

# **Wheel Damage Prediction Using Wayside Detector Data for a Cross-Border Operating Fleet with Irregular Detector Passage Patterns**

Johan Öhman, Wolfgang Birk,  
Jesper Westerberg

# Wheel damage prediction using wayside detector data for a cross-border operating fleet with irregular detector passage patterns.

Johan Öhman  
Predge AB  
SE-97236 Luleå, Sweden  
+46 725 09 43 49  
johan.ohman@predge.se

Wolfgang Birk  
Predge AB,  
SE-97236 Luleå, Sweden  
+46 725 390 909  
wolfgang.birk@predge.se

Jesper Westerberg  
Predge AB  
SE-97236 Luleå, Sweden  
+46 706 878 691  
jesper.westerberg@predge.se

## ABSTRACT

Wheel damages on railway vehicles caused by rolling contact fatigue or blocked wheels can cause severe problems for railway operators and infrastructure owners. Wheel impact load wayside detectors (WILD) are one of the means to assess the condition of a wheel in operation, but varying operating routes, irregular traffic patterns, and especially cross-border operations make this quite challenging. While the condition updates occur randomly, the detectors themselves are managed by different owners and principles. Thus, using the same type of data from not only different wayside locations but also different providers and authorities with varying fidelity and operational practices introduces uncertainties in data quality and consistency.

This paper presents an approach for predicting wheel damage severity on a wagon fleet with irregular cross-border operations, achieving similar confidence levels as for regular traffic patterns on a national scale. The different sensor characteristics are explored between countries and within each country. The approach is implemented as a cloud-based solution which integrates wayside detector data from multiple locations provided by two different infrastructure owners in two countries. The solution estimates remaining useful life based on data from both countries and aggregates this to a single indication for the decision maker. The algorithm's performance is showcased for vehicles with cross-border operations. The results indicate that the proposed approach confirms that irregularly provided measurement data with data quality and consistency issues are manageable and adequate decision-making performance.

## Keywords

Wheel damages, railway, condition-based maintenance, predictive maintenance, wayside detectors, data fusion, statistical fusion

## 1. INTRODUCTION

The efficiency of railway operations depends on a well-working infrastructure and a minimal number of disruptions of operations. Wheel damages are one of the factors that can cause harm to the infrastructure, derailments, stops in line and acute shunting of wagons in operation. These damages are often caused by rolling contact fatigue or blocked wheels. Wayside detectors measuring the impact forces between rail and wheel aid in determining damage severity during the passage in normal operation on a main line, as pointed out by Alemi et al. (2017), who also indicates that further research needs to be conducted to consider multiple detector

stations. Clearly, if such detectors are distributed in the infrastructure, they can be used to capture growth in forces for individual wheels and to characterise growth patterns enabling condition-based and predictive maintenance approaches. High detector coverage on the operating routes would then be required, but irregular traffic patterns and especially cross-border operations are challenging boundary conditions. Integrating and synchronising different data types from multiple sources and the same type of wayside data from multiple locations are common tasks today, as shown by Birk et al. (2019). But using the same type of data from not only different wayside locations but also different providers and authorities with varying fidelity and operational practices introduces data quality and consistency uncertainties. Subsequently, the design of automated decision-making algorithms achieving consistent confidence levels becomes a more complex task.

In this study, data fusion from multiple wayside detector stations located in two different countries measuring wheel impact forces is used to detect damages and predict wheel damage severities. While this is in line with the proposed research tracks in Alemi et al. (2017), there are numerous challenges that need to be considered, which are already shortly summarised by Karim et al. (2015).

In previous research, there is a consensus that the wheel-to-rail impact forces due to wheel defects are dependent on passage speed and severity of the damages, as shown in studies based on actual field tests as well as simulations of theoretical models, Maglio (2022). Therein and claimed by Dong & Shankar (1994) and by Steenbergen. et al. (2004), the shape and size of the wheel defects, axle load, train speed, and contact patch stiffness are factors that influence impact loads the most. Further, the impact of the ambient condition is investigated and reported by Olofsson and Sundvall (2004).

