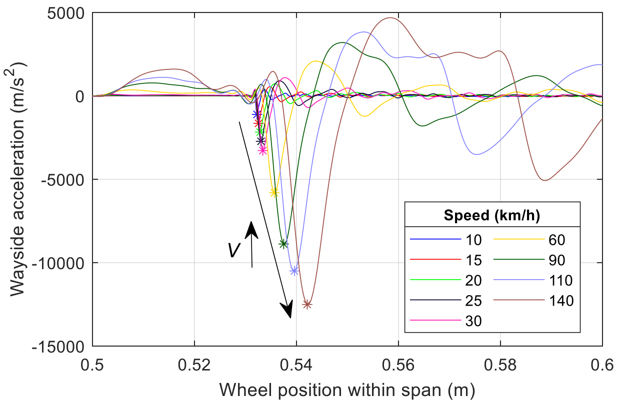

Figure 5. WA change due to the vehicle speed.

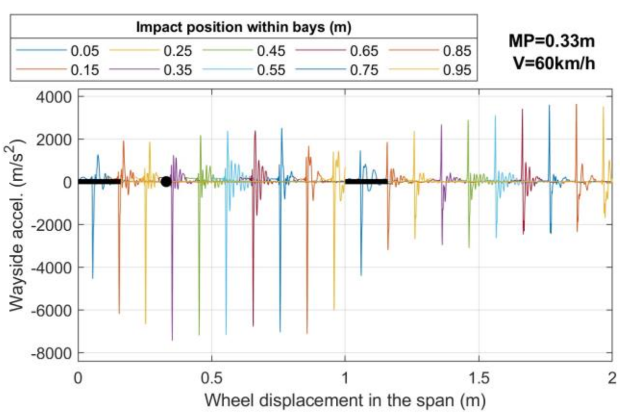

Figure 6. WA change due to the impact position.

4. FEATURES FOR DETECTION AND ANALYSIS OF FLATS

By using the simulated WA signals, features for the detection and analysis of flats were defined. Some of the signal features that present a significant dependency on flat characteristics, like depth and evolution stage, are:

- WA peak value
- WA root-mean-square (RMS) value
- RMS value of the WA band-pass filtered around the first pinned-pinned frequency (from now on denoted as RMS_{1pp}).
- RMS value of the WA band-pass filtered around the second pinned-pinned frequency.

Pinned-pinned frequencies are characteristic features of the track caused by the support periodicity. The values of the 1st and 2nd pinned-pinned frequencies are determined by the frequencies for which the rail vibrates with wavelengths that are twice the distance support and the distance support, respectively. In Figure 7 a qualitative representation of the modal shape of these resonances is presented.

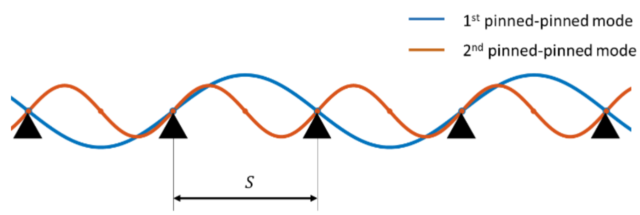

Figure 7. Qualitative representation of the first and second pinned-pinned resonances.

As an example of the features dependency on the flat shape, Figure 8 shows RMS_{1pp} versus the impact position for impacts with different flat depths. The influence of the flat shape, in this case, its depth, is clear with higher magnitude of this feature as the depth increases. Nevertheless, the impact position has also a dominant role. Strong dependencies between the features and vehicle speed are also detected. This makes necessary a normalization of the features to subtract the effect of these operational conditions.

