Leveraging Machine Learning to Predict Rail Corrugation Level from Axle-Box Acceleration Measurements on Commercial Vehicles

Wael Hassanieh[a], Abdallah Chehade[a], Alan Facchinetti[b], Mark Carman[c], Marco Bocciolone[b] and Claudio Somaschini[b]

[a]Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn, Dearborn, Michigan, USA; [b]Department of Mechanical Engineering, Politecnico di Milano, Milan, Lombardy, Italy; [c]DEIB, Politecnico di Milano, Milan, Lombardy, Italy

ABSTRACT
Rail corrugation is a prominent degradative problem in the health monitoring of railway systems. Monitoring process is dependent on use of a diagnostic trolley, which is expensive and needs the track to be out-of-service. Alternatively, in-service rail vehicles with Axle-Box Acceleration measurement systems installed, have shown success in detecting rail corrugation levels based on physical models, albeit with limitations. Extending this approach, we build a Machine Learning model, represented by a tuned Random Forest regressor, trained on collected accelerometer signals along with other offline and/or static features. We also propose a method to engineer acceleration-based features which nullifies the aggregated acceleration vibrations inherited from the other rail due to dynamically coupled vibrations between the left and right rials. The resulting model is able to recreate the moving RMS irregularity profile at bandwidth 100-300 mm, especially in highly corrugated sections, with an R^2 score of 0.97-0.98. The results show that the suggested data-driven approach outperforms a state-of-the-art model-based benchmark.

1. Introduction

One of the most prominent problems in the railway industry is rail corrugation. It is a degradative phenomenon in which rail lines exhibit a cyclic wearing process due to rail-wheel interactions. It is demonstrated by quasi-sinusoidal irregularities of wavelengths less than one meter along the rail's running surface; in Figures 1a & 1b, we can spot

CONTACT Wael Hassanieh. Email: waelh@umich.edu

(a)   (b)

Figure 1.: Rail corrugation demonstrations on sections of rail from a metropolitan rail network.

on the individual rails corrugations of different wavelengths. The persistence of corrugations paves the way for further high dynamic loading, which leads to unfavourable noise pollution, vibrations transmitted to nearby buildings and infrastructure, discomfort to rail passengers and nearby inhabitants, and increases the probability of vehicle and track component failure.

Rail corrugation is a problem that is substantially understood in the railway literature. An important source is by Grassie and Kalousek [1] who classified different corrugation types according to the wavelength of the irregularity, the corrugation locations, its appearance on the rail running surface and the main cause of its formation. Corrugation forms as a result of a 'wavelength-fixing' mechanism, that occurs when the original longitudinal rail profile, which exhibits vibrational components at all wavelengths, excites the dynamics of the vehicle–track system. This mechanism produces certain dynamic loads that eventually 'fix' at certain wavelengths and certain positions of the formed corrugations along the rail lines. These dynamic forces will instigate changes and damages on the rail profile; this process is called 'damage mechanism'. With the recurring train-track interactions, a positive feedback effect occurs on the degradation process of the rail profile. This phenomenon is further aggravated at certain wavelengths especially if the rail vehicles passing over a specific track section share similar architecture, speed and dynamics.

With the plethora of information regarding the characteristics and cause of corrugation formation, there has also been available a diverse range of treatment processes [see 2]. They include: change into 'hard' rails, control of friction, and 'improved steering'. However, rail corrugation's most effective and widely used treatment remains the application of periodic grinding - also called rail profiling - over the rail's damaged surfaces. This maintenance operation is done via the use of dedicated vehicles on the corrugated parts of the track, only when the track is available and out-of-service. Moreover, with the substantial cost associated with this operation, it becomes more important to schedule conditional maintenance operations upon detection of high irregularity levels that ought to be re-profiled.

2

Current best practice for tackling the problem of rail corruption level detection is for railway line operators to run a diagnostic trolley periodically over the line to gather measurements of the rail profile, which are then used for determining a schedule for performing the grinding process on the line. Given the ongoing trend to instrumentalise public transport vehicles on such networks, there is the enticing possibility to make use of sensors (positional encoders, gyroscopes and accelerometers) on commercial vehicles to monitor the status of the line. As suggested by Grassie [3], the Axle-Box Accelerometer (ABA) system - connected to an in-use rail vehicle - is one of the more robust techniques to recreate the rail profile via a model-based transfer function. Work done by Karaki et al. [4] shows that the ABA system can produce encouraging results in monitoring rail corrugation, albeit with some difference between the system's estimated signal and the true measured signal. On the other hand, the use of data-driven techniques and machine learning tools for predictive maintenance in railway engineering has garnered increasing interest over the past few years [see 5, 6, 7]. Accordingly, there's motivation to attempt to take advantage of the ABA use and minimise its shortcomings using machine learning techniques.

In this work, we develop a Machine Learning approach to estimate rail irregularity profiles using the multi-sensor data signals produced by an ABA system installed on an in-service rail vehicle of a Milan metropolitan line managed by Azienda Trasporti Milanesi S.p.A. (hereafter ATM). This approach would result in significant advantages over the current best practice:

- Reduced costs for measurements: there is no need for the use and maintenance of a diagnostic measurement vehicle, along with expert workers or overnight operators to collect measurements.
- More frequent measurements: fleet vehicles pass multiple times per day over the same lines, rather than once every few months for typical maintenance measurement schedule.
- Possibility to schedule conditional maintenance instead of random periodical maintenance.
- Ability to track evolution of corrugation in real-time.
- Ability to identify the main factors (time of run, speed etc.) affecting the evolution of the corrugation at certain locations.

With the goal of detecting corrugation formation on the rail lines, it's imperative to define the scope of our work. Rail corrugation is a particular problem for underground rail systems due to the tight turns that such systems have, especially with their need to follow the above-ground road network. For this reason, in this work we concentrate our efforts on building a learning model based on a underground metropolitan railway line in Milan. Moreover, we focus on re-creating the moving RMS of the irregularity profile on the running surface of the rail without the use of a measurement diagnostic vehicle. Accordingly, the scope of the problem is the ability to correlate the input, represented by on-board measurements of the moving vehicle, to the output, represented by the irregularity profile measurements provided. From there, the purpose is to be able to obtain a learning model with high credibility and accuracy to predict the rail irregularity measurements, solely from the Axlebox Accelerometer System, and consequently provide a predictive detection tool for corrugation.

2. Related Work

We now discuss general corrugation detection techniques and the use of machine learning techniques in railway systems.

2.1. Corrugation Detection Techniques

With the scope of interest for this work being the detection of rail corrugation, we focus on specific techniques that measure rail irregularities. Grassie [3] examines three different techniques: chord-based measuring systems, inertial systems and ABA systems. His work explores the characteristics, benefits and limitations of each technique.

Axlebox Accelerometer systems are not a particularly new technique. Indeed, they have existed since the the 1980s for the purpose of monitoring rail corrugation, e.g. on the UK network [8]. One of the successful applications in the field include the HSRCA system, presented by Grassie [9]. This system is an application of the ABA system mounted on a road-rail track recording vehicle; it produces measurements of longitudinal irregularities on rails at vehicle speeds of 40-60km/h. The HSRCA does not directly produce measurements of the rail irregularity profile; however, it outputs an estimation of the irregularity profile based on acceleration signals.

In addition, Karaki et al. [4] have leveraged the use of the ABA system installed on an in-service rail vehicle in the purpose of continuous monitoring of the rail roughness. In comparison to the HSRCA system [9], Karaki et al. [4]'s system also produces an estimation of the irregularity profile based on collected acceleration signals rather than outputting a direct measurement. This technique's credibility is highly related to the confidence in the dynamics of the train-track interaction. In other words, it is based on an input-output function represented by a linearized model of the vertical wheel-rail interaction in frequency domain, much like the one incorporated by [9], who instead has the ABA system on a dedicated recording vehicle. This Frequency Response Function filters the acceleration signal that the system collects, and then it outputs an estimate of the irregularity profile. This means that this estimation process highly depends on the reliability of the input-output function. Encouraging results were obtained through the ability of the system in recreating the rail-head profile; however, there are certain limitations and small discrepancies in relation to direct rail profile measurements. This is due to the fact that the input-output function neglects some sophistications that include: complex characteristics of the dynamic behaviour of the wheelset, and dynamics coupling of the two rails. Accordingly, in this work, we aim at incorporating the ABA system with the use of a data-drive ML method, instead of a model-based transfer function in frequency domain, to further enhance the performance of recreating the moving RMS of the rail irregularity profile, from which reliable corrugation level detection may be obtained. In Section 6, a comparison is shown between the model-based transfer function output and that of the paper's suggested data-driven method.