Moreover, Kalay et al. (1995) investigated impact loads from wheel flats as a function of train speed in the interval 30-100 km/h. The results show a slight increase in peak impact force with increased speed for short (25-40 mm) and long (75-100 mm) wheel flats.

The results in Bogdevicius et al. (2016) also show that, for all-wheel flat lengths, the increase in impact force slows down at speeds over 80 km/h, which is in contradiction to the field tests made by Nielsen & Johansson (2000) with a wheel flat of length 100 mm and depth 0.9 mm showed a negative correlation between impact force and speed in the investigated speed interval, 30-100 km/h. A further challenge for a multiple detector station fusion is the local character of a damage and that there can be multiple damages with varying causes present both along the circumference

and laterally, meaning that damage might not show consistently increased wheel impact forces at all detector stations.

To mitigate the effect of the above challenges, an analytics scheme is suggested that can compensate for the above-mentioned effects and harmonises the detector data prior to the data fusion, making the subsequent predictive maintenance algorithms resilient to fluctuations in the reported wheel impact forces and able to detect wheel damages and forecast the future damage severities.

The paper is organised as follows. First, the problem is more formally outlined, and then the measurement characteristics are discussed in the following section. The analytics approach to harmonise and fuse the data is then presented. Thereafter, the case study is presented where the approach is applied and compared with the naïve approach of using the data with a data harmonisation. The paper ends with some conclusions and an outlook on future work.

## 2. PROBLEM OUTLINE

A train operator is operating a cross-border fleet in two countries: Sweden and Norway. During regular operation, the fleet passes several wheel impact load detectors (WILD), Swedish Transport Administration (2013). These sensors measure the force from the wheels to the rail during the passage. The result from a passage is the mean force and the peak force for each wheel. Denote the measurement of a wheel  $w_i$  at location  $j$  and time  $t$  with

$$z_{i\text{PEAK}}^j(t) = f(t; w_i, \theta_j), \quad (1)$$

$$z_{i\text{MEAN}}^j(t) = f(t; w_i, \theta_j), \quad (2)$$

where  $\theta_j$  are the location-dependent parameters. The measurement function  $f$  is stochastic, so the same wheel state and the parameters will generally not reproduce the same measurement. The time  $t$  is the large-scale time typically in hours. The difference in the sampling parameters depends on the WILD type, its calibration, and the track's location. For instance, there will be a difference if the WILD is placed on a straight track or in conjunction with a turn. In most cases, the WILD detectors are placed on a straight track. There are two different governing agencies in the two countries, Trafikverket in Sweden and Bane NOR in Norway. These agencies are responsible for installing, calibrating and operating the detectors within each country. To an external part, the calibration and maintenance actions are unknown, and we do not know the status of the detectors. Another aspect that can vary between measurement locations is the WILD's measurement technique. Common configurations are either based on strain gauges or fibre optics (Kouroussis et al., 2015).

We assume that there are differences between detectors in the two countries and differences between individual detectors within each country. Historically, the processing for each country has been made individually. Joint processing is preferable since all data influence decision-making in conjunction.

The mean value mainly depends on the wagon load, while the peak value depends on the wheel and loading conditions. It is preferable if the measurements are independent of the loading conditions. The dynamic force is calculated as the difference between the peak and mean force

$$z_{i\text{DYN}}^j(t) = z_{i\text{PEAK}}^j(t) - z_{i\text{MEAN}}^j(t), \quad (3)$$

making it independent of the loading conditions. For the rest of the paper, we will exclusively use the dynamic force and drop the DYN notation from the variable.