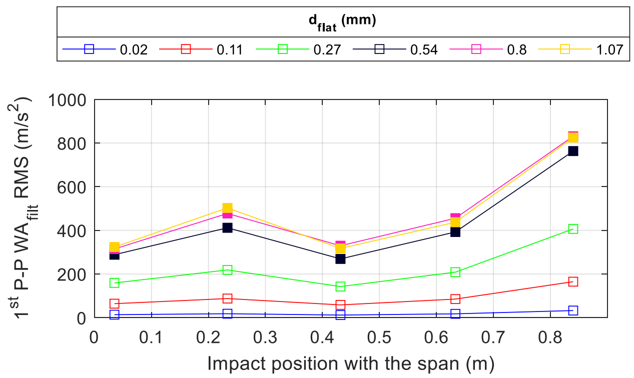

Figure 8. $\mathbf{RMS_{1pp}}$ versus the impact position and for different flat depths.

Correction functions remove the dominant role of operational parameters. Following, the normalization of $\mathrm{RMS_{1pp}}$ in relation to the impact position is detailed. The correction function is obtained by normalizing curves in Figure 8 to make the energy of each curve unitary, which results in the curves presented in Figure 9. The resulting curves are very similar, therefore, their shape is independent of flat depth. It was proved that these curves are weakly coupled with flat shapes and vehicle speeds. This fact can be used to define a function that determines a correction factor for the impact position. The resulting correction function for $\mathrm{RMS_{1pp}}$ vs. impact position is shown in Figure 10. This function is scaled to make its minimum value, which in this case is at 0 m, equal to one. Figure 10 also shows the correction functions for two very different flat profiles, from which it is noticed that the correction functions do not differ significantly. This function is used by dividing the raw value of $\mathrm{RMS_{1pp}}$ by the function value at the position where the impact occurs.

For the sake of clarity, correction functions for the other three features defined in this section are not shown here but they are calculated following a similar procedure. Moreover, the correction functions for the speed and the vehicle load are also obtained.

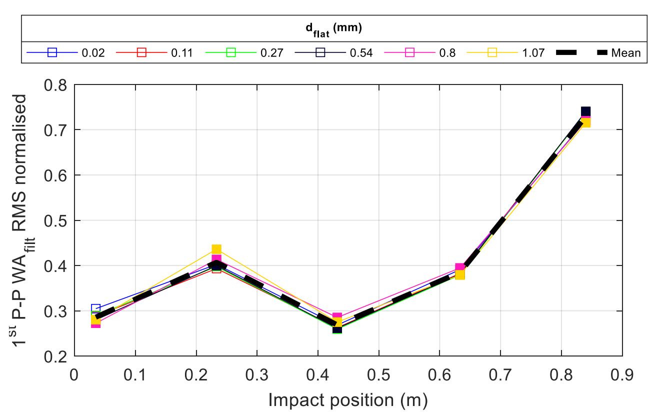

Figure 9. Normalized $\mathbf{RMS_{1pp}}$versus the impact position and for different flat depths.

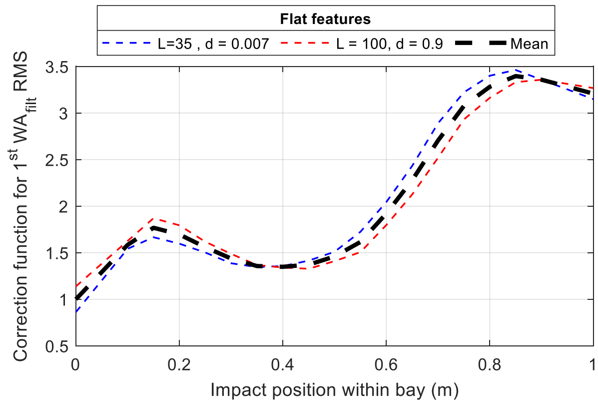

Figure 10. Correction function for $\mathbf{RMS_{1pp}}$ of the impact position.

Once the features are corrected, a severity parameter is defined as their weighted sum. In this work, weights are adjusted after the first series of field measurements. The value of each weight will depend on different factors, like the distance between track supports and it is based mainly on the operator requirements (severity parameter is related to maintenance needs). The severity parameter does not refer, therefore, to any physical magnitude.

However, the theoretical model also provides the contact force response. The monitoring strategy can take advantage of it by labelling the impacts with the predicted maximum contact force, F_{max}, of each flat. In this case, the label is the load ratio, F_{max}/F_{sta} , where F_{sta} is the static contact force. By using the corrected features of the WA and the labeled data, ML regression models can be trained for prediction of F_{max}/F_{sta}.