2.2. Machine Learning in Railway Systems

ML is an important tool that optimises performances of certain methods through a process of "learning" from a certain set of data [10]. This allows the understanding and analysis of data in order to make discoveries and conclusions in the field of study utilising these methods. In our case, the field of study of interest is railway engineering, where ML techniques are becoming increasingly imperative in processing and analysing relevant track geometry data.

According to Xie et al. [5], publications regarding the use of data-driven models in predictive maintenance of railway track increased from 3.1 per year from 1991-2012 to an average of 40 per year in the last three years. Among the top three issues addressed in the literature regarding the applications of data-driven models was rail geometry irregularity.

One of the particular examples is the work of Martey et al. [11] who have integrated several unsupervised and supervised learning techniques in estimating and predicting track geometry quality. In specific, they compared the different learning regression techniques to predict the wavelength variation of the surface parameter of the rail based on geometrical inspection features collected by a track geometry car. They found that Random Forest regression model produces the best performance based on global RMSE metric. As for Ma et al. [12], they targeted the opposite of what we focus on in this study; they predicted the vertical and lateral body acceleration on high-speed railway using deep learning methods, that include CNNs and LSTMs, from track geometry inspection data.

During the past years, much of the rail defect detection studies focus on the use of computer vision. In particular, machine learning methods were applied on images of the rails in order to obtain reliable and accurate detection of defects. For example, Ma el al. [13] worked on giving a general classification of rail defects, not only corrugation, into 8 severity levels. They used images extracted from a state-of-the-art system installed on a railway vehicle. It is considered to be the first fully automatic system that classifies overall condition of rail surfaces, not just a particular defect. Wei et al. [14] apply a multi-defect classification, that includes broken fasteners, missing fasteners, and rail corrugations etc., on selected images taken by a camera. They combine diverse spatial feature extraction methods and then apply an improved YOLOV3 model to accurately classify the different defects found in the images of the rail. Wei et al.[15] collected images of rails obtained from a handheld camera, and successfully identified only one form of rail defect - rail corrugations. Their method consisted of spatial feature extraction, represented by an improved Spatial Pyramid Matching model, followed by a linear Support Vector Machine classifier. They extend their work into also properly assessing the severity of the corrugations.

In comparison, our work uses signals collected from the ABA system installed on an in-service running railway vehicle, in order to recreate the moving RMS of the irregularity profile that can be measured by a diagnostic vehicle - from which a single rail defect, namely corrugation, can be easily detected. In contrast to the diagnostic vehicle's measurement outputs, the model should be able to recreate the irregularity profile much more frequently.

3. Materials and Methods

Let us now discuss the methods employed in this work, describing the axle-box derived acceleration data, the diagnostic vehicle measurements and the models employed to make predictions.

3.1. Axle-Box Acceleration System

One of the in-service vehicles of a subway line in Milan had the ABA system installed on one of its trailer bogies, for the purpose of acquiring the acceleration signals for the estimation of the rail roughness. The advantage of this system is that data can be collected at regular working speeds of the commercial vehicles, providing information regarding the rail profile every time the vehicle is driven on these rails. It should be noted that these commercial vehicles have autonomously driving controlled speed; this means the speed profile, which is an important feature in corrugation evolution, is smooth and often repeatable on all its runs. This system is comprised of: six accelerometers (four of which measure the vertical accelerations, and two measure the lateral accelerations), a gyroscope and an encoder. In Figure 2, we can see some of the mounted accelerometer sensors on the wheelset of an ATM commercial vehicle. One of the challenges is detecting the absolute position of the vehicle along the track of the subway line; it is computed by combining the speed derived from the encoder with data collected by the gyroscope. Taking into consideration that the distance covered by the vehicle over two consecutive stops is in the range of a single kilometer, the error of computation of the vehicle's position is below a meter, and thus acceptable.
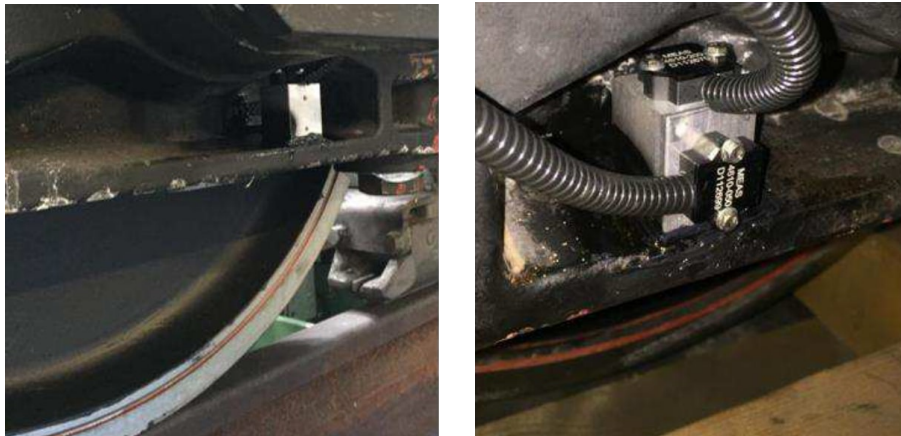

Figure 2.: MEMS accelerometers mounted on a commercial vehicle for vertical and lateral accelerations measurements.

3.2. Measurement Diagnostic Vehicle

ATM, which is both the railway operator and the infrastructure manager of Milan's local public transport, periodically measures vertical irregularities using a measurement

diagnostic vehicle and provided the data used in this work. The diagnostic vehicle is a trolley, moving at a speed of 3-4 km/h, and it collects direct measurements of the rail surface profile using lasers. The profile, acquired at 3 kHz, is filtered in different wavelength bands. Statistical values are then calculated for each band. As for the results that will be shown in this work (focused on the 100-300 mm band), the RMS is calculated on a 1.5 m moving window and it is re-sampled each 250 mm.

Because of the cost and time the operation of this vehicle takes, the vehicle is run approximately every two months during off-service hours. Moreover, with its limited speed and infrequent usage, it cannot be used as a continuous monitoring system.

3.3. Prediction Model Description

The problem addressed in this study is obtaining a prediction model which can guarantee a well enough accurate estimate of the irregularity profile. In this model, the data collected from the ABA system connected to a commercial vehicle is collected along with other features to recreate the direct measurements collected from the diagnostic vehicle. Accordingly, as depicted in Figure 3 we define this model as a black-box model with a vector input containing acceleration signals (along with with other data) and an output represented by the direct measurements of the rail surface.
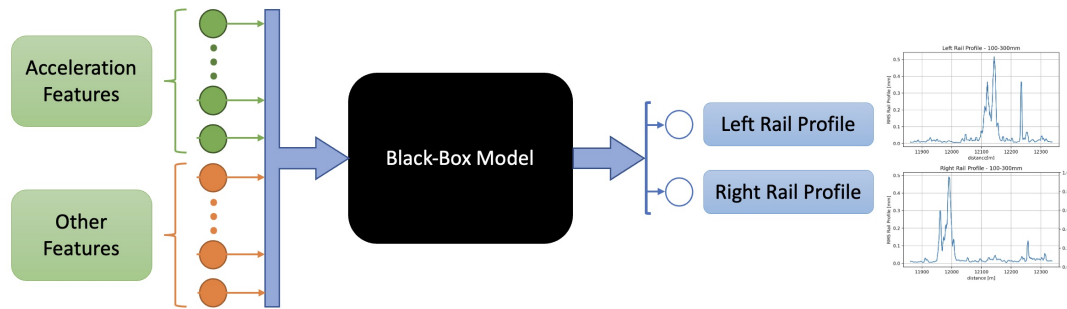

Figure 3.: The data-driven learning model under study.

This model can be characterised by a Supervised Learning (SL) model, which can be mathematically expressed as per the following [11]: Let X be the training set comprised of pairs independent variables (X) and response (y). Assuming there's a dependence relation between the response and its corresponding variables:

$$\text{find} \quad f : X \mapsto y = f(X), \quad \forall \ (X_i, y_i)_{i=1}^{l} \subset \texttt{X} \tag{1}$$

The objective is to fit a model that learns the mapping function f from input X to the output y, as per the following:

$$\hat{y}_i = g(X_i \mid \theta) \tag{2}$$

where: $g(.) = $ fit model
　　　　$\theta \quad = $ the model's parameters

During the learning process on the training set X, the parameters which minimize a loss function that calculates the error between the prediction $\hat{y}_i$ and the real output y_i are obtained by:

$$\arg\min_{\theta} \sum_i L(y_i, \hat{y}_i) = \arg\min_{\theta} \sum_i L(y_i, g(X_i \mid \theta)) \tag{3}$$

The model $g(.)$ and loss function $L(.)$ vary according to the type of learning model used and its corresponding method to find its optimized parameters θ.