Given a set of historic force measurements from different locations, we would like to estimate the current force level and predict the time until maintenance is required. The different characteristics of different locations will influence the outcome of the prediction. Predge's algorithm estimates the current force level and predicts future trends. Denote this algorithm function by  $g$ . Then the processing is described by

$$[\hat{z}_i, h] = g([z_i^j(t_0), z_i^j(t_{-1}), z_i^j(t_{-2})]) \quad (4)$$

where  $\hat{z}_i$  is the estimated current force level, and  $h$  is the time estimate until a maintenance action is required. Note that  $i$  is constant, but  $j$  varies in the equation above. Different locations may have different characteristics. These differences are further explored in the next section.

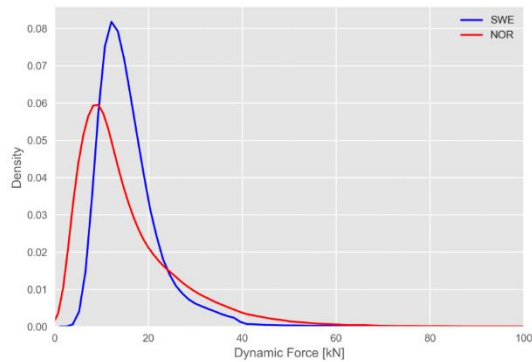
## 3. MEASUREMENT CHARACTERISTICS

If all measurements in both countries are included, we can look at the group statistics of both countries. We are limiting the analysis to one wagon type, *Sggmrs*. The data is from the time period of 2022-01-01 to 2022-04-30 and includes wagons in both Sweden and Norway. In total, 396 wagons and 263548 detections are included in the analysis. Start by comparing the group statistics in both countries. Figure 1 shows the distributions of detections in Sweden and Norway. There is a skewed distribution in both countries where lower force detections are most common. This behaviour is expected since the wheels operate correctly for longer periods than in a damaged state. The mode of the Norwegian data is lower than in Sweden, but in Norway, more data is located in the right tail. The low values are, in reality, uninteresting as long as they are below approximately 25kN, so it is the data in the right tails that are of interest. It is this data that is interesting for estimating the wheel condition. It is noted that there are some minor differences between countries for values above 25kN

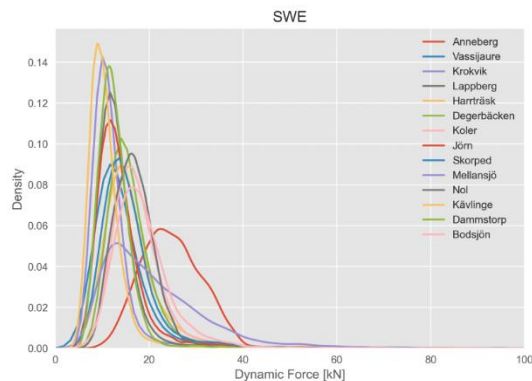
If the data is further split into measurement locations in each country as depicted in parts (b) and (c) of Figure 1, it is clear that the changes within a country are equal or greater than the changes between countries. In Sweden, two locations deviate from the majority. One of the locations also deviates in the shape of the pdf, having a more symmetric and gaussian-like profile. There are only five WILD detectors in Norway on the operated routes. Two deviate from the others, as shown in part (c) of Figure 1. The data shows that differences within each country are already on the same scale as those between countries. We will approach the problem twofold:

1. Since differences already exist within the countries, joint processing could be done by running all data without any correction.
2. Normalise the data based on the group statistics of each location before performing the estimation using equation (4).

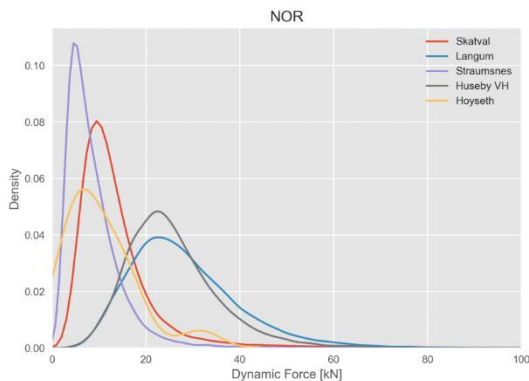
The second approach requires some assumptions that the difference between the measurement locations depends only on the location's characteristics, not on any external source. Using normalisation, the data is harmonised into a joint representation. From the joint representation, the data can be transformed into the configuration of a specific location.