Several ML models were trained by using simulated data encompassing a wide range of flat profiles and operational conditions. The Gaussian Process Regression algorithm is the one resulting in the lowest root-mean-squared error for the estimation of F_{max}/F_{sta} , 0.043. Figure 14 shows the predicted response vs. the true response of the F_{max}/F_{sta} for the mentioned ML regressor.

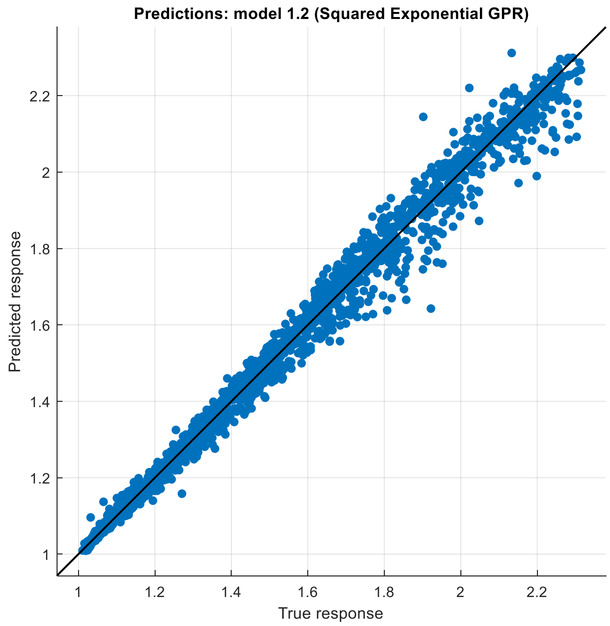

Figure 11. Predicted load ratio values versus the true one.

5. SYSTEM VALIDATION IN A REAL ENVIRONMENT.

A wayside system has been installed in a real track with the objective of validating the proposed methodology to detect flats. **¡Error! No se encuentra el origen de la referencia.** shows a layout of the wayside measurement site: it consists of three accelerometers per rail located in consecutive spans.

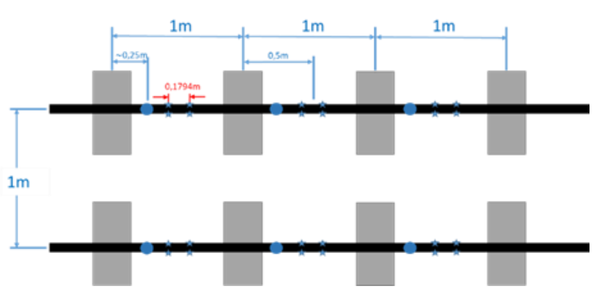

Figure 12. Layout of the wayside measurement site

For the methodology validation, a vehicle with a small flat has circulated over the wayside system several times. It has been checked that the developed algorithm has always detected the flat. Additionally, it has been seen that if the severity of the flat is just obtained with the acceleration peak very different results are obtained for each passage. Figure 13 shows the peak value of the acceleration measured in the first 20 passages. It can be checked that this parameter has a large variability: it goes from around 60m/s2 to 120m/s2. The differences can be found in the fact that each passage has a different speed, different payload and the impact is produced in different points.Figure 13Figure 13

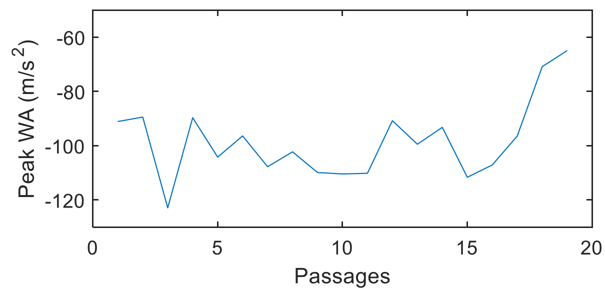

Figure 13. Peak acceleration values caused by the flat over different passages.