3.4. Adopted Learning Models

In this subsection, we present the ML model that represents a good and adequate candidate to properly estimate rail irregularity profile: Random Forest Regression. A Random Forest Regressor (RFR) is a non-linear ML model that predicts a continuous target variable. It is an ensemble model of a collection of a base estimator, Decision Tree Regressor.

A Decision Tree is another ML algorithm that depends on binary partitioning features to develop a series of simpler rules/decisions after splitting more complex ones. Through repetitive process of splitting, which increases the depth level of the tree, a series of decisions are set to predict the target variable. However, a main concern for this algorithm is over-fitting of training data and weak generalization of results to testing data. [11]

In a RFR, an ensemble of trees is collected, with each tree trained with different series of decisions. The algorithm averages out the predicted outputs across all trees, which in return enhances the predictive performance in comparison to only one tree, while also avoiding over-fitting and benefiting from a much better generalization of results. [11, 16]

In our work, RFR is successfully used with an input vector of data (more details in the next section) highlighted by acceleration signals to predict the continuous rail profile irregularity. In training this model, two of its many hyperparameters are taken into consideration. First is the number of basic estimators used, i.e. number of diverse Decision Trees used in prediction. Second is the maximum depth set for each estimator/tree. In tuning these two hyperparameters, we optimize the RFR model used in minimizing the loss function (as stated in Equation 3), which in this case is the mean squared errors:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{4}$$

where: $\hat{y}_i$ = prediction of ith data point
y_i = observed target of ith data point
N = number of sample data points

4. Dataset

The study is based on one of the lines operated by ATM in Milan. It should be noted that the data provided is restricted to a number of consecutive sections of a bigger line, in which high levels of rail corrugations are usually witnessed. The datasets available on this track are the ground truth data provided by the measurement diagnostic vehicle used by ATM, the data signals produced by the ABA system installed on one of the in-service commercial vehicles on this track, and some information that are classified as offline and static features.

4.1. Target Data: RMS Measurements from Diagnostic Vehicle

The diagnostic vehicle produces as output the moving average RMS of the irregularity profile according to Standard EN 13231-3 [17], that takes into consideration the acceptance criteria for rail grinding. This trolley is driven over the track's rails, and it consequently measures the main profile for the wavelength bands: 30-100 mm and 100-300 mm. The data provided from this vehicle is a series of data for each of the left and right rails over the two wavelength bands. The declared measurement uncertainty is ±0.006 mm for the 30-100 mm band while it is 0.012 mm for the 100-300 mm band. The estimated roughness profile is re-sampled at a constant space resolution of 25 cm. Due to the high fluctuations in these measurements, the target data is a smoothed version of this data, namely the moving average with a window of 5 meters. In Figure 4, the ground truth moving RMS of the irregularity profile measurements are displayed for a section of the track. Signals are shown for the two mentioned wavelengths for the left rail and the right rail for the months of March, April and June of 2021. It should be noted few sections witness high levels of rail corrugation across the three months where measurements were collected. Also displayed is the curvature of the track in red dotted points (with positive values corresponding to a curve in the right direction) and the switching location in pink shaded area, which are both thought to greatly influence corrugation evolution on the rails. Figure 4 shows corrugation on narrow curves for sections with wavelength of 100-300 mm. The figure also shows that corrugations are not the same on parallel rail profiles; specifically, corrugation appears on the low rail of narrow curves. For example, in the left rail, high level of rail irregularity is located at distances 12050-12175 m where the second big curve (left direction) of the section is located. On the other hand, for the same wavelength band, corrugation has evolved on the right rail at the distances 11950-12050 m where there happened to be a big significant right curve. The data corresponding solely to 100-300 mm bandwidth are taken as the output dataset in the following sections, due to corrugation dominantly evolving only in this bandwidth in comparison to others.

4.2. Input Data

We now describe the two different types of input data used to predict the rail profile.
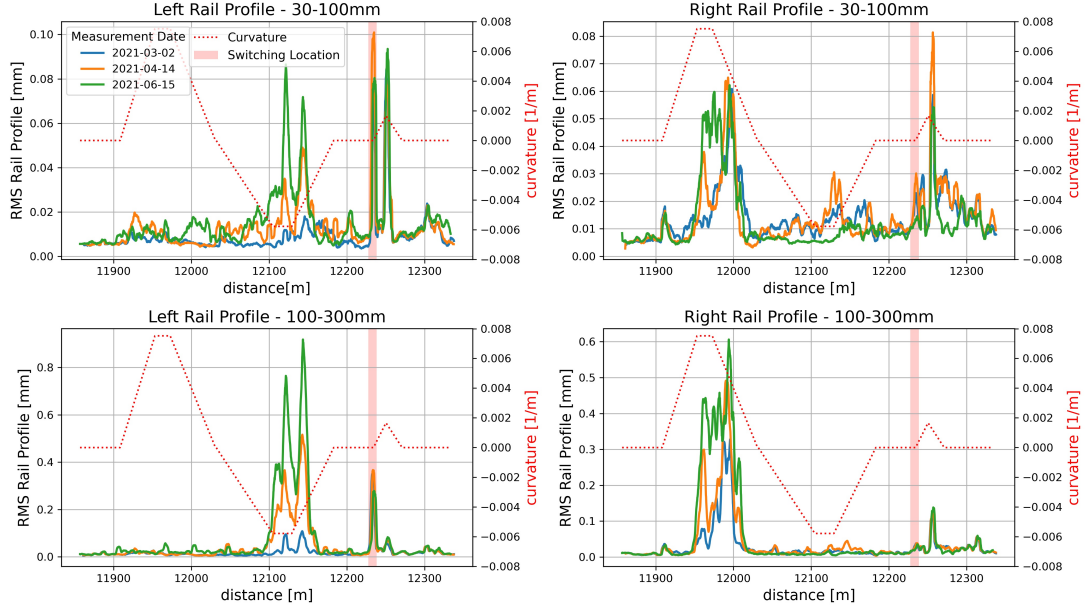
Figure 4.: Direct Measurements of the the moving RMS of the irregularity rail surface profile collected by the Diagnostic Vehicle.

4.2.1. Axle-Box Accelerometer Extracted signals

Measurements are collected from the ABA system installed on one of the everyday operating rail vehicles. This data is available almost everyday with the frequent use of the rail vehicle; however, as attested in Figure 4, measurements are collected by the diagnostic vehicle only rarely. Accordingly, with the purpose of correctly matching the input and the target, we only make use of the ABA's data that correspond to within a few days of the operation date of the diagnostic vehicle. This is because there is previous evidence that shows rail corrugation evolves rapidly while rail vehicles frequently pass over the same rails causing the prominence of the self-exciting phenomenon for fast corrugation growth. The data extracted from the ABA system consists of the following:

- Four corner vertical accelerations (one for each wheel of the axle set as denoted in Figure 5 in green).
- Two lateral accelerations (one on the front of wheel-set and another on the rear side as as denoted in Figure 5 in red).
- Speed and longitudinal acceleration of the vehicle (derived from the position encoder sensor in the system).

Signals extracted from ABA system are time-series measurements. Due to the purpose of predicting the moving RMS of the irregularity profile at specific location on the rail, these measurements are mapped into absolute position of the wheel-set with the multi-sensor box on the track. In Figure 6, rear left (RL) and rear right (RR) vertical acceleration signals are plotted against absolute distance on the left and right rails respectively in comparison to the measured irregularity profiles for the respective
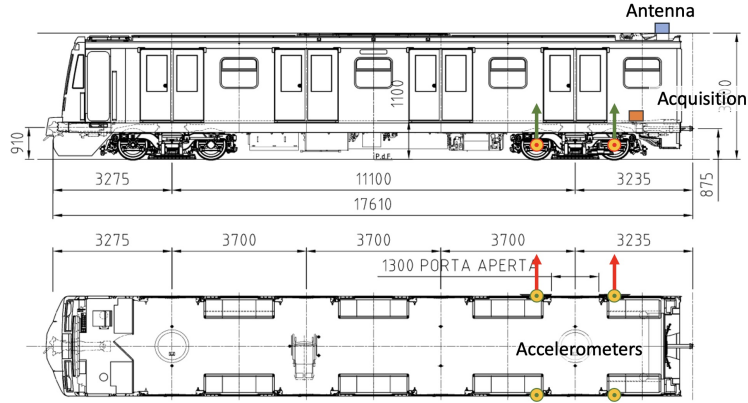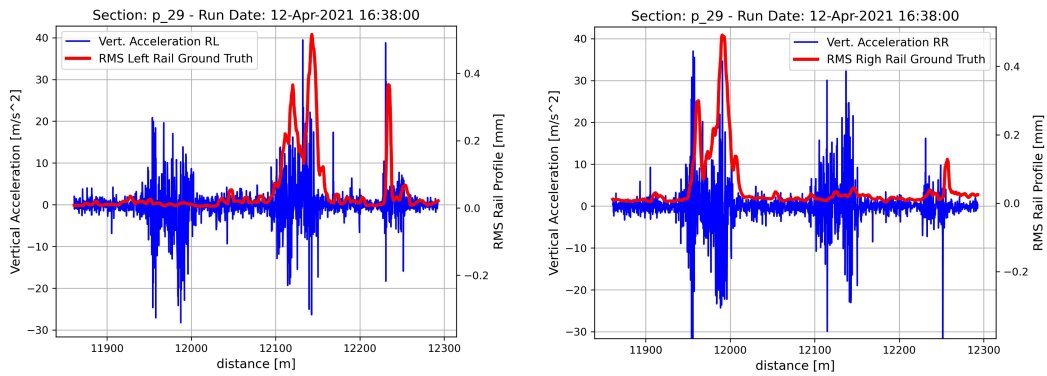
Figure 5.: Accelerometers location in the ABA system installed in a bogie of a rail vehicle.

