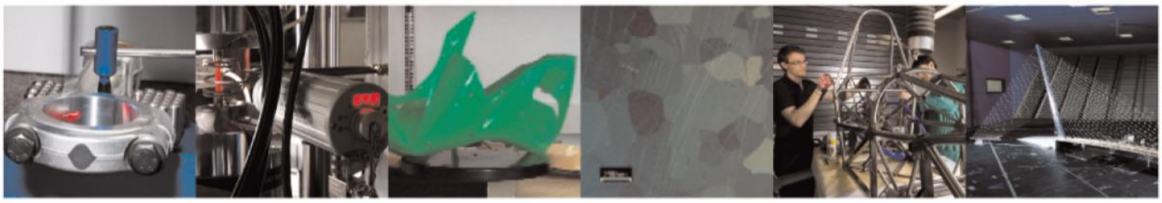

Acceleration-based condition monitoring of track longitudinal level using multiple regression models

I. La Paglia, E. Di Gialleonardo, A. Facchinetti, M. Carnevale & R. Corradi

This is a post-peer-review, pre-copyedit version of an article published in Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. The final authenticated version is available online at: https://doi.org/10.1177/09544097231201513

Acceleration-based condition monitoring of track longitudinal level using multiple regression models

I. La Paglia[1], E. Di Gialleonardo[1], A. Facchinetti[1], M. Carnevale[2], R. Corradi[1]

[1] Politecnico di Milano, Department of Mechanical Engineering,
Via La Masa 1, 20156 Milano, Italy

[2] Università degli Studi di Pavia, Dipartimento di Ingegneria Industriale e dell'Informazione,
Via A. Ferrata 5, 27100 Pavia, Italy

Corresponding author: ivano.lapaglia@polimi.it

Abstract

In this paper, a condition monitoring system for railway track geometry is presented. The methodology has been designed for high-speed application, where the train travels at the maximum allowed speed for most of the trip. The system is designed to rely on acceleration data recorded by in-service vehicles to provide estimations of the track longitudinal level, based on pre-built regression models. It exploits synthetic indicators sampled over predefined track sections 100 m long. Different predictors are considered, computed both from acceleration data and from track geometry measured by the diagnostic train. The proposed modelling strategy allows distinguishing between isolated and distributed defects that populate the railway track as well as reproducing the evolution over time of the maximum longitudinal level registered in the considered track section; moreover, also accurate predictions of the defect amplitude are made. The results have been validated against track geometry data recorded by the diagnostic train during a monitoring period of two years. It is proven that the proposed system could support current maintenance strategies, providing a continuous flow of data to monitor the track infrastructure.

Keywords: condition monitoring, longitudinal level, bogie vertical acceleration, in-service vehicle, multiple regression models.

1. Introduction

Maintenance strategies currently adopted by railway infrastructure managers to guarantee the integrity and safety of the network rely on special purpose diagnostic trains. These trains are exploited for periodic inspection runs, planned according to a predefined schedule, during which both acceleration data and track geometry parameters are recorded [1]. The frequency of train passages should be sufficient to detect any occurring change of the track conditions, and it normally depends on the class of the railway line. For instance, high-speed lines are typically inspected every two weeks, whereas less frequent inspections (e.g., monthly) are carried out on commuter lines. However, diagnostic trains do not allow a continuous monitoring of the track state, that may lead to downtimes and potentially critical situations in the unlikely event of very fast or sudden degradation of the track conditions.

In the last decade several works have been proposed, proving in-service vehicles to be suitable for infrastructure monitoring purposes and progressively leading to condition-based maintenance strategies (CBM). Inertial sensors have been widely adopted in this field of research, as they are robust, reliable, versatile and rather cheap. They can be installed on different components of the vehicle [2], their position, type and full scale being dependent on the target of the specific application, i.e., the defect type. For instance, axle-box mounted accelerometers were employed to successfully monitor rail corrugation and roughness [3][4]. Bogie mounted sensors were adopted to monitor vertical and lateral track alignment [5][6][7]. Finally, applications relying on carbody mounted sensors were proposed, in particular for low speed and urban applications [8][9].

Once data are recorded on-board and transmitted to a ground server for analysis, different methodologies could be employed. The acceleration time histories can be used to solve an inverse problem, willing to identify the input track profile generating the acquired signal. Carbody acceleration was used in [10] to solve an inverse problem adopting a Kalman Filter (KF) with a simplified yet fast rail vehicle model. A mixed filtering approach allowed estimating track irregularities [11], combining KF for the displacement estimation, band-pass filters for wavelength classification and a compensation filter for amplitude and phase estimation.

Vehicle accelerations can also be treated at the post-processing stage to identify the track geometry profile by adopting signal processing techniques. Double integration of the acceleration signals was proven to be an effective solution. Axle-box vertical acceleration allows reconstructing the track geometry of a high-speed line [12], adopting a 10 m versine processing. The results were validated in [13] against track geometry data measured by a Track Recording Vehicle (TRV), showing satisfactory results. Other attempts have been made in [14][15], where vertical and lateral track irregularity were identified.

The methodologies previously mentioned aim at the reconstruction of the track geometry profile and generally require a relevant number of sensors and post-processing effort [16]. Although accurate results have been reached, such a system may be not adequate for commercial applications, especially in case a fleet of vehicle is considered, in light of the significant amount of data to be managed, analysed, and aggregated to provide condensed information about the track state. In this respect, the possibility to monitor synthetic indicators may represent an effective alternative. In fact, current maintenance procedures consider the peak and rms values as parameters to trigger the necessary operations [17]. For instance, the degradation of the track conditions was monitored considering the standard deviation of the track geometry measured by the TRV [18]. A step forward would consist in considering indicators directly computed on-board the in-service vehicle, prior to transmission to the ground control room. In [19], the rms of bogie vertical acceleration has been adopted to monitor the vertical track alignment, in terms of rms and peak values. However, over and underestimations of the defect amplitude has been observed in the research work.

In this framework, different machine learning approaches have been proposed in recent years [20]. Probabilistic models based on Markov chains [21] and classification algorithms (Support Vector Machine, decision tree, augmented Bayesian) [22][23] were used to predict the evolution of synthetic indexes related to track geometry.

In this paper, a condition monitoring system for the track longitudinal level is proposed, starting from the methodology we presented in [19]. Specifically, an update of the modelling strategy is here provided, with the aim of improving the accuracy of the predictions of the vertical track alignment. To this end, the designed system relies on multiple regression models to infer the maximum longitudinal level in D1 range (3-25 m wavelength, in accordance with the reference standard EN 13848-5 [17]). Different regressors are considered, ranging from acceleration rms, railway track morphological characteristics (i.e., curves and straight sections), up to direct information of track longitudinal level coming from the diagnostic vehicle. The attention is limited to high-speed applications, where this type of defect represents the major reason for track repair. To this end, data collected during a long-term monitoring campaign of two years along a reference high-speed line (300 km/h) are considered. To verify the accuracy reached, the predictions are validated against the geometry records taken by the TRV. At the present stage of this research, also the acceleration data are measured on the same TRV, that has in fact exactly the same architecture of the commercial high-speed vehicle.