In order to get more stable results, the severity parameter introduced in the previous section is calculated. As explained this severity parameters consider not only the peak value but the RMS in the whole frequency range and the signal energy around the peak resonances. Figure 14 includes the severity results obtained for the previous flat (blue crosses). The results now are more stable than when considering just the peak value. Finally, the correction factors have been applied for all the variables that can modify the phenomena (impact point, speed…) and the severity factor is calculated with these corrected values (Figure 14, red squares).

When the correction is applied, the impact position dependence becomes considerably lower, and the variability is largely diminished (due to the speed and load corrections). There is still some variability caused by different reasons, such as the wear of the flat in time, measurement and discretisation errors or not totally accurate severity parameter corrector functions.

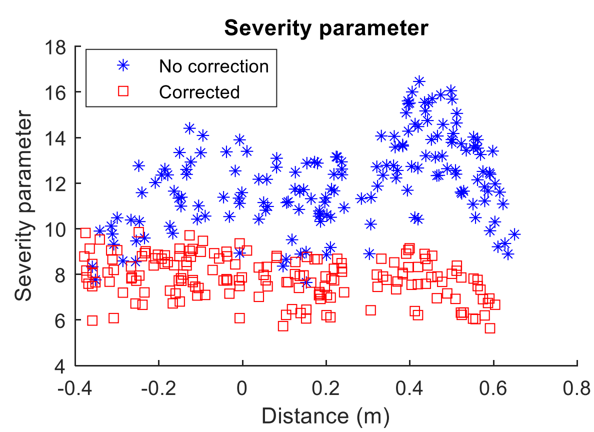

Figure 14. Measured severity parameter calculated combining the peak, the RMS and the RMS for a flat.

Finally, an estimation of the impact force caused by the flat is obtained by applying the developed Machine Learning model. Figure 15 shows the peak force estimation in the different passages as a function of the impact point. The following considerations can be done:

- The peak force estimation is quite constant for the different passages being a minimum of around 1.00 and a maximum of around 1.03.
- The analyzed flat is extremely small. It only causes an increase of around 2% of the nominal force. This indicates the high sensibility of the developed methodology.
- The estimation of the peak force gives a valuable indicator of the severity of the flat easing the maintenance decision-making.

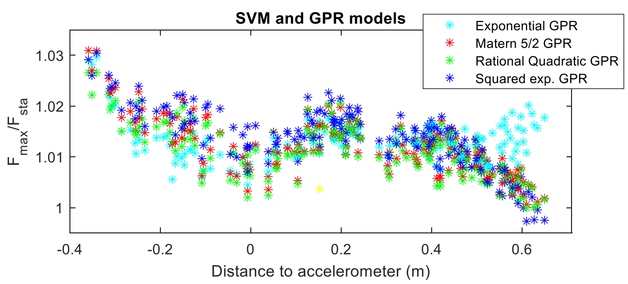

Figure 15. Estimation of peak force with respect to the nominal one.

6. CONCLUSIONS.

This work presents the design and validation of a new wayside detection system for wheel defects. More generally, the aim is to show that the combined use of the wealth of existing knowledge on physics-based models of complex phenomena can be efficiently combined with new machine learning techniques.

This combination of physical modelling and data-based techniques has to be done in a particular way in order to benefit from the best of both approaches. In this particular work:

- Physical models and a deep understanding of the phenomena allow defining optimized sensor systems to develop HMS and guarantee the interpretability of the results.
- Results indicate RMS value of the WA band-pass filtered around the first pinned-pinned frequency is an important feature in assessing impact severity.
- Features of acceleration measurements are highly influenced by speed and impact position. Correction functions obtained via numerical model enable the standariazaiton of the features.
- The use of key parameters and correction functions for the contour variables (mass, speed …) allow obtaining an estimation of defect severity independent of the operational conditions.
- Machine Learning algorithms are key to adjusting the correction parameters in a very efficient way: considering the high complexity of the relationships between magnitudes, machine learning techniques offer a way to reach optimum results.
- Stable prediction of the contact force peak value due to wheel flats is enabled by ML algorithms in combination with defined corrected features of the rail accelerations. Theoretical results point out a RMSE of less than 5%.
- The complete methodology has been checked in a real environment (metro track) giving promising results.