rails (Figures 6a & 6b). The space resolution adapted is that of the diagnostic trolley's measurements in order to correctly match the input signals' positions to the output measurements' positions in the learning model.
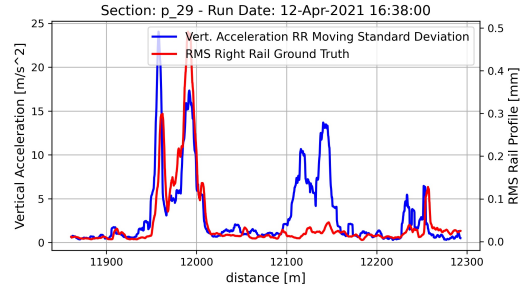

(a) Rear Left wheel vertical acceleration signal and Left Rail RMS irregularity profile

(b) Rear Right wheel vertical acceleration signal and Right Rail RMS irregularity profile

Figure 6.: Rear Wheelset Raw Vertical Acceleration signals plotted against distance in comparison to ground truth signals collected by diagnostic trolley at 100-300 mm bandwidth for left and right rails.

Before applying this mapping, these signals require thorough pre-processing. As evidenced in Figure 6, the vertical accelerations are noisy signals which garner higher energy and noise amplitude at specific locations on the track. Accordingly, a second order Butterworth band-pass filter is designed with a sampling frequency of 1000 and cutoff frequencies of 10 Hz and 120 Hz. This band is aimed at addressing corrugation wavelengths in the 100-300 mm range. In fact, this range is the one mainly influencing running comfort and the considered line is actually mostly affected by this kind of corrugation. The filter is applied on the vertical acceleration signals twice, once forward and once backwards, to maintain a zero phase and avoid lags/delays. Then, a moving standard deviation with a 500 ms window is applied on the filtered signal, which is then mapped to absolute position. Figures 7a & 7b show the rear wheel vertical acceleration
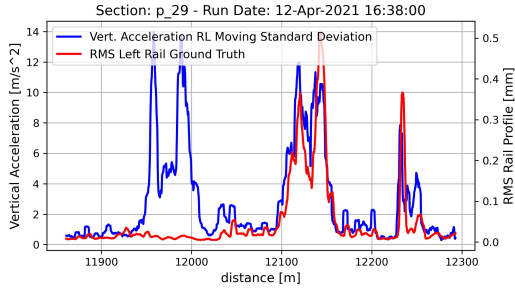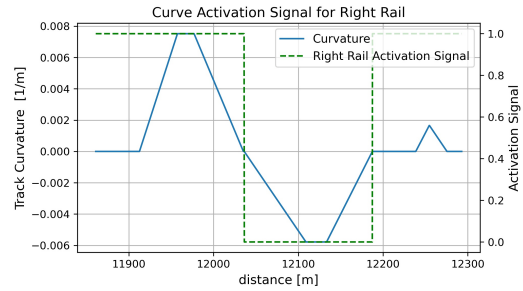
(a) Moving Standard Deviation of Filtered Rear Left Vertical Acceleration signal and Left Rail RMS Irregularity Profile
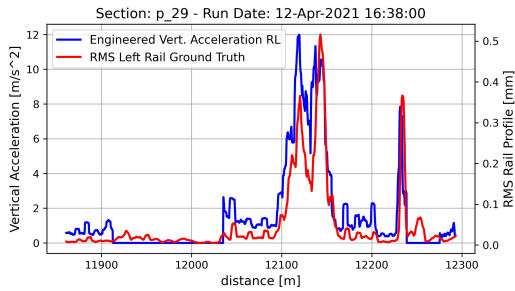
(b) Moving Standard Deviation of Filtered Rear Right Vertical Acceleration signal and Right Rail RMS Irregularity Profile
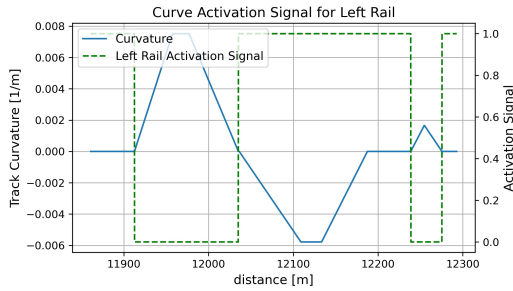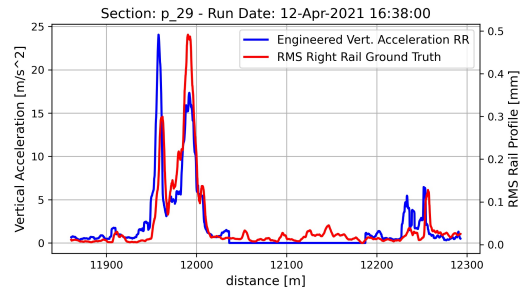
(c) Left Rail Activation Signal

(d) Right Rail Activation Signal

(e) Engineered Rear Left Vertical Acceleration and Left Rail RMS Irregularity Profile

(f) Engineered Rear Right Vertical Acceleration and Right Rail RMS Irregularity Profile

Figure 7.: Development of Engineered Rear Wheel Vertical Acceleration signals for both left and right rails.

signals after being processed and mapped. At specific locations on the track, the signals show a replicate dynamics to that demonstrated by the diagnostic trolley's measurements especially at locations with high levels of corrugations (i.e. at locations 12100-12175 m for the left rail and 11950-12025 m for the right rail).

In reference to Figures 7c & 7d which show the curvature of the section of the track under study, the dynamics of the signals shown in Figures 7a & 7b match at locations where the corresponding rail is the low rail at specific curves. This is an important observation as corrugation is prone to grow on inside rails at a curve. At other locations where the processed acceleration signal demonstrates high energy that is not matched by the diagnostic trolley's measurements, the rail under study happens to be the high/outside rail at a specific curve. The sensors at the side of the high rail capture the vibrations from the other rail due to coupling effects in the whole wheel-set.

For this reason, the importance arises for decoupling in order to separately predict the moving RMS of the irregularity profile on each rail and engineer acceleration features that depend on whether the rail under study at ta specific position is a high or low rail.

Accordingly, a curve activation signal for each rail is designed to rid of the coupling effect in the signals. The activation signal is a binary signal which is deactivated when the rail is the high rail (i.e. curvature > 0 for left rail and curvature < 0 for right rail) and activated otherwise. For each rail side, the curve activation signal is multiplied by the corresponding acceleration signals on the wheels in contact with these rails. The obtained signal are the engineered acceleration signals. The RL and RR engineering vertical acceleration signals can be seen in Figures 7e & 7f, with similar dynamics to the left and right measurements of the diagnostic trolley, respectively, at all locations of the section. These pre-processing steps followed by spatial mapping are applied to all vertical accelerations.

4.2.2. Static & Offline features

Along with those extracted from the ABA system, there are additional features available that might have certain effect on corrugation growth. These data are put into two categories: static and offline. The static signals are those that are a constant along a rail vehicle's run and only change from run to run. For each run, they are transformed into a non-dynamic vector with one constant value and a length equal to that of the output vector. The offline signals are those that do not change from run to run, but rather change along the position of the rail. They are inherently dynamic vectors that are mapped to the spatial position of the track and are a-priori information that are available regardless of the run. Moreover, in this application the features related to one category of the two are exclusively not in the other. This means that static features are not offline as they are not a-priori to the recorded run, and the offline features are not static as they change in value during the same run. The following are considered Static features:

- *Orientation of the commercial vehicle.* Depending on the run recorded, the commercial vehicle can be driven forwards or backward. This is an important information because dynamics recorded by the ABA system can differ as the orientation of the vehicle differs. This is related to the position of the bogey holding the multi-sensor box with respect to the rail. For example, a manoeuvre the rail vehicle can perform is slowing down on a curve, which will have its bogeys at different rail positions experiencing the same speed at a specific moment; now, if the vehicle's orientation is reversed while also performing the same manoeuvre at the same moment, each bogey will have also the same speed but at different locations over the rail. Since we are more interested about the position-based data over the rail, the bogey holding the system can have different sensor measurements based on its orientation. Moreover, absolute position corrections are applied on all signals given the orientation. This feature is represented by a binary (dimension=1) - forwards or backwards - data series that has the same value along the absolute position on the track and might differ only from run to run.