2. Condition monitoring system

In this section, the working principle of the condition monitoring system first presented in [19] is summarized. The system is specifically designed to predict the track longitudinal level based on acceleration measurements taken by an in-service vehicle. However, at the present time, the condition monitoring system is not installed on a commercial vehicle yet. Therefore, the acceleration data measured by the very same diagnostic vehicle performing the track inspection are considered. Major attention is devoted to the longitudinal level in D1 range due to the frequency of these defects, that typically drive maintenance operations along high-speed lines. Synthetic indicators as RMS and the peak longitudinal level (MAX in the following) are computed, considering the RMS bogie vertical acceleration (0-40 Hz frequency range) and pre-built linear regression models. The indicators are sampled along 100 m windows, after that a precise data positioning along the line is achieved by the specifically designed geo-localization algorithm [24]. For a better comprehension of the system architecture, a schematic representation is provided in Figure 1.

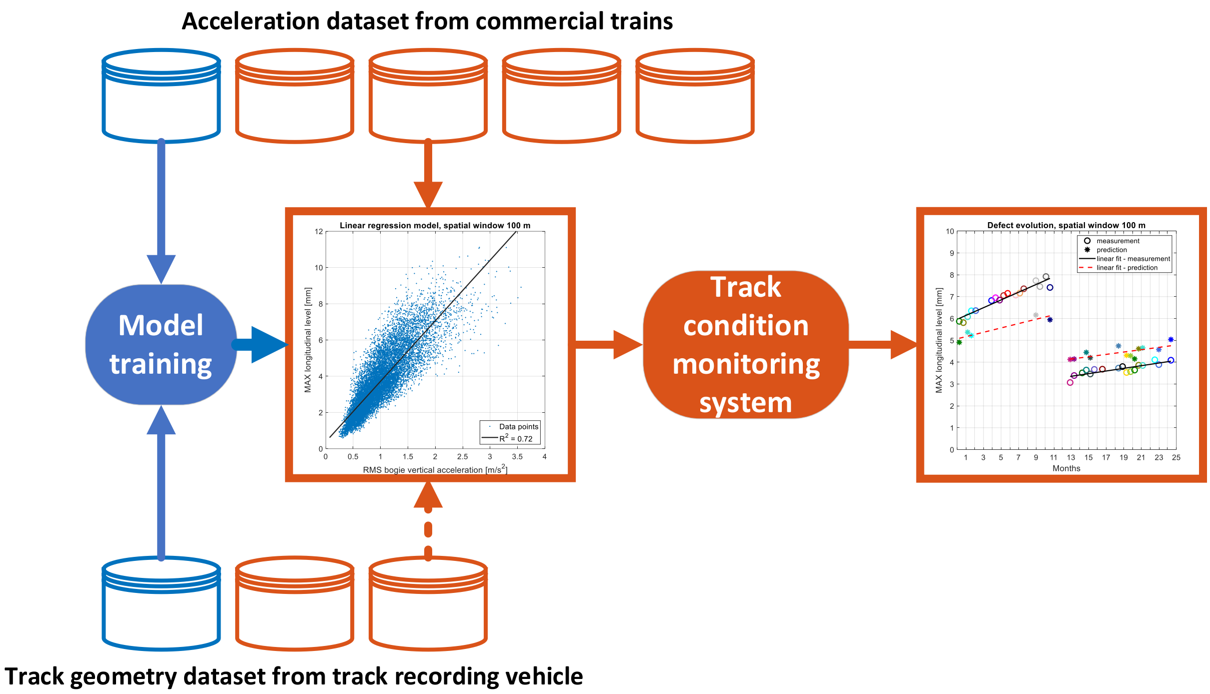

Figure 1. Workflow of the condition monitoring system.

The system consists of two steps. In the training phase, acceleration data gathered by the commercial vehicle are combined to the direct measurements of the track geometry recorded by the diagnostic train. Data are registered along a reference high-speed line of 200 km, considering the vehicle target speed of 300 km/h, during a long-term monitoring period of 18 months. This way, the statistical relevance is guaranteed by including different track conditions (i.e., degraded and renewed track profiles) as well as wheel profiles (that is worn and reprofiled ones). Linear regression models are built considering track geometry and acceleration data registered along the corresponding 100 m windows. Two models are built, to predict both the RMS and the MAX longitudinal level considering the RMS bogie vertical acceleration as input. In the implementation foreseen in [19], the track condition monitoring system takes advantage of the dataset from the TRV only in the training phase of the model, thus a dashed line is adopted in Figure 1.

The second step follows, that adopts the designed models to predict the track geometry parameter of interest. Very accurate results are achieved when inferring the RMS longitudinal level, given the significant value reached by the coefficient of determination $R^2 = 0.93$. A lower degree of correlation is instead registered in the case of the MAX longitudinal level, since the corresponding model is characterized by $R^2 = 0.72$. For completeness, the considered regression model is presented in Figure 2a.

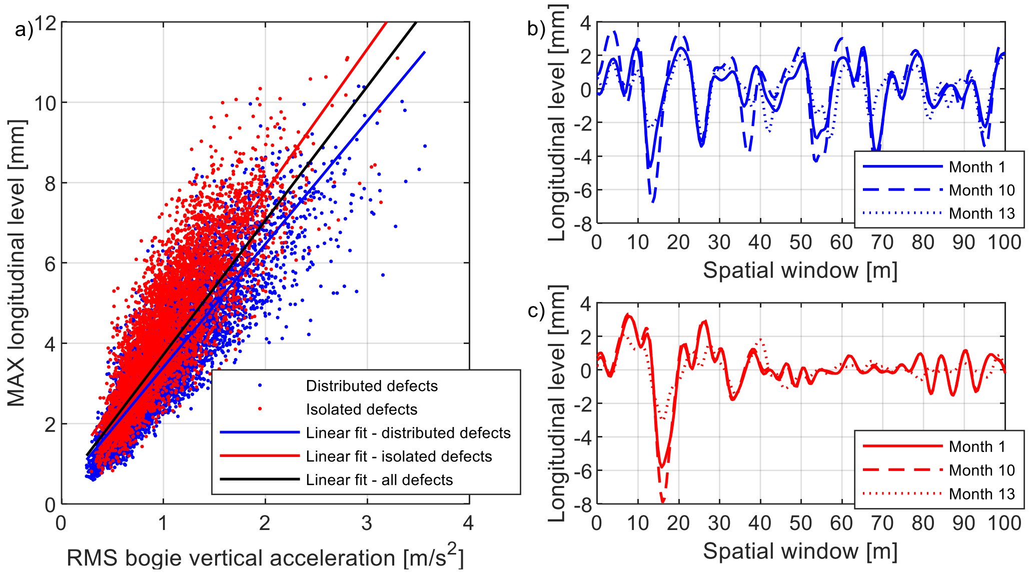

Figure 2. Linear regression model to predict the MAX longitudinal level from the RMS bogie vertical acceleration. a) Regression models for all defects, distributed defects and isolated defects. Data gathered at 300 km/h along the considered high-speed line. Longitudinal level in D1 range (left rail) measured by the diagnostic vehicle along a spatial window of 100 m: b) distributed defect (defect A in the following); c) isolated defect (defect B).