The main limitation of the given approach is that the wayside flat detection algorithm should be tuned for each different track it is installed. Future work will dive into the experimental validation of the impact force estimation performed via ML algorithms using rail accelerations.

ACKNOWLEDGEMENT

The authors would like to acknowledge the final support to perform this work through the grants Project 101101973 — FP1 - MOTIONAL and PID2020-118013RB-C21 (funded by MCIN/AEI/10.13039/501100011033).

REFERENCES

Alemi, Alireza, Francesco Corman, Yusong Pang, and Gabriel Lodewijks. 2019. «Reconstruction of an Informative Railway Wheel Defect Signal from Wheel–Rail Contact Signals Measured by Multiple Wayside Sensors». *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 233(1): 49-62.

Blanco, B. et al. 2019. «Implementation of Timoshenko Element Local Deflection for Vertical Track Modelling». *Vehicle System Dynamics* 57(10): 1421-44.

Blanco, B., N. Gil-Negrete, L. Kari, and A. Alonso. 2022. «On the Correction of Rail Accelerations Predicted by Numerical Track Models Based on Timoshenko Beam Theory». *Vehicle System Dynamics* 60(6): 1993-2017.

Gao, Run, Qixin He, Qibo Feng, and Jianying Cui. 2020. «In-Service Detection and Quantification of Railway Wheel Flat by the Reflective Optical Position Sensor». *Sensors* 20(17): 4969.

Iwnicki, Simon, Jens C. O. Nielsen, and Gongquan Tao. 2023. «Out-of-Round Railway Wheels and Polygonisation». *Vehicle System Dynamics* 61(7): 1785-1828.

Jergéus, J. 1998. «Martensite Formation and Residual Stresses around Railway Wheel Flats». *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 212(1): 69-79.

LBFoster. «WILD IV Product Brochure». https://m6g3q6n5.stackpathcdn.com/perch/resources/wild-mk-iv-product-brochure.pdf (3 de julio de 2023).

Maglio, Michele, Astrid Pieringer, Jens C.O. Nielsen, and Tore Vernersson. 2021. «Wheel–Rail Impact Loads and Axle Bending Stress Simulated for Generic Distributions and Shapes of Discrete Wheel Tread Damage». *Journal of Sound and Vibration* 502: 116085.

Mazilu, Traian. 2007. «Green's Functions for Analysis of Dynamic Response of Wheel/Rail to Vertical Excitation». *Journal of Sound and Vibration* 306(1-2): 31-58.

Mosleh, Araliya et al. 2023. «Early Wheel Flat Detection: An Automatic Data-Driven Wavelet-Based Approach for Railways». *Vehicle System Dynamics* 61(6): 1644-73.

Mosleh, Araliya, Pedro Montenegro, Pedro Alves Costa, and Rui Calçada. 2021. «An Approach for Wheel Flat Detection of Railway Train Wheels Using Envelope Spectrum Analysis». *Structure and Infrastructure Engineering* 17(12): 1710-29.

Nielsen, JCO, JW Ringsberg, and L Baeza. 2005. «Influence of railway wheel flat impact on crack growth in rails». En *Proceedings of the Eighth International Heavy Haul Conference (IHHC8) in Rio de Janeiro*,.

Schenck process. «MULTIRAIL Improving safety and reliability in rail processes». https://www.schenckprocess.com/download?id=139&lang=en&pid=123 (3 de julio de 2023).

Stratman, Brant, Yongming Liu, and Sankaran Mahadevan. 2007. «Structural Health Monitoring of Railroad Wheels Using Wheel Impact Load Detectors». *Journal of Failure Analysis and Prevention* 7(3): 218-25.

voestalpine. «WDD/WIM wheel impact load detection». https://cdnstorevoestalpine.blob.core.windows.net/documents/792615/original/railwaysystems_factsheet_WIM-WDD_en.pdf (3 de julio de 2023).