13

- *Time and date of the commercial vehicle's run.* Each set of measurements recorded by the ABA system for a certain run is associated by the time and date the recording occurred. This is because depending on the time of the day and and the day of the week, the commercial vehicle can vary in on-board crowdedness. This would vary the mass held in the cart which in return affect the vertical dynamics between the wheels and the rails. Each of the time of the day and the day of the week features for each run are cyclically encoded to preserve the cyclical significance of each; i.e. dimension=4. This is a series of data that holds the same value along all the absolute positions on the track but differ from run to run.

On the other hand, Offline features are data provided by ATM and are constituted of the following:

- *Curvature of the track* (reciprocal of the radius of curve). It is an offline data series along the absolute position of the track; i.e. dimension=1. A sample of this signal is shown as a blue curve over one of the sections of the track in Figures 7c & 7d.
- *Location of switches on the track.* It is also an offline data series which locates switches along the track; i.e. dimension=1. This considered as valuable information because vertical accelerations on the wheel-set can experience a sudden spike at these locations; these locations do not necessarily experience corrugation levels. The signal is represented by a binary data series that determines where an absolute position on the track is an area of switch or not.
- *Properties of the rails along the track.* Information regarding the track's rails differ from position to position. In particular, at specific locations there are diverse combinations of rail design specs, sleeper material, bed material (ballast or concrete) over which the rails and sleepers are installed, and presence of elastomers that isolate and mitigate vibration. This combination of properties regarding the track's structure and materials is one-hot encoded for use as binary features (dimension=4) along the absolute position of the track. These are prior-known features that do not differ from run to run.

4.3. Training, Validation and Test Split of data

With the abundance of the available runs recorded by the ABA system for a frequently used commercial vehicle, the quantity of usable data points is restricted by the availability of a matching target dataset, i.e. the measurements of the diagnostic trolley. As mentioned before, the input vector, which is constructed to match with each of the targets, is collected from runs close in terms of time (within a few days) to the date the diagnostic trolley was used. As attested in Figure 4, the trolley's measurements are extracted three times in total, once for each of the following months: March, April and June. Consequently, each target measurement of the available three is associated to several and different corresponding input vector runs; the number of runs also differs in availability from section to section.

Given this into consideration, the dataset is split into a training set to fit the learning

model and determine its parameters and a testing set that evaluates the built model. The training set corresponds to all the runs from March and April along with half of June's runs. The other half of June's runs are kept as a testing dataset, which represent around 17.5% of the whole available and usable dataset. In summary, the training sample size is 673575, while the testing set size is 143123.

An important factor is building the input vector of the RFR model. As seen in the previous two subsection, diverse information are available. However, not all of them might be necessary or important for building an adequate ML model capable of predicting the moving RMS of the irregularity profile of the rails. Several models with different combinations of features in the input vector are trained; the input vector of each of the models is:

- *Raw ABA signals model*: signals generated by the ABA system, namely, raw vertical accelerations (4), lateral accelerations (2), longitudinal acceleration (1) and longitudinal speed (1).
- *Moving Standard Deviation ABA signals model*: signals generated by the ABA system, with the raw vertical accelerations being replaced by their subsequent filtered moving standard deviation versions as shown in Figures 7a & 7b.
- *Engineered ABA signals model*: signals generated by the ABA system, with the raw vertical accelerations being replaced by their subsequent engineered versions as shown in Figures 7e & 7f.
- *Engineered ABA signals with Static Features model*: engineered ABA signals in addition to the Static features mentioned in Section 4.2.2.
- *Engineered ABA signals with Offline Features model*: engineered ABA signals in addition to the Offline features mentioned in Section 4.2.2.
- *Engineered ABA signals with Static and Offline Features model*: engineered ABA signals in addition to both Static and Offline features mentioned in Section 4.2.2.

The above models were formed and studied for two main purposes. The first one is to better understand the benefit of engineering the acceleration signals. The second one is to evaluate whether adding additional features is advantageous. The input and output dimensions are summarized in Table 1.

Table 1.: Trained Random Forest Regression Models input & output dimensions.

ML Model's Input Vector	ABA Signals	Static Signals	Offline Signals	Input Dimension	Output Dimension
Raw ABA	Yes	No	No	8	2
Moving Standard Deviation ABA	Yes	No	No	8	2
Engineered ABA	Yes	No	No	8	2
Engineered ABA + Static	Yes	Yes	No	13	2
Engineered ABA + Offline	Yes	No	Yes	14	2
Engineered ABA + Offline + Static	Yes	Yes	Yes	19	2

No dataset is kept aside for validation, instead a cross-validation technique is followed. Each of the ML models mentioned above undergo a thorough grid search to find their optimal pair of hyperparameters: total number of base estimators ([50,75,100,125,150]) and maximum depth of each estimator ([10,12,14,16,18,20]). Each of the models is tuned via 5-fold cross-validation on the training dataset, picking the pair of hyperparameters that minimise the mean squared error, as per Equation 4.

5. Results

We now discuss the evaluation metrics employed and results obtained for various predictive models.

5.1. Evaluation Metrics

To illustrate and assess the performance of the proposed approach, the metrics detailed in the following are adopted:

- RMSE: root mean squared error. It represents the standard deviation of the predicted errors. Thus, the lower the better. It is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{5}$$

- MAPE: mean absolute percentage error. It is another measurement of prediction accuracy for regression problems, in which the average normalized difference between the actual value and predicted value is obtained across the whole sample. Thus, the lower the better. It is defined as follows:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \frac{|y_i - \hat{y}_i|}{y_i} \tag{6}$$

- R^2 score: also called the coefficient of determination, represents the percentage of learned variability using the model. It measures how good of a fit the model is. Thus, the higher the better; it ranges between [0, 1]. Small or negative values indicate the presence of a bias in the prediction resulting in poor accuracy but it does not indicate anything about signal shape identification. It is defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \tag{7}$$

  where: $\bar{y}$ = mean of observed target data

- Cosine Similarity: It is a measure of similarity between two sequences or vectors of similar size; amplitude difference is disregarded, and only dynamic changes of signals are captured. It is the un-centered version of Pearson Correlation. In contrast to Pearson Correlation, Cosine Similarity is invariant to shifts, which is important as we are interested in detecting similar dynamics at same positions on the rail between the observed and predicted target. Due to the signals exclusively being positive, the performance measure is bounded between [0, 1]. High values indicate that the model captures the shape of the true signal. It is defined as

16

follows [18]:

$$\text{CosineSim} = \frac{\sum_{i=1}^{N} y_i \hat{y}_i}{\sqrt{\sum_{i=1}^{N} y_i^2}\sqrt{\sum_{i=1}^{N} \hat{y}_i^2}} \tag{8}$$

5.2. Different models Performance Comparison

All RFR models with different input vectors described in Section 4.3 are trained as described and their optimal hyperparameters are obtained as shown in Table 2.

Table 2.: Trained Random Forest Regression Models with their optimal hyperparameters for each of the left and right rail.

RFR Model's Input Vector	Left Rail		Right Rail	
	Base Estimators Number	Maximum Depth	Base Estimators Number	Maximum Depth
Raw ABA	100	10	100	10
Moving Standard Deviation ABA	125	16	50	10
Engineered ABA	150	12	100	10
Engineered ABA + Static	100	10	150	10
Engineered ABA + Offline	50	10	125	20
Engineered ABA + Offline + Static	150	10	50	16

Each of the models' predictions is evaluated on the testing set. The result on the right rail irregularity predictions are shown in Table 3. It can be seen that building a model with only raw signals from the ABA produces predictions with a low R^2 score of 0.3149, which is an unsatisfactory result. Improvement is seen when taking the moving standard deviation of the acceleration signals as both R^2 score and CosineSim increase while RMSE and MAPE decrease. Further improvement is recorded with the application of the engineering method (as mentioned in Section 4.2.1) of the acceleration features as R^2 score, RMSE and MAPE are improved by 72.72%, 18.29% and 29.55% relative to the model with Raw ABA signals. The 17.22% improvement in CosineSim shows that not only errors between prediction and target is decreased, but also the dynamics of the predicted profile is closer to imitating that of the target. This validates the use of engineered features.