The solid black line in Figure 2a represents the linear regression model adopted to infer the MAX longitudinal level from the RMS bogie vertical acceleration ($R^2 = 0.72$). The model is built considering pairs of acceleration and track geometry data recorded along corresponding spatial windows of 100 m, i.e., the whole dataset reported in Figure 2a (blue and red data contemporarily considered). It turns out that the adoption of this regression model to infer the MAX indicator may lead to significant estimation errors, with both over and under-estimations of the MAX longitudinal level.

This result can be qualitatively explained referring to a preliminary analysis carried out to distinguish the defect nature, that relies on the crest factor (i.e., the ratio of the peak and rms value, referred to as CF) of the acceleration data, that is a measure of how extreme the peaks in a waveform are. In this first analysis, an arbitrary threshold value set to 3.5 was adopted (note that the comments in the following are not depending on this specific threshold); a defect is considered as isolated in case the threshold is exceeded, while lower values identify distributed defects. Therefore, the data available in Figure 2a have been separately considered to realize two different regression models: blue and red colours respectively identify distributed and isolated defects.

For the sake of clarity, Figure 2b shows an example of a distributed defect (CF = 2.9 in Month 1), that consists in a series of adjacent defect inside the considered 100 m window; whereas in Figure 2c an isolated defect is presented (CF = 4.1 in Month 1), that is one single defect inside the window.

From Figure 2a, it is then evident that for the same acceleration rms considered as input, the adoption of a specific model (related to the defect nature) instead of a unique one (representative of the whole railway line) will lead to significantly different results. For instance, consider an acceleration of 2 m/s^2 as input for the prediction of the MAX longitudinal level: a defect of 7.2 mm will be recognised by the adoption of a unique regression model (black line), while 6.4 mm and 7.8 mm will be respectively predicted considering the defect as a distributed (blue line) or an isolated one (red line). This result exemplifies the need to design a modelling strategy that is sensitive to the defect nature, as well as to distinguish the defect before predicting the track geometry from the measured acceleration. As a result, higher prediction accuracy can be reached, making the system more reliable to support the current maintenance strategy.

Willing to increase the model accuracy, an attempt to reduce the window size from 100 m down to 25 m was considered in [19]. Improved results were achieved in terms of coefficient of determination (R^2 passing from 0.72 to 0.77), as well as by correctly predicting the evolution over time of the MAX longitudinal level of a specific 25 m window (limiting the prediction error to 1.5 mm in the analysed cases). However, the reduction of the window size would require a very accurate geo-localization of the data, that may be hard to be met in a commercial application.

Therefore, in this paper an alternative modelling strategy to improve the prediction accuracy is proposed, based on the distinction of the defect type as discussed in Figure 2. To this end, the solution relies on multiple regression models, still adopting a 100 m window to sample the data. In Section 3, different regressors available both from the commercial vehicle and from the diagnostic train itself will be considered, discussing their effectiveness.

3. Multiple regression models

The considered high-speed vehicle is equipped with bogie accelerometers to sense the vertical and lateral acceleration. During the train run, the acquisition board computes synthetic indicators as the rms and peak values in a limited frequency band (0-40 Hz) over 100 m windows, that are then saved and transmitted to the ground server for the analysis. Moreover, the geo-localization algorithm allows associating the track nominal characteristics (i.e., curves and straight sections) to the acquired acceleration data. These parameters have been considered as possible meaningful contributors to the prediction of the MAX longitudinal level, as detailed below:

- the crest factor of the bogie vertical acceleration (computed as the ratio of the peak and rms value at the post-processing stage), that as mentioned provides an insight of how extreme peaks are in a signal;
- the RMS bogie lateral acceleration;
- a categorical variable (Boolean, that assumes values 0 or 1) to distinguish curves and straights tracks.

These variables have been considered as candidate predictors for a multiple regression model, adopting the forward selection procedure. According to the methodology, given N candidate variables, N simple regression models are built (one per each predictor), to select the one that minimise the residual sum of squares. This variable is picked up to build the model, and the procedure is iterated including each one of the remaining N-1 variables, until a stopping criterion is satisfied. To this end, a p-value of 0.05 for the null hypothesis (i.e., no relationship between two variables) was considered, that is the approach commonly adopted from a statistical point of view. In case a p-value higher than the threshold is registered, the variable is proven to be unfit and discarded.

According to the forward selection criterion, the crest factor computed from the bogie vertical acceleration ($CF_{\ddot{z}}$ in the following) proves the highest relevance, and was thus considered in the regression model, that reads like:

$$MAX_{LL} = b_0 + b_1 RMS_{\ddot{z}} + b_2 CF_{\ddot{z}} \quad (1)$$

where the subscript LL stands for the longitudinal level, while the bogie vertical acceleration is identified by z double-dotted. Equation 1 represents the plane that best fits the experimental data by minimising the residual sum of squares, shown in Figure 3. Specifically, a 3D view is proposed in Figure 3a, and side views aligned with the best fit plane are shown in Figure 3b and Figure 3c to better observe the data dispersion.

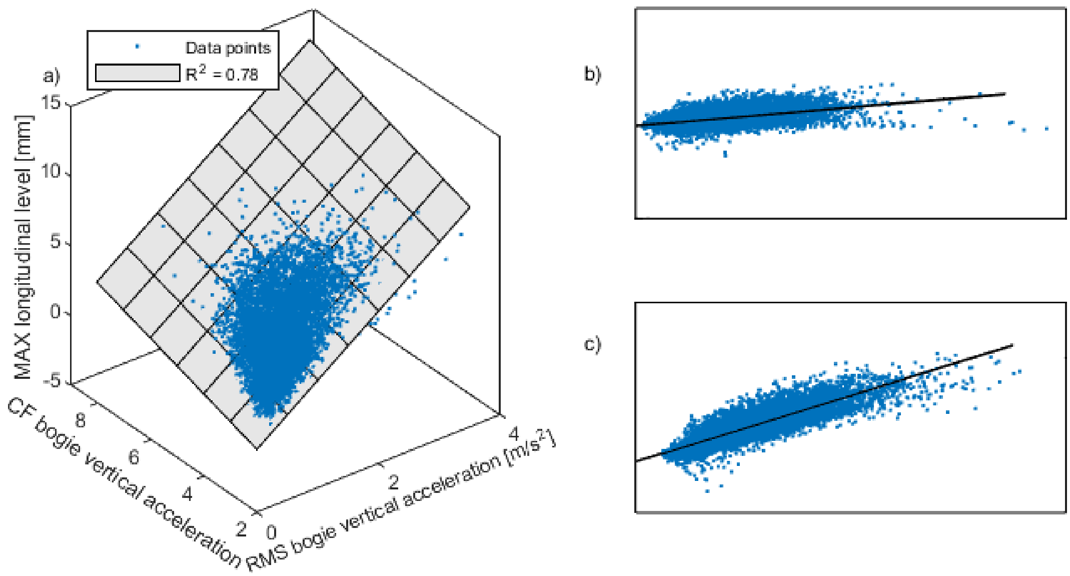

Figure 3. Multiple regression model adopting the RMS and the CF from bogie vertical acceleration to predict the MAX longitudinal level (spatial window 100 m, R^2 = 0.78). a) Data distribution in 3D view; side views aligned to the fitting plane in b) and c).