(a)



(b)



(c)

**Figure 1. Distributions of detections in Sweden and Norway. Part (a) shows the histogram with the data divided into two groups. Parts (b) and (c) show the data further split into each country's locations.**

### 3.1 Data Normalisation

From the characteristics shown in Figure 1, the detections at a single location can be assumed to have a log-normal distribution and parametrised by the mean  $\mu$  and standard deviation  $\sigma$ . The following pdf models the data for each location

$$f(z) = \frac{1}{\sigma z \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\ln(z) - \mu)^2\right), \quad (5)$$

where the distribution is given in a logarithm scale. Using this parametrisation, the location parameters  $\theta_j = [\mu_j, \sigma_j]$  describes the distribution of each location. These parameters are calculated for each location by minimising the negative log-likelihood function. Given a detection and the location parameters, the value can be transformed into a normal  $N(0,1)$  distribution in the logarithm scale using

$$\log z'_j = \frac{\log z_j - \mu}{\sigma} \quad (6)$$

If one location is chosen as the reference, the detections from other locations can be mapped to that location's characteristics.

$$z'_j = \exp\left[\frac{\log z_j - \mu}{\sigma} \sigma_{ref} + \mu_{ref}\right], \quad (7)$$

where  $\sigma_{ref}$  and  $\mu_{ref}$  are the parameters at the reference location.

## 4. Case Study

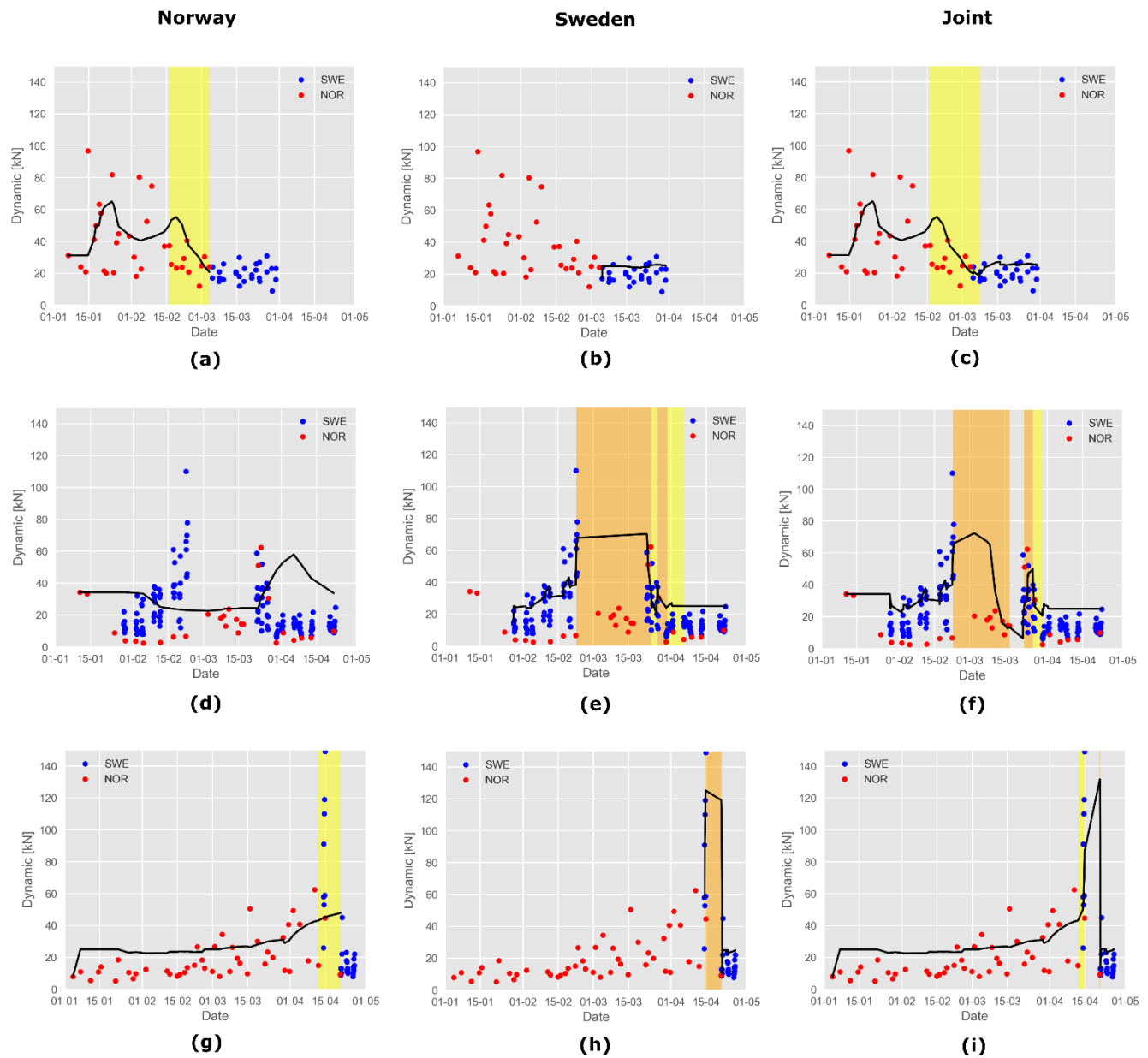
To showcase the different processing methods suggested, data for three different wheels are processed in four different ways:

1. Using only the Swedish data.
2. Using only the Norwegian data.
3. Using the data from both countries.
4. Using the normalised data from both countries.

For all four cases, the data is processed using equation (4). The result is in the form of an estimated force level and potential warnings. Depending on the severity, the warnings are on two different scales, yellow and orange. Figure 2 shows the result for the first three processing methods. Each row is a wheel, and each column is a processing method. The first column uses only Norwegian data, the centre uses only Swedish data, and the right column uses both. The solid black line is the estimated current force level representative of the damage severity.

There is no apparent damage to the first wheel on the first row, as can be seen by the decreasing force. An incorrect warning is generated in the processing using the Norwegian and combined data due to the consistent moderate-force measurements. This warning is later removed when the Swedish data is fed to the algorithm. Of course, there is no incorrect warning when using only Swedish data since it is solely based on Norwegian data.

For the second vehicle on the second row, there are two sections with increasing force levels. For the first increase, there is data from both Norway and Sweden. However, the increase is only present in the Swedish data. So, in part (d), there is no warning in processing only Norwegian data, and the estimated force level is low. An orange warning is raised correctly in both the Swedish and joint processing. After the first increase, the vehicle travelled to Norway, and the damage was corrected. For joint processing, the estimated force level decreases. Since this period only uses Norwegian data, the Swedish processing does not note this decrease and stays on a high level. About one month after the first increase, there is a second increase. This time, the Norwegian data also shows an increasing force level. However, Norway has fewer measurements, and the estimated force level increases slowly. The Swedish processing never noticed a decrease between the two peaks and continues to report a high force level and orange warning.

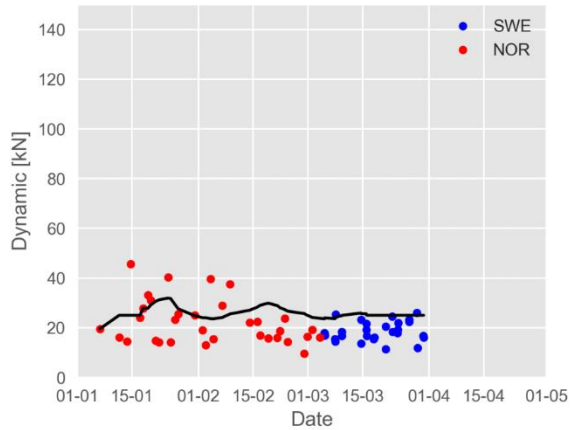


**Figure 2. Example processing of three different vehicles. Each row is a vehicle and each column is a processing method. I.a parts (a)-(c) is one vehicle processed in three different ways, parts (d)-(f) is the second vehicle and parts (g)-(f) is the third vehicle.**