In the last three models, they all include the engineered ABA signals as inputs. Adding the static features has proved to be redundant if not disadvantageous as both R^2 score, RMSE and MAPE got a little worse with respect to the model that includes only the engineered ABA features; the similarity metric only marginally improved. On the other hand, the addition of the offline features has shown greater improvements as all error-based and similarity-based metrics were the best relative to all models. The improvements of R^2 score, RMSE, MAPE and CosineSim with respect to the model with Raw ABA signals are 216%, 44.84%, 28.52% and 39.71%, respectively. In the last model, we consider adding all available features in the input vector, but there is no improvement; as a matter of fact, adding the Static features to the best model yet, which includes in its input vector the engineered ABA signals and the Offline features, is not beneficial as all metrics slightly regressed.

17

Table 3.: Right Rail Trained Random Forest Regression Models performance evaluation on whole line of Testing set.

RFR Model's Input Vector	Performance Metrics			
	RMSE	MAPE	R^2 score	CosineSim
Raw ABA	0.0339	0.8464	0.3149	0.7071
Moving Standard Deviation ABA	0.0321	0.6554	0.3869	0.7865
Engineered ABA	0.0276	0.5963	0.5439	0.8289
Engineered ABA + Static	0.0277	0.6244	0.5426	0.8300
Engineered ABA + Offline	**0.0157**	**0.3608**	**0.8521**	**0.9452**
Engineered ABA + Offline + Static	0.0167	0.3995	0.8324	0.9378

These results show that taking into consideration Static features, i.e. orientation of the vehicle or the time of the run, is unnecessary as it doesn't add anything to the model. These features do not change during the same run and thus cannot help in detecting areas of corrugations. Also, they are actually corrections that are already indirectly included in the acceleration signals. On the other hand, the Offline features have proved that they are valuable information for the learning model to acquire to produce more accurate predictions as opposed to only gain knowledge on the irregularity profile of the rails from just the signals extracted from the ABA system. The model with engineered ABA signals and Offline signals as inputs is adopted from here on.

Next, we compare RFR with: Decision Tree Regression [19], Support Vector Regression (SVR) [20], and Deep Neural Network (DNN) [21, 22]. All benchmark ML models are constructed with same inputs, i.e. Engineered ABA signals and Offline signals. The test results are summarized in Table 4 and it shows that the RFR outperforms other models. It is worth noting that the decision trees and DNN did a reasonably good job in capturing the dynamics with a CosineSim higher than 0.85; however, both models failed to robustly predict the tails of the corrugation distributions (peaks). This is because the first generalizes weaker than the RFR, while the DNN, although powerful, are restricted by their batching mechanism during learning leading to underestimating of the peaks. SVR has noisy (low CosineSim values) and bad (low R^2 score and high RMSE) predictions due to the fact it does not perform well in large datasets. More details in Appendix A.

Table 4.: Comparison of ML Models performance performance evaluation on right rail of whole line in the Testing set.

ML Models	Performance Metrics			
	RMSE	MAPE	R^2 score	CosineSim
Random Forest Regression	**0.0157**	**0.3608**	**0.8521**	**0.9452**
Decision Tree Regression	0.0207	0.5808	0.7443	0.9019
Support Vector Regression	0.0913	6.0278	0[*]	0.3587
Deep Neural Network	0.0242	0.6282	0.6497	0.8664

[*] huge biases; mean of observed target is better than prediction

5.3. Comprehensive Results

As mentioned in the previous section, the runs corresponding to the testing dataset are recorded on dates within a small period of time of the ground-truth measurement taken in July. Some runs are recorded just before and others are recorded just after the target data was extracted from the diagnostic trolley. Accordingly, for fair comparison with the target, the average of the of all runs' predictions around the measurement date of interest are taken against a single target data measurement. The process is highlighted in Figure 8.
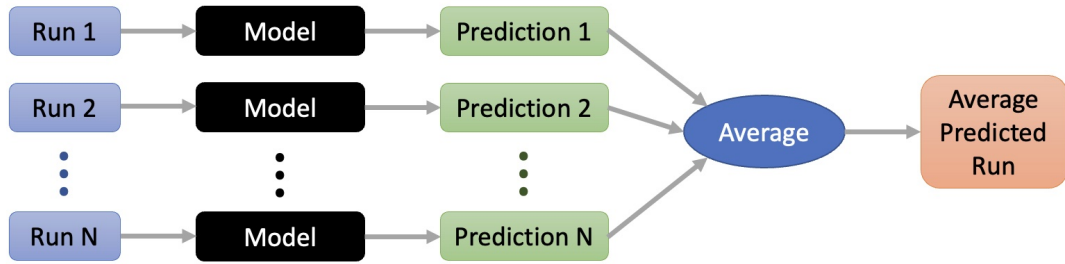

Figure 8.: The average prediction process of the runs corresponding to a diagnostic trolley's measurement date.
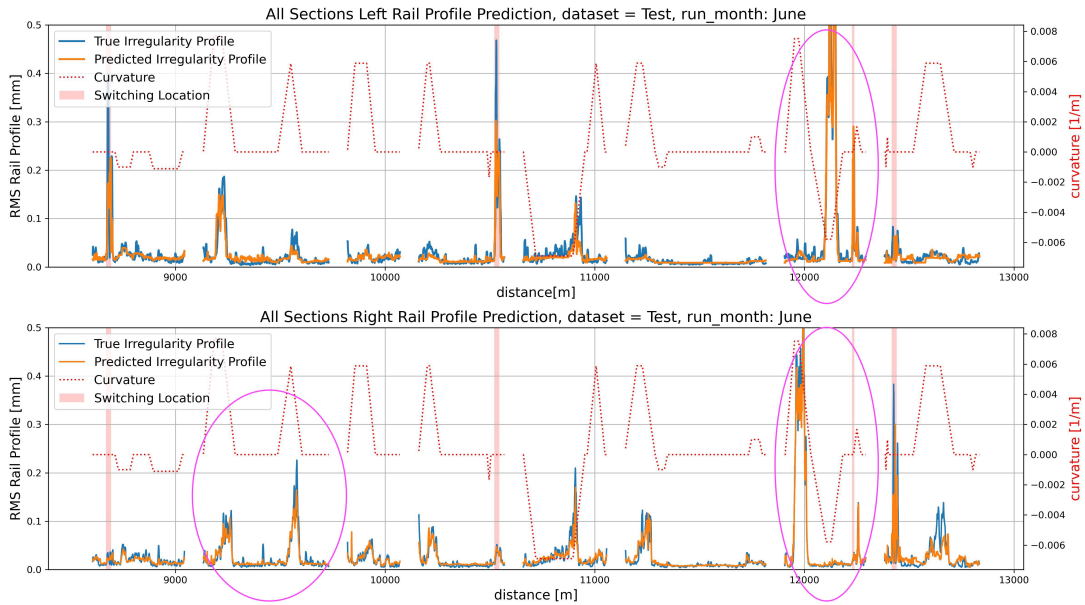

Figure 9.: Irregularity profile prediction using Random Forest Regression model vs. diagnostic trolley's measurements for all available consecutive sections of the line for both rails at a 100-300 mm bandwidth.

The average predictions of the runs corresponding to the testing dataset are plotted along consecutive sections of the track corresponding to a distance of around 4 km, as shown in Figure 9. The graph on top shows the rail profile for the left rail, while the bottom one shows that of the parallel right rail. In both graphs, the curvature of the consecutive sections is drawn for reference. The blank spaces between the 8 total

sections represents locations of the stations that the rail vehicle stops at along the track, and they are not of interest in rail corrugation detection. The predictions show, along all the sections, a strong ability to trace the irregularity profile measurements collected by the diagnostic vehicle. This is also represented by its performance metrics displayed in Table 5. Most notable result is the high R^2 scores: 0.9076 for the left rail and 0.8758 for the right rail. Moreover, the low RMSE and MAPE metrics along with the almost perfect CosineSim metric prove the credibility of the model in re-producing the measurements in the both the Left and Right rail.

Table 5.: Right and Left Rails Models' Average Predictions' Test Performance on all available consecutive sections of the line.