The considered modelling strategy leads to a coefficient of determination R^2 = 0.78. This result can be regarded as satisfactory if compared to the simple regression model, with an increase with respect to the 100 m window (R^2 = 0.72) and 25 m window as well (R^2 = 0.77).

The results of Figure 3 prove the multiple regression to be a promising strategy to improve the model accuracy, so that the forward selection criterion is then considered to verify the possibility to adopt any additional predictor. The p-value suggests the relevance of both the track characteristics (i.e., curves and straight sections) and the RMS of bogie lateral acceleration. However, no significant increase in the coefficient of determination is reached by their adoption, as R^2 reaches a value of 0.79. Therefore, they have been considered as negligible and discarded for the following analysis.

Once the crest factor was recognised to be the most suited regressor, a step forward is provided in an attempt to better capture the defect nature. The proposed solution directly adopts the track longitudinal level measured by the TRV as the parameter to compute the crest factor, referred to as CF_{LL} in the following. CF_{LL} is computed at the post-processing stage as the ratio between the peak and rms values of the longitudinal level, considering 100 m windows along the railway line, and it is a direct index of the defect nature, not affected by the vehicle dynamics as in the case of $CF_{\ddot{z}}$. Then, it is adopted to build up a multiple regression model, that reads like:

$$MAX_{LL} = b_0 + b_1 RMS_{\ddot{z}} + b_2 CF_{LL} \tag{2}$$

The algorithm schematised in Figure 1 is essentially modified accounting for the fact that the track condition monitoring system will take advantage also of the dataset of the TRV, so that the dashed line will become a “solid” line.

The results of the adoption of the crest factor of the longitudinal level as a predictor are shown in Figure 4, according to the same data representation of Figure 3.

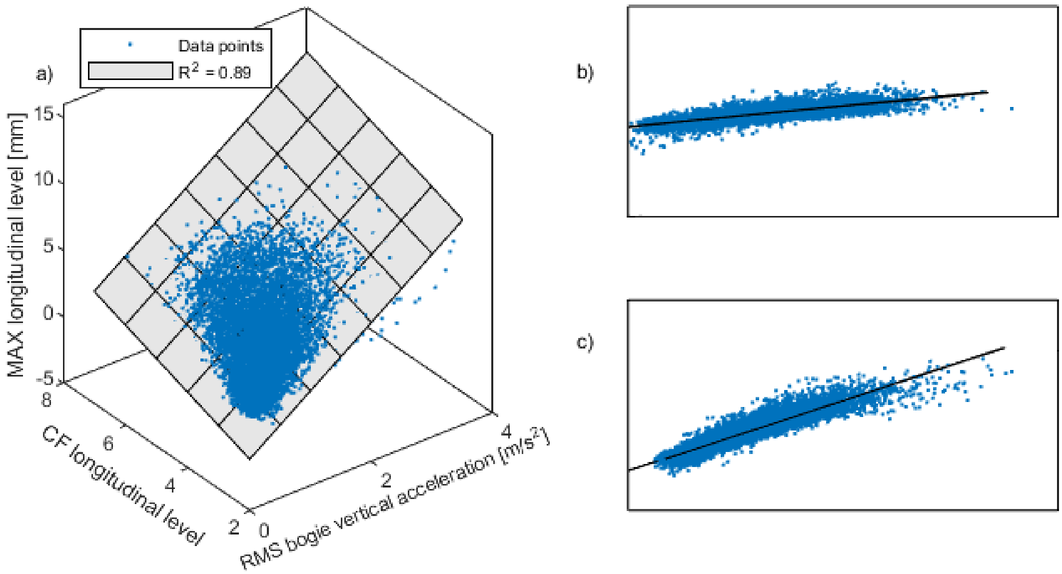

Figure 4. Multiple regression model adopting the RMS bogie vertical acceleration and the CF from longitudinal level to predict the MAX longitudinal level (spatial window 100 m, R^2 = 0.89). a) Data distribution in 3D view; side views aligned to the fitting plane in b) and c).

A significant increase in the coefficient of determination R^2 can be observed, that passes from 0.78 (Figure 3) to 0.89 (Figure 4). The increase in the degree of correlation can be also inferred observing the dispersion of the data around the best fitting plane in Figure 4b and Figure 4c.

To deepen the results, Figure 5 shows the distribution of the residuals as a function of the MAX longitudinal level (Figure 5a) and in terms of probability of occurrence (Figure 5b). In each diagram, the results achieved adopting the regression model of Equation 1 and Equation 2 are shown for comparison.

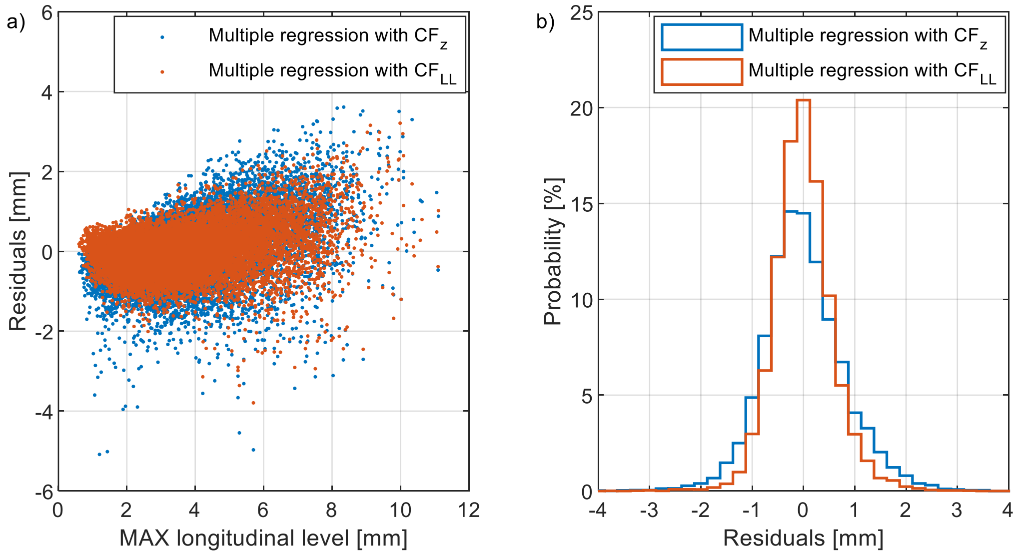

Figure 5. Comparison of the residuals achieved adopting the crest factor from bogie vertical acceleration and from track longitudinal level as additional predictor. a) Residuals as a function of the MAX longitudinal level; b) statistical distribution.