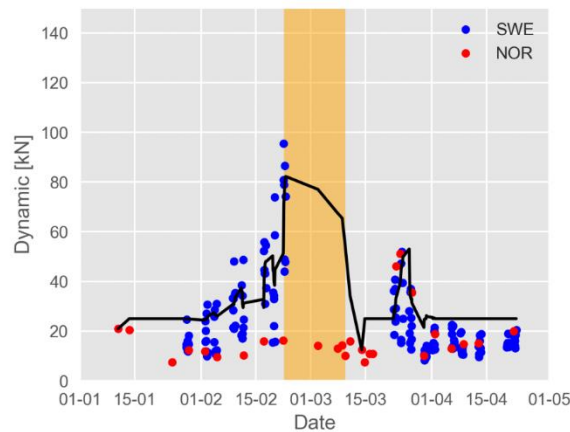
The final wheel has a gradual transition starting in Norway with an increase from low force levels. When it enters Sweden around the fifteenth of April, there is an increase in force. The Norwegian processing does not have these measurements and estimates a lower value. The Swedish processing has no prior information on the wheel, and the high measures skyrocket the estimated force. The combined processing yields a smooth transition from Norwegian to Swedish data in this case.

The final part of this case study is data normalisation, as described in section 3.1. Before equation (4) estimates the wheel state, the data is normalised using equation (7). The processing is evaluated for the same three wheels but only with the joint processing. Figure

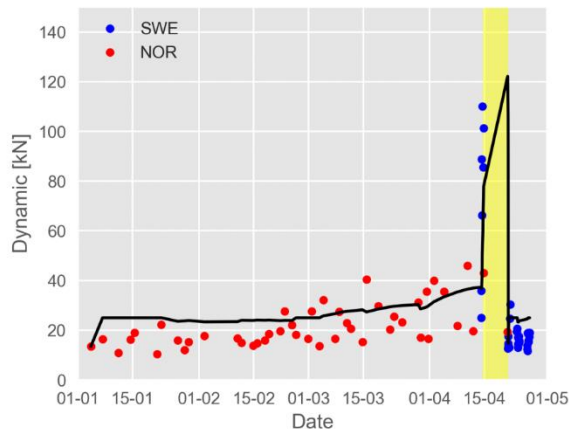
3 shows the results of this processing. This Figure should be compared to the right column in Figure 2. The first wheel values in part(a) are drastically lower than in the corresponding non-normalised case in Figure 2c. The value reduction is due to a high mode and standard deviation for the Norwegian data. With the normalisation, the values are more in line with the Swedish results. On this data, there are no warnings, which is the desired outcome. The changes are minor for parts (b) and (c) compared to the corresponding data in Figure 2. In part (b), a warning is raised correctly for the first increase, while the second is missed. In part(c), the wheel performs similarly to the normalised and non-normalised data.



(a)



(b)



(c)

**Figure 3. Processing of normalised data for the same vehicles as in Figure 2. Part (a) corresponds to part (c) in Figure 2, part(b) corresponds to part(f) in Figure 2, and part(c) corresponds to part(I in Figure 2.**

## 5. Discussion & Conclusions

Although most warnings are captured with Swedish or Norwegian processing, joint processing has advantages. If damage is initiated in one country and the vehicle travels to the other, it will appear as a step-function-like increase, triggering a more severe response. This effect is seen in parts (g) and (h) of Figure 2. The joint processing in part (i) offers a smoother transition between the data in the two countries. The joint processing seems to have no inherent drawback since all data better reflect the wheel state. This is likely due to the already existing differences in location character within each country, as depicted in Figure 1.