	Performance Metrics			
ML Model's Corresponding Rail	RMSE	MAPE	R^2 score	CosineSim
Left Rail	0.0199	0.3957	0.9076	0.9632
Right Rail	0.0185	0.3565	0.8758	0.9532

Three specific sections in Figure 9 are circled. They represent the top three sections with most corrugations in the whole testing set; that does not mean the highest amplitude reached, but rather the largest areas under the curve. They are displayed each (as seen in Figure 9 in the order of top-to-bottom and left-to-right) in Figure 10 in a clearer fashion. Figures 10a and 10c represent the Left and Right rails respectively of the same section, which happens to be the most corrugated section of the whole track. The trained model is able to predict, reach and dynamically imitate "interesting outliers", which in our case are highly corrugated areas. Specifically, the left rail corrugation on the big left curve (at locations 12100-12175 m) in Figure 10a and the right rail corrugation on the first right curve (at locations 11950-12025 m) in Figure 10c are accurately predicted, not only in magnitude but also in their local peaks. Moreover, the small spike in the left rail at the switching location in 10a at 12230 m is also accurately predicted; it can be observed that the switching location only affected the left rail in this case, and this is because it happens to be switching on the left rail. On the other hand, Figure 10b displays the Right rail of a different section. At each of the two existing right curves, there can be seen developing corrugations with lower amplitudes than ones witnessed in the other two figures, and the trained model can also accurately trace them. The strong prediction performances of these sections is also evidenced in Table 6, particularly with high R^2 scores of 0.97 and 0.98 for the left and right rail respectively for the most corrugated section, especially in comparison with those of the whole Left and Right rails seen in Table 5. Moreover, the lower MAPE and higher CosineSim metrics also show that the model performs better in the sections of higher corrugations, which are sections of higher interest for condition based maintenance purposes.

(a) Section *p_29*, Left Rail

(b) Section *p_24*, Right Rail
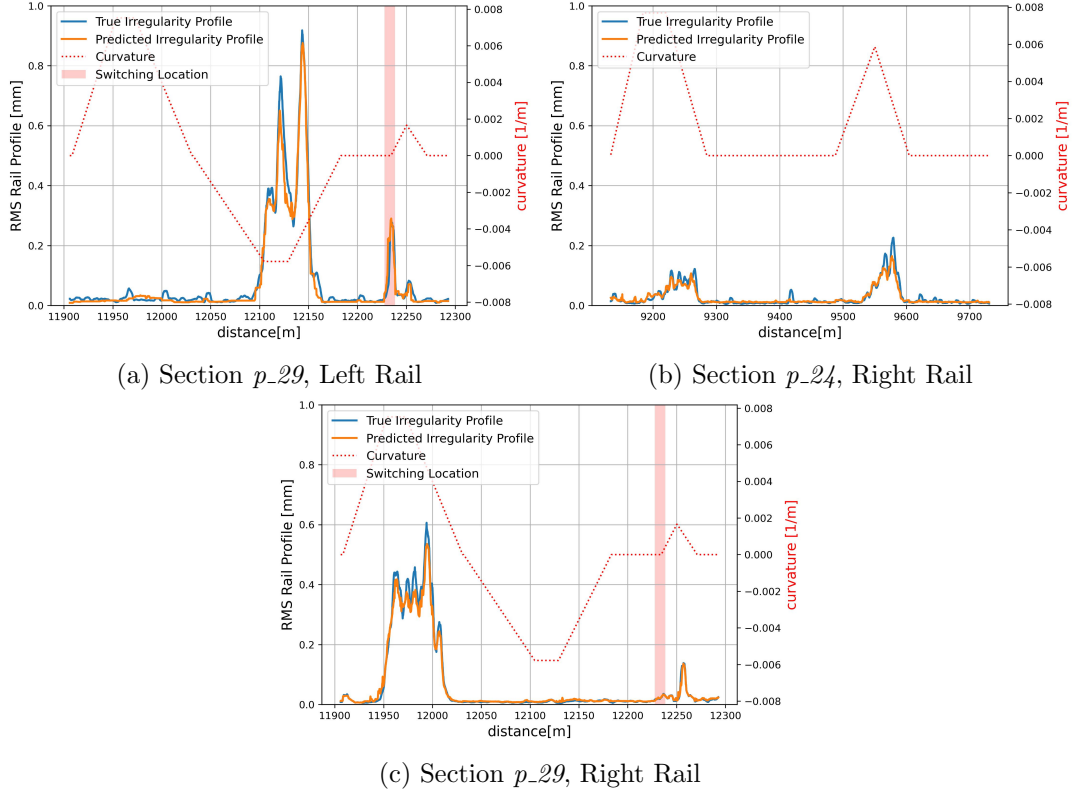
(c) Section *p_29*, Right Rail

Figure 10.: Moving RMS Irregularity profile prediction using Random Forest Regression model vs. direct measurements for three section for both rails at a 100-300 mm bandwidth.

6. Discussion

6.1. Comparison with Benchmark

An important validation of the technique followed in this paper is to compare it with another benchmark that also uses inputs from the ABA system. In this case, the prediction obtained by the Random Forest Regression model is compared with that of the technique used in the work of Karaki et al. [4] (in which the deformability of the system is taken into account thanks to analytical models optimized through experimental tests). In particular, in their findings, they obtain an estimation of the rail profile using a model-based transfer function that only inputs the vertical acceleration signal of the left front wheel in the case of the left rail and the right front wheel in the case of the right rail. However, since the model presented in this paper directly predicts the 5 m moving average of the true rail profile (provided by the diagnostic trolley), the output of their model has a 5 m moving average also applied to it for fair comparison.

Both models of interest are applied on a testing run in June, as seen in Figure 11, where a particular segment of a section of the track with apparent corrugation is shown for both the left and right rails. The segment represents a left curve, where corrugation

21

Table 6.: Model Average Predictions Test Performance on sections shown in Figure 10.

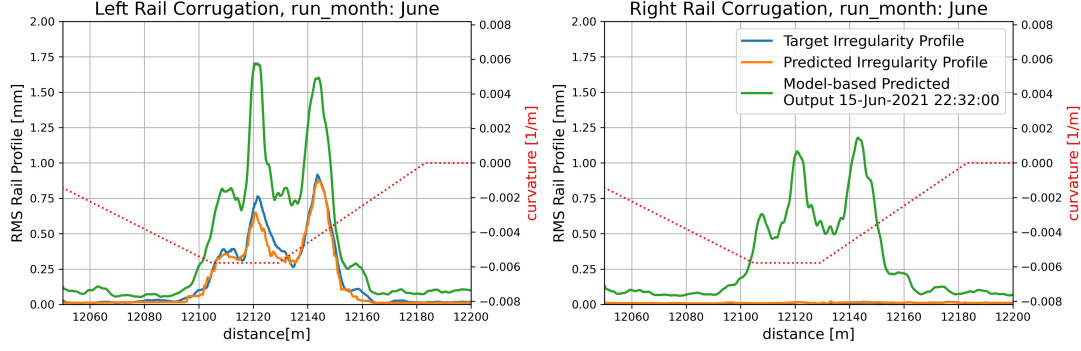| Rail of Figures	Performance Metrics			
	RMSE	MAPE	R^2 score	CosineSim
Left Rail Section p_29, Figure 10a	0.0283	0.3101	0.9708	0.9910
Right Rail Section p_24, Figure 10b	0.0124	0.3198	0.8812	0.9694
Right Rail Section p_29, Figure 10c	0.0164	0.2216	0.9838	0.9968

Figure 11.: Comparison between the moving RMS irregularity profile predictions by the Random Forest and by the model-based transfer function model of the left curve of a section.

has only evolved on the left rail. However, the model-based function has mistakenly estimated a high corrugation level on the right rail, while there was none as represented by the target RMS irregularity profile. The false irregularity estimation on the right rail in fact resembles the correct profile in the left rail at the same locations. This is due to the parallel rails coupling effect that the model fails to eliminate. Conversely, the data-driven model developed includes acceleration engineered features that remove the coupling effect. Moreover, in predicting the high corrugation levels, the Random Forest Regression model correctly reaches the correct magnitudes of the true moving RMS irregularity profile. The model-based function's output dynamically resembles the true target measurement in case of the left rail, but it overestimates in magnitude along the whole segment.

Table 7.: Data-driven vs. Model-based function [4] Prediction Performance of Segments shown in Figure 11

Rail of Segment in Figure11	Performance Metrics			
	RMSE	MAPE	R^2 score	CosineSim
Left Rail Data-driven model	**0.0396**	**0.2242**	**0.9716**	**0.9938**
Left Rail Model-based model	0.3117	2.7645	0^{*}	0.9885
Right Rail Data-driven model	**0.0029**	**0.2674**	0^{*}	**0.9779**
Right Rail Model-based model	0.4248	31.847	0^{*}	0.7056

[*] huge biases; mean of observed target is better than prediction

The results are further highlighted in Table 7. The Data-driven model shows a

perfect R^2 score on the left rail, representing ability of the model in capturing the high corrugation values. On the right rail, the R^2 score is null and redundant as the RMSE values are almost zero, while also showing an even better RMSE value than the one in the left rail due to the very low values of the signals; there's no corrugation occurring on this rail in this segment of the section. On the other hand, the model-based function shows a high CosinSim but low R^2 and high RMSE, which indicates that the model captured the shape of corrugation but failed to accurately predict it. This is also clear from Figure 11, where it shows much higher values than the true corrugation values but still captures the shape for the left rail. The null R^2 scores indicate the model predictions are biased resulting in poor prediction accuracy but not necessarily failing to capture the dynamics. As for the right rail, the model predicts a corrugation that does not exist, which is why even the CosineSim metric is not as good.