Focusing the attention on Figure 5a, a significant reduction of the absolute error can be recognised in case CF_{LL} is adopted as a predictor, with the datapoints reported in red that are much closer the null value. However, the residual shows a slight dependency with the defect amplitude. This result can be related to the fact that large MAX values are uncommon, the railway line being kept in healthy conditions by maintenance interventions, so that inferring large defects will generally lead to larger prediction errors. In any case, it is worth noting from the distribution of the residuals of Figure 5b that the benefits from the adoption of CF_{LL} as a regressor (red bars) can be also registered in the significant reduction of the tails. For instance, a 14% cumulative probability of getting prediction errors larger than 1 mm is registered in case $CF_{\ddot{z}}$ is adopted, while the probability drops to 5% in case CF_{LL} is instead considered. Therefore, more precise estimates of the MAX longitudinal level are expected by the adoption of the regression model of Equation 2, as will be shown in Section 4.

The significant improvement reached can be associated to the data fusion with direct track geometry measurements. The proposed solution can be regarded as a compromise between an autonomous condition monitoring system (based on acceleration data from the fleet) and the need to improve the model accuracy. In this respect, it is worth recalling that the aim of the condition monitoring system is to support the current maintenance strategy, providing a continuous flow of data in terms of reliable estimations of the track conditions.

From the point of view of the implementation of the methodology, this leads to some implications that are hereafter presented supposing the system to be installed on a commercial vehicle. The regressor $RMS_{\ddot{z}}$ of Equation 2 will be made available on a daily basis, for each train run along the considered railway line. Conversely, CF_{LL} will be at disposal once every diagnostic train run, typically once every two weeks along high-speed lines. Therefore, to infer the MAX longitudinal level at a specific track section adopting the model of Equation 2, CF_{LL} must be kept constant in between two subsequent diagnostic train runs and will be updated as soon as more recent data will be recorded. The proposed strategy relies on the assumption that the evolution rate of the defects can be considered as almost constant in a short time period of two weeks and will be verified and discussed in the next section, where the proposed model will be adopted to predict the MAX longitudinal level.

4. Results and discussion

The capability of the multiple regression model to predict the MAX longitudinal level is hereafter analysed. To this end, the most promising modelling strategy is considered, that relies on the adoption of the crest factor of the longitudinal level (CF_{LL}) together with the RMS bogie vertical acceleration. Several defects were analysed. Applying the methodology, the obtained results are common for any defect belonging to one of the following categories: distributed defect, isolated defect and rapidly evolving isolated defect. For this reason, in the following three defects are discussed, representative of the three typologies. The same defects already analysed in [19] are considered in order to perform a comparative analysis.

At first, in Figure 6, the attention is paid to a distributed defect observed along the considered high-speed line (referred to as defect A). The corresponding signal measured by the diagnostic train along the considered spatial window can be observed in Figure 2b.

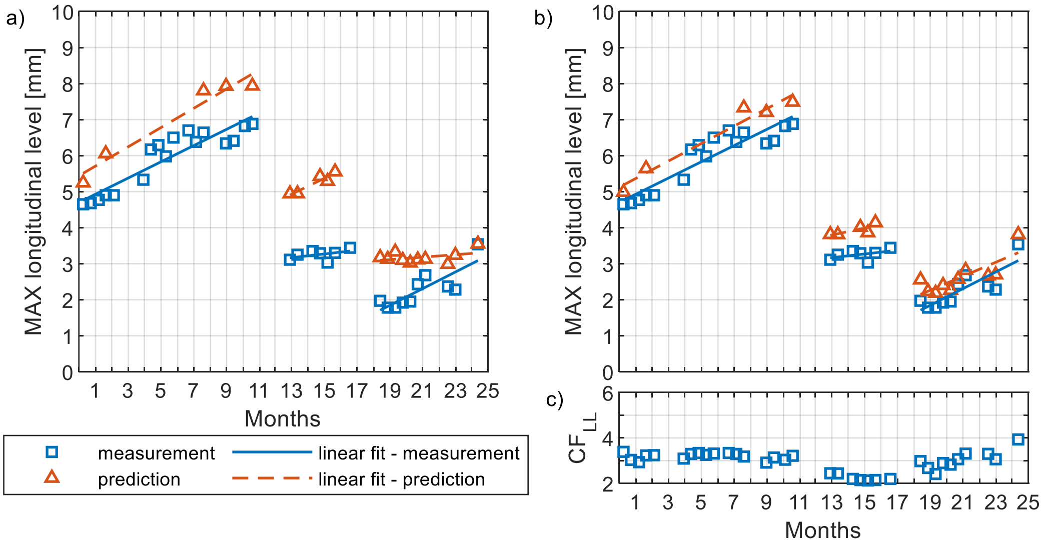

Figure 6. Defect A: distributed defect. Comparison of the MAX longitudinal level measured by the diagnostic train and predicted by the regression models. a) Simple regression model; b) multiple regression model adopting CF_{LL} as additional regressor; c) CF_{LL} as a function of the recording time.

The direct measurements taken by the diagnostic vehicle during the whole monitoring period of two years are reported as blue squared markers as a function of the inspection time. Two maintenance interventions can be identified respectively in Month 11 and 17, so that three separate time periods can be distinguished, characterized by significant reductions of the MAX index. Data belonging to each period are fitted with solid regression lines to give evidence of the rate of defect growth.

In addition, the predictions of the MAX longitudinal level are also shown as triangular red markers. With reference to Figure 1, the first 18 months constitute the training dataset, while the following months belong to the validation set. It can be observed that less predictions than the direct measurements are actually available. This is due to the fact that only acceleration data at the train maximum speed (300 km/h) can be adopted, in light of the dependency of vehicle acceleration over the train speed. When the system will be fully operational, more predictions than direct measurements will be instead available.

More in detail, Figure 6a shows the results achieved adopting the simple regression model with the RMS bogie vertical acceleration as predictor. In Figure 6b, the same defect is inferred with the new proposed multiple regression model (exploiting the RMS of bogie vertical acceleration and the CF from longitudinal level, as shown in Figure 4). Comparing the estimated indexes in Figure 6a and Figure 6b, a significant improvement in the prediction accuracy can be recognised by the adoption of a multiple regression model, since the triangular markers are much closer to the squared ones at any time record available. A slight overestimation of the defect amplitude can be still observed, but when adopting a multiple regression model, the rate of defect growth is correctly predicted during the entire monitoring period, solid and dashed regression lines being always parallel. This significant improvement in terms of both prediction and evolution rate was observed for all the distributed defects.