The normalised data in Figure 3 offers some additional benefits. The noisy data for the first row in Figure 2 comes from a detection site with high variance in the detections. The high variance is compensated for in Figure 3a, and no incorrect alarm triggers. It is important to note that all algorithm processing has been made with the same parameter settings. Future work will include finding optimal parameters for each processing method. Unoptimised parameters could be a reason for the miss of the second increase in Figure 3b.

This paper only showcases the behaviour of three wheels. The wheels chosen are representative of the data in the two countries. However, a larger dataset with associated evaluation metrics is required to fully benchmark the performance gain with the joint and normalised processing.

Using normalised data comes with a risk. The data may represent the wheel state better, but the infrastructure owner will still act on high measurements regardless of its group statistics. Therefore, all stakeholders must know what is being done and what risk is involved.

## 6. Acknowledgement

The authors thank Cargonet for providing the data for the analysis.

## 7. References

- Alemi, A., Corman, F., & Lodewijks, G. (2017). Condition monitoring approaches for the detection of railway wheel defects. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 231(8), 961-981.
- Birk, W., Dittman, T., Karim, R., Westerberg, J., (2019), Experiences from the detection and prediction of wheel damages on railway vehicles in operation. In Proceedings of the 13th International Heavy Haul Association STS Conference, June 10-15, Narvik.
- Bogdevicius, M., Zygiene, R., Bureika, G., & Dailydka, S. (2016). An analytical mathematical method for calculation of the dynamic wheel-rail impact force caused by wheel flat. Vehicle System Dynamics, 54(5), 689-705.
- Dong, R. G. & Sankar, S. (1994). The characteristics of impact loads due to wheel tread defects. RTD Rail Transpn ASME, 8, 23-30.
- Johansson, A. & Nielsen, J. O. (2000). Out-of-round railway wheels – wheel-rail contact forces and track response derived from field tests and numerical simulations. Proc IMechE, Part F: J Rail and Rapid Transit, 217, 135-146.

- Kalay, S., Tajaddini, A., Reinschmidt, A. & Guins, A. (1995). Development of performance-based wheel-removal criteria for North American Railroads. In Proceedings of 11th International Wheelset Congress, 227-233.
- Karim, R., Birk, W., & Larsson-Kräik, P. O. (2015). Cloud-based emaintenance solutions for condition-based maintenance of wheels in heavy haul operation. In International Heavy Haul Association: The 11th International Heavy Haul Association Conference will be held 21-24 June 2015 in Perth 21/06/2015-24/06/2015. International Heavy Haul Association.
- Kouroussis, G., Caucheteur, C., Kinet, D., Alexandrou, G., Verlinden, O., & Moeyaert, V. (2015). Review of Trackside Monitoring Solutions: From Strain Gages to Optical Fibre Sensors. *Sensors (Basel, Switzerland)*, 15(8), 20115–20139.
- Maglio, M., Vernersson, T., Nielsen, J. C., Pieringer, A., Söderström, P., Regazzi, D., & Cervello, S. (2022). Railway wheel tread damage and axle bending stress—Instrumented wheelset measurements and numerical simulations. *International Journal of Rail Transportation*, 10(3), 275-297.
- Olofsson, U., & Sundvall, K. (2004). Influence of leaf, humidity and applied lubrication on friction in the wheel-rail contact: pin-on-disc experiments. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 218(3), 235-242.
- Steenbergen, M. J. M. M. (2007). The role of the contact geometry in wheel-rail impact due to wheel flats. *Vehicle System Dynamics*, 45(12), 1097-1116.
- Swedish Transport Administration, (2013). BVF 592.11 – Detektorer. Hantering av larm från stationära detektorer samt åtgärder efter upptäckta skador vid manuell avsyning. (TDOK 2014:0689).