6.2. Importance Features Case Study: Lateral Accelerations

Most institutions that attempt to install ABA systems on their vehicles do not have accelerometers to test lateral acceleration. This raises the question of the importance of these measurements towards predicting a statistic of the rail irregularity profile that would help identify rail corrugation. To this end, we attempt to build data-driven models without the lateral acceleration features in an attempt to provide explainability and interpretability of the features' effect in the problem at hand. In Table 8, we compare the average predictions of the model without the lateral acceleration features to the full model for both rails. While accounting for lateral acceleration features in the left rail gave marginally better results, the right rail results of model without those features was marginally better. Overall, the addition of lateral acceleration features did not significantly augment (or decrease) the prediction performance, and thus these features are either redundant or provide information that is already accounted for by other features.

Table 8.: Comparison between Models' Average Predictions Test Performance on all available consecutive sections of the line for Right and Left Rails, with and without Lateral Acceleration Features.

ML Model's Corresponding Rail	Performance Metrics			
	RMSE	MAPE	R^2 score	CosineSim
Left Rail Model w/ Lateral Accelerations	**0.0199**	**0.3957**	**0.9076**	**0.9632**
Left Rail Model w/o Lateral Accelerations	0.0205	0.4004	0.9019	0.9621
Right Rail Model w/ Lateral Accelerations	0.0185	**0.3565**	0.8758	0.9532
Right Rail Model w/o Lateral Accelerations	**0.0183**	0.3646	**0.8777**	**0.9539**

7. Conclusion

In this paper, a ML model, namely a Random Forest Regressor, is constructed for recreation of moving RMS irregularity rail profile, based on ABA system measurements

installed on an in-service rail vehicle. The purpose of this model is to predict rail corrugation level measurements that can replace those produced by the measurement diagnostic trolley. Left and right rails of a Milan rail track are separately predicted, and thus engineered features were extracted from the acceleration signals to remove coupling effects. A comparative study was done by designing several models with diverse input features to validate the use of engineered features. Another purpose of this study was to show that adding the so-called Offline features will build a better model, while Static features are redundant.

The estimated irregularity profiles of the rails were compared with the direct target measurements. Encouraging results have shown an R^2 score of 0.9076 and 0.8758 for the left and right rail respectively along all the sections under study, with highly corrugated section reaching an R^2 score of 0.97 and 0.98 respectively. Comparisons to a successful frequency-based model, which also uses signals from an ABA system, show that the ML model can be more accurate in recreating the measurements of the diagnostic trolley on both left and right rail, while also clearly identifying corrugated areas in need of treatment.

It is worthy to note that the framework proposed was not built to only accustom one case study, but to also be generalized for different railways that might encounter different kinds of commercial vehicles. However, the optimal model built in the case study is not generalizable to other railway sections; the generalizable framework entails extracting ABA signals, proper pre-processing, building model and training before obtaining credible future predictions of RMS irregularity profile. For future work, we would like to stretch the envelope of robustness of the framework proposed by investigating on different railways with different commercial vehicles and interpret how feature importance for prediction might change across these environments.

8. Acknowledgements

9. Declaration of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Grassie SL, Kalousek J. Rail Corrugation: Characteristics, Causes and Treatments. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. 1993;207(1):57-68.

[2] Grassie SL. Rail Corrugation: Characteristics, Causes and Treatments. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. 2009;223(6):581-96.

[3] Grassie SL. Measurement of railhead longitudinal profiles: a comparison of different techniques. Wear. 1996;191(1-2):245-51.

[4] Karaki J, Faccini L, Di Gialleonardo E, Somaschini C, Bocciolone M, Collina A. Continuous Monitoring of Rail Corrugation Growth Using an In-Service Vehicle. In: the 27th IAVSD Symposium on Dynamics of Vehicles on Roads and Tracks. vol. 243; 2021. .

[5] Xie J, Huang J, Zeng C, Jiang SH, Podlich N. Systematic Literature Review on Data-Driven Models for Predictive Maintenance of Railway Track: Implications in Geotechnical Engineering. Geosciences. 2020;10(11):425.

[6] Khajehei H, Ahmadi A, Soleimanmeigouni I, Haddadzade M, Nissen A, Jebelli MJL. Prediction of track geometry degradation using artificial neural network: a case study. International Journal of Rail Transportation. 2022;10(1):24-43.

[7] Zhan Y, Dai X, Yang E, Wang KCP. Convolutional neural network for detecting railway fastener defects using a developed 3D laser system. International Journal of Rail Transportation. 2022;9(5):424-44.

[8] Lewis RB, Richards AN. A new method for the routine measurement of railhead corrugations. Rail International. 1986:37-41.

[9] Grassie SL. Measurement of Longitudinal Irregularities on Rails Using an Axlebox Accelerometer System. In: Noise and Vibration Mitigation for Rail Transportation Systems. Springer Nature; 2021. p. 320-8.

[10] Attoh-Okine NO. Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering. John Wiley & Sons, Inc.; 2017.

[11] Martey EN, Ahmed L, Attoh-Okine N. Track geometry big data analysis: A machine learning approach. In: 2017 IEEE International Conference on Big Data (Big Data); 2017. p. 3800-9.

[12] Ma S, Gao L, Liu X, Lin J. Deep Learning for Track Quality Evaluation of High-Speed Railway Based on Vehicle-Body Vibration Prediction. IEEE Access. 2019;7:185099-107.

[13] Ma K, Yago Vicente TF, Samaras D, Petrucci M, Magnus DL. Texture classification for rail surface condition evaluation. In: IEEE Winter Conference on Applications of Computer Vision (WACV); 2016. p. 1-9.

[14] Wei X, Wei D, Suo D, Jia L, Li Y. Multi-Target Defect Identification for Railway Track Line Based on Image Processing and Improved YOLOv3 Model. IEEE Access. 2020;8:61973-88.

[15] Wei D, Wei X, Liu Y, Jia L, Zhang W. The Identification and Assessment of Rail Corrugation Based on Computer Vision. Applied Sciences. 2019;9(18):3913.

[16] Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32.

[17] BSI: Railway applications - Track - Acceptance of works - Part 3: acceptance of reprofiling rails in track. The British Standards Institution.; 2012.

[18] Han J, Kamber M, Pei J. Getting to Know Your Data. In: Data Mining. Elsevier; 2012. p. 39-82.

[19] Torgo L. Regression Trees. In: Encyclopedia of Machine Learning. Springer US; 2011. p. 842-5.

[20] Awad M, Khanna R. Support Vector Regression. In: Efficient Learning Machines. Apress; 2015. p. 67-80.

[21] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.

[22] Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.

Appendix A.

Figure A1 shows the predictions of each of the optimized ML models for the left and right rails for a segment of a section at a narrow corner. It can easily be observed that the Random Forest Regression prediction has the least deviation to the true target and outperforms the other three ML models.
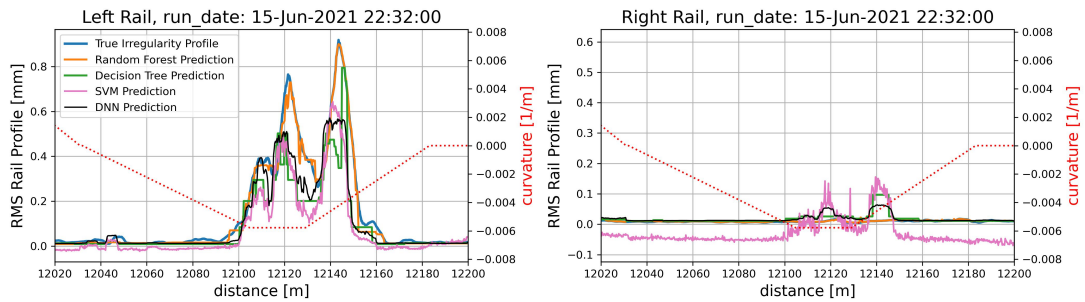

Figure A1.: Comparison between ML models' prediction of moving RMS irregularity profile for a single run shown on a segment of a narrow corner in Section *p_29* for Left and Right Rail.