In Figure 6c the CF_{LL} parameter is also reported in correspondence of each diagnostic train run. Also in this case, it can be observed that track renewal defines three regions. Considering a specific monitoring period, CF_{LL} assumes values that can be regarded as almost constant. For instance, CF_{LL} assumes values close to 3 in the first period, slightly above 2 in the second one, and back to 3 in the last monitoring period (exception made for the very last data at disposal, approaching a value of 4). The change in the average CF_{LL} values corresponds to the variation of the rate of defect growth inferred by the solid regression lines: the higher the CF_{LL} index, the faster the defect evolution. If the attention is now paid to two subsequent time records belonging to the same monitoring period, the CF_{LL} values are very close to one another. This result demonstrates the model assumption to be reasonable, that is to consider constant CF_{LL} values between two runs of the TRV, to infer the MAX longitudinal level considering the acceleration from the commercial fleet.

Figure 7 proposes the same kind of analysis for defect B, an isolated defect whose measurements from the TRV are available in Figure 2c.

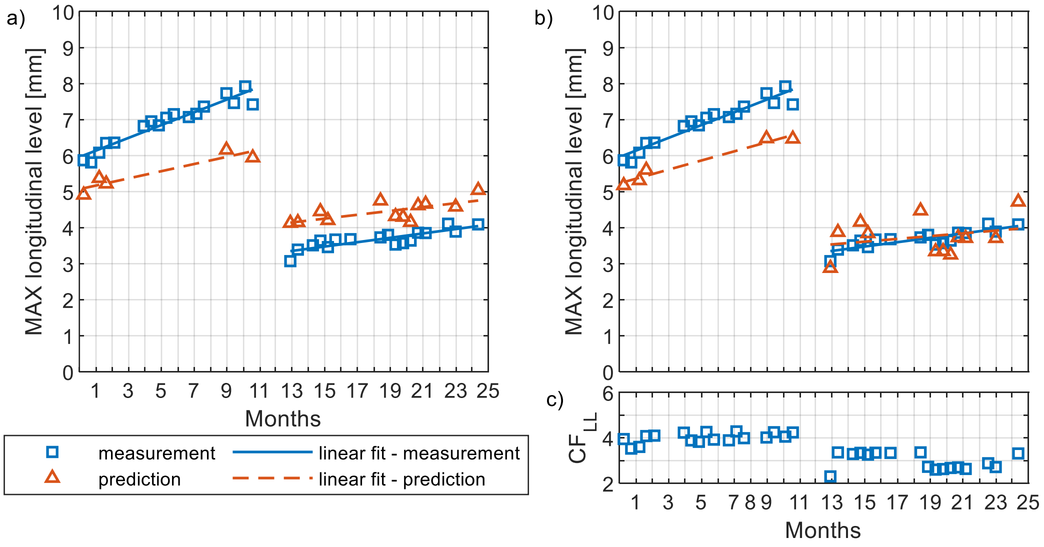

Figure 7. Defect B: isolated defect. Comparison of the MAX longitudinal level measured by the diagnostic train and predicted by the regression models. a) Simple regression model; b) multiple regression model adopting CF_{LL} as additional regressor; c) CF_{LL} as a function of the recording time.

Comparing the results shown in Figure 7a and Figure 7b, the multiple regression model provides better results, since it allows reducing the estimation error. For instance, a maximum error of 1 mm is registered in Month 11 in Figure 7b, whilst 1.5 mm is achieved adopting the simple regression model in Figure 7a at the same time record. The benefits are even higher if the monitoring period after the track intervention is analysed, where the predicted values are closer to the measured ones and a more precise rate of defect growth is identified, with the dashed line that is in very good agreement with the solid one.

Concerning the CF_{LL} values shown in Figure 7c, almost constant values (about 4) are observed in the first monitoring period. This value is consistently higher than the one registered in Figure 6c, that complies with the nature of the considered defect (i.e., isolated). Lower CF_{LL} values are then reached in the period after maintenance operation. This result can be associated to the defect getting closer to a distributed one with small amplitude, that is also an indication of the effectiveness of the track renewal. As a confirmation, attention can be paid to the signals reported in Figure 2c, where the longitudinal level in D1 range recorded before (i.e., Months 1 and 10) and after tamping operation (i.e., Month 13) is shown. It can be observed that before maintenance intervention, the defect can be regarded as an isolated one, with the longitudinal level that shows one single defect (15 m in the considered spatial window). Conversely, after maintenance, the amplitude of the waveform is significantly reduced, making the defect a distributed one of small amplitude. In the end, the signal histories of Figure 2c confirm the change of defect nature inferred by the CF_{LL}, and therefore its capability to distinguish isolated defects (solid and dashed lines in Figure 2c) from distributed ones (dotted line in Figure 2c). The change of defect nature is also the reason for the different prediction accuracy before and after maintenance intervention. In fact, as previously pointed out, the multiple regression model is particularly effective in case of distributed defects.

As a final example, defect C is analysed in Figure 8, that is another isolated defect registered along the considered high-speed line. Out of several defects, it was selected due to its peculiar evolution in time, that shows a piece-wise linear trend. First, a rapid deterioration is observed in the initial phase of the index evolution; then, a significant reduction in the degradation rate can be recognised when the index reaches 6 mm in amplitude. The observed variation in the rate of defect growth can be related to the different phases of the ballast degradation, as recognised in [25][26]. Note that this type of behaviour is preserved also after maintenance took place in Month 11.

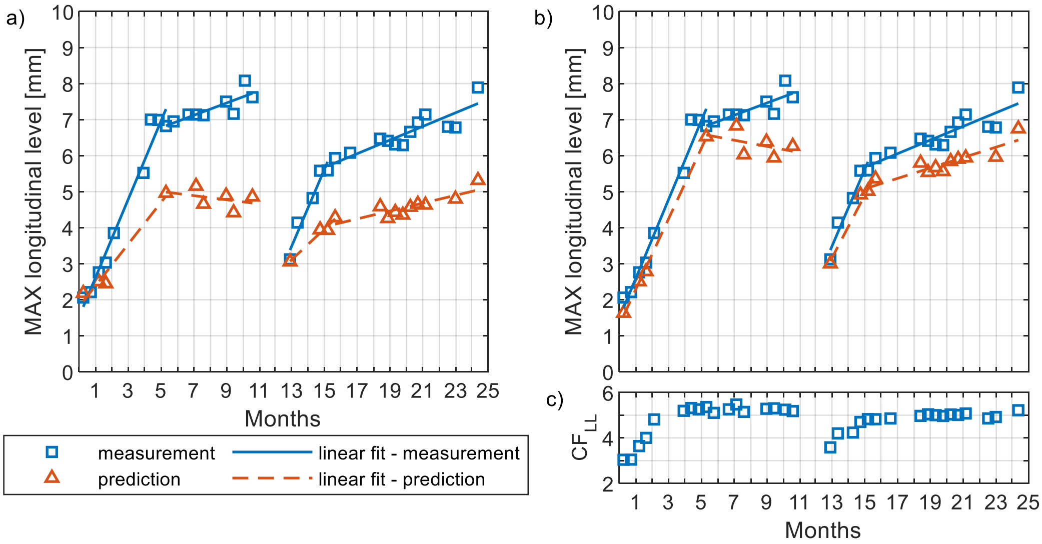

Figure 8. Defect C: isolated defect. Comparison of the MAX longitudinal level measured by the diagnostic train and predicted by the regression models. a) Simple regression model; b) multiple regression model adopting CF_{LL} as additional regressor; c) CF_{LL} as a function of the recording time.

The results achieved by the multiple regression model shown in Figure 8b outperform the ones that relies on just the RMS of bogie vertical acceleration. In fact, the predicted values are in good agreement with the measured data in the entire monitoring period of two years. The only exception is represented by Months 6 to 11, where larger deviations of the predictions are observed. As a result, also the degradation rate predicted by the model is lost, with the dashed red line that is far from the solid blue one, also showing a negative slope that corresponds to an infeasible improvement of the track conditions.

If reference is made to Figure 8c, CF_{LL} shows a similar piece-wise linear evolution both before and after the maintenance intervention. Referring to Months 1 to 11, at the beginning of the monitoring period a significant increase of the CF_{LL} indicator is registered, that passes from 3 (Month 1) to 5 (Month 4). Note that a constant CF_{LL} of about 5.5 is preserved up to the tamping operation carried out in Month 11.

The observed results, together with a fast degradation of the track condition, can be associated to a change in the defect nature, as confirmed by the longitudinal level measured by the diagnostic train in correspondence of the 100 m window under analysis, presented in Figure 9. Out of the period before maintenance, three train runs are considered, respectively associated to Month 1 (solid line), Month 4 (dashed line) and Month 11 (dotted line).

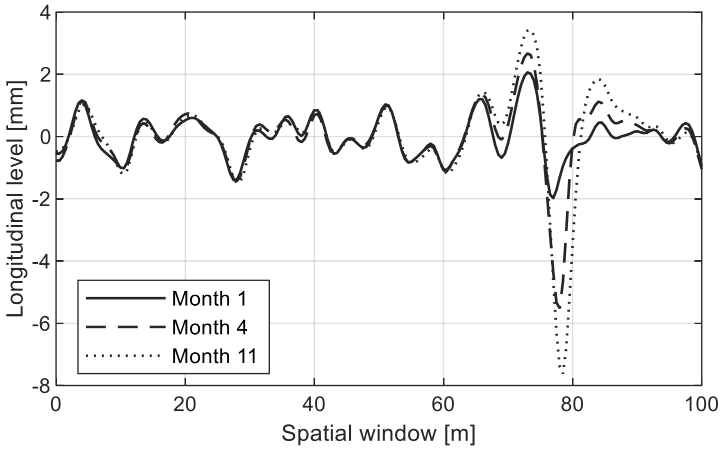

Figure 9. Defect C: isolated defect. Longitudinal level in D1 range (left rail) measured by the diagnostic vehicle along a spatial window of 100 m.

The signals shown in Figure 9 confirm that the low CF_{LL} value observed at the beginning of the evolution (Month 1, solid line) is associated to a distributed defect of small amplitude (2 mm). In the data recorded during Month 4 (dashed line), the defect amplitude presents both a significant increase (5.5 mm) at 80 m of the window, and the peculiarities of an isolated defect. These results provide experimental evidence of the arise of an isolated defect, identified by the increase of CF_{LL}. Finally, the last record before maintenance (Month 11), confirms the presence of an isolated defect of about 8 mm in the considered window.

Moving towards the conclusion, the comparison between the CF_{LL} values of the three defects analysed and the time histories of the corresponding defects prove the capability of the proposed CF_{LL} indicator to distinguish the nature of the defects.

In the end, the multiple regression model proposed in this work can be regarded as a significant improvement of the condition monitoring system presented in [19]. The proposed modelling strategy allows predicting the MAX longitudinal level based on the RMS bogie vertical acceleration (commercial train) and the crest factor of the track longitudinal level (diagnostic train). The achieved results further confirm the possibility to monitor the track longitudinal level based on the acceleration data measured by an instrumented commercial vehicle, that can be adopted to support the current maintenance strategy by means of a continuous flow of data and possibly a more efficient intervention scheduling.

It is worth mentioning that the accuracy of the proposed methodology is expected to be mainly affected by correct data positioning (as addressed in [24]) rather than by the adopted acceleration transducers, given that a proper measurement range is selected. In fact, any accelerometer would offer accuracy and uncertainty suitable for the application. On the other hand, the application on a commercial vehicle may affect the availability of the monitoring system, that can be lower with respect to that installed on a TRV on account of the lower priority of transducers' maintenance. However, the possibility to instrument different bogies of the same vehicle, or different trains, may overcome this limitation.

Conclusions

Based on a previously designed condition monitoring system suitable for high-speed applications, this paper proposes an upgrade of the modelling strategy to predict the track longitudinal level. The aim is that of predicting and monitoring the evolution over time of synthetic indicators, i.e., the MAX longitudinal level in D1 range, in predetermined spatial windows of 100 m length.

In the paper, a multiple regression model is presented, considering different additional regressors other than the RMS bogie vertical acceleration. Out of several candidates, the crest factor of the longitudinal level best improves the model accuracy, adopting a data fusion approach with the direct track geometry measurements. This reflects into a significant increase of the coefficient of determination R^2, that passes from 0.72 (simple regression [19]) to 0.89. In addition, also the prediction error is significantly reduced, leading to more accurate estimations of the MAX index.

The increase in the model accuracy is related to the capability of the CF_{LL} predictor to distinguish the defect nature, given that the railway line is populated by distributed and isolated defects. Regardless of the defect type, the designed system allows to correctly reproduce the rate of defect growth. Moreover, also the quantitative estimations of the MAX indexes are accurate, with a maximum error of 1 mm in the analysed cases. This improvement is particularly significant in the case of distributed defects.

Currently, acceleration data are coming from the same TRV providing track geometry measurements. In the future stage of the research, the designed system could be tested considering the first acceleration data coming from an instrumented in-service vehicle. This would allow discussing the need to manage data redundancy from different bogies of the same train, or even data coming from a fleet of trains. In this respect, attention should be paid to the design of methodologies to integrate the information coming from the diagnostic train, both in terms of CF_{LL} and of inspections timings and locations.

The proposed condition monitoring system could be adopted to support the current maintenance strategy. In case the diagnostic train is not available to operate on a specific railway line (due to maintenance of its own equipment, or in case urgent interventions are required elsewhere), an estimation of the track condition can be achieved. Moreover, given that reliable estimations of the track conditions and of the degradation rate are made available, the time lapse in between the diagnostic train runs could be increased, with benefits in terms of easiness of the intervention scheduling. Finally, in case of sudden changes in a portion of the railway line, daily estimations would allow the identification of possible critical track sections in advance with respect to the diagnostic train, allowing timely interventions and preventing dangerous situations.

References

[1] P. Weston, C. Roberts, G. Yeo, and E. Stewart. “Perspectives on railway track geometry condition monitoring from in-service railway vehicles”. In: Vehicle System Dynamics 53.7 (2015), pp. 1063–1091.

[2] C. P. Ward, P. F. Weston, E. J. C. Stewart, et al., Condition Monitoring Opportunities Using Vehicle-Based Sensors. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. 2011;225(2):202-218. doi:10.1177/09544097JRRT406.

[3] S. L. Grassie, Measurement of railhead longitudinal profiles: a comparison of different techniques. Wear. 1996, 191, 245-251.

[4] H. E. McAnaw. The system that measures the system. NDT&E, 2003, 36, 169-179.

[5] P. Ackroyd, S. Angelo, B. Nejikovsky, J. Stevens. Remote ride quality monitoring of Acela train set performance, Proceedings of the IEEE/ASME Joint Railroad Conference, (2002) pp. 171-178.

[6] P. F. Weston, C. S. Ling, C. Roberts, C. J. Goodman, P. Li, R. M. Goodall, Monitoring lateral track irregularity from in-service railway vehicles. (2007). Proceedings of The Institution of Mechanical Engineers Part F-journal of Rail and Rapid Transit - PROC INST MECH ENG F-J RAIL R. 221. 89-100. 10.1243/0954409JRRT64.

[7] P. F. Weston, P. Li, C. S. Ling, C. Goodman, R. Goodall, C. Roberts, et al., Track and Vehicle Condition Monitoring during Normal Operation Using Reduced Sensor Sets. (2006) HKIE Transactions. 13. 47-53. 10.1080/1023697X.2006.10668031.

[8] S. Nicks. Condition monitoring of the train/track interface. In Proceedings of the IEEE Seminar – Condition Monitoring for Rail Transport System, London, 10 November 1998, pp. 7/1-7/6.

[9] M. Bocciolone, A. Caprioli, A. Cigada, A. Collina, A measurement system for quick rail inspection and effective track maintenance strategy, Mech. Syst. Signal Process. 21 (3) (2007) 1242–1254.

[10] H. Tsunashima, Y.Naganuma, T. Kobayashi. Track geometry estimation from car-body vibration. (2014). Vehicle System Dynamics. 52:sup1, 207-219, DOI: 10.1080/00423114.2014.889836.

[11] J. S. Lee, S. Choi, S.-S. Kim, C. Park, and Y. G. Kim. "A mixed filtering approach for track condition monitoring using accelerometers on the axle box and bogie". In: IEEE Transactions on Instrumentation and Measurement 61.3 (2011), pp. 749–758.

[12] H. Tsunashima, Y. Naganuma, A. Matsumoto, T. Mizuma, H. Mori. Japanese railway condition monitoring of tracks using in-service vehicle. (2011). IET Conference Publications. 2011. 1-6. 10.1049/cp.2011.0587.

[13] H. Tsunashima, Y. Naganuma, A. Matsumoto, T. Mizuma, H. Mori. Condition monitoring of railway track using in-service vehicle. In: Perpinya X, editor. Reliability and safety in railway. InTech; 2012.

[14] E. J. OBrien, P Quirke, C Bowe, et al. Determination of railway track longitudinal profile using measured inertial response of an in-service railway vehicle. (2017) Struct Health Monit.;17:1425-1440.

[15] P. F. Weston, C. S. Ling, C. Roberts, C. J. Goodman, P. Li, R. M. Goodall, Monitoring vertical track irregularity from in-service railway vehicles. (2007). Proceedings of The Institution of Mechanical Engineers Part F-journal of Rail and Rapid Transit - PROC INST MECH ENG F-J RAIL R. 221. 75-88. 10.1243/0954409JRRT65.

[16] A. De Rosa, S. Alfi, S. Bruni. Estimation of lateral and cross alignment in a railway track based on vehicle dynamics measurements. (2019). Mechanical Systems and Signal Processing, Volume 116, Pages 606-623, https://doi.org/10.1016/j.ymssp.2018.06.041.

[17] EN 13848-5:2017. Railway applications – Track – Track geometry quality. Part 5: Geometry quality levels – Plain line, switches and crossings; 2017.

[18] T. Karis, M. Berg, S. Stichel, M. Li, D. Thomas, B. Dirks. Correlation of track irregularities and vehicle responses based on measured data. (2018). Vehicle System Dynamics, 56:6, 967-981.

[19] I. La Paglia, M. Carnevale, R. Corradi, E. Di Gialleonardo, A. Facchinetti, S. Lisi. Condition monitoring of vertical track alignment by bogie acceleration measurements on commercial high-speed vehicles. (2023). Mechanical Systems and Signal Processing, Volume 186, 2023, https://doi.org/10.1016/j.ymssp.2022.109869.

[20] Y. Shang, M. Nogal, H. Wang, A. R. M. Wolfert. Systems thinking approach for improving maintenance management of discrete rail assets: a review and future perspectives. (2023). Structure and Infrastructure Engineering, 19:2, 197-215, DOI: 10.1080/15732479.2021.1936569.

[21] L. Bai, R. Liu, Q. Sun, F. Wang, P. Xu. Markov-based model for the prediction of railway track irregularities. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. 2015;229(2):150-159. doi:10.1177/0954409713503460.

[22] A. De Rosa, R Kulkarni, A Qazizadeh, M. Berg, E. Di Gialleonardo, A. Facchinetti, S. Bruni (2021). Monitoring of lateral and cross level track geometry irregularities through onboard vehicle dynamics measurements using machine learning classification algorithms. In Proceedings of the Institution of Mechanical Engineers. Part F: Journal of Rail and Rapid Transit - ISSN:0954-4097 vol. 235 (1) pp.107-120. DOI:10.1177/0954409720906649.

[23] L. Bai, R. Liu, Q. Sun, F. Wang, F. Wang. Classification-learning-based framework for predicting railway track irregularities. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. 2016;230(2):598-610. doi:10.1177/0954409714552818.

[24] M. Carnevale, I. La Paglia, P. Pennacchi. An algorithm for precise localization of measurements in rolling stock-based diagnostic systems. (2021). Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 235 (7) 827–839.

[25] J. Nielsen, E. Berggren, A. Hammar, F. Jansson, R. Bolmsvik. Degradation of railway track geometry – Correlation between track stiffness gradient and differential settlement. Proceedings

of the Institution of Mechanical Engineers, 095440971881958, Part F: Journal of Rail and Rapid Transit. 234 (2020), https://doi.org/10.1177/0954409718819581.

[26] S. Jovanovic, G. Evren, H. Guler. Modelling railway track geometry deterioration. (2011). Proceedings of the ICE – Transport. 164. 65-75. 10.1680/tran.2011.164.2.65.