

## INDICATORS FOR COMMON CROSSING STRUCTURAL HEALTH MONITORING WITH TRACK-SIDE INERTIAL MEASUREMENTS

MYKOLA SYSYN<sup>a,\*</sup>, ULF GERBER<sup>a</sup>, OLGA NABOCHENKO<sup>b</sup>, YANGYANG LI<sup>a</sup>,  
VITALII KOVALCHUK<sup>b</sup>

<sup>a</sup> Technical University of Dresden, Hettnerstrasse 1/3, Dresden, Germany

<sup>b</sup> Dnipropetrovsk National University of Railway Transport, Blagkevich 12a, Lviv, Ukraine

\* corresponding author: mykola.sysyn@tu-dresden.de

**ABSTRACT.** This paper focuses on the experimental study of an alteration in the railway crossing dynamic response due to the rolling surface degradation during a crossing's lifecycle. The maximal acceleration measured with the track-side measurement system as well as the impact position monitoring show no significant statistical relation to the rolling surface degradation. The additional spectral features are extracted from the acceleration measurements with a wavelet transform to improve the information usage. The reliable prediction of the railway crossing remaining useful life (RUL) demands the trustworthy indicators of structural health that systematically change during the lifecycle. The popular simple machine learning methods like principal component analysis and partial least square regression are used to retrieve two indicators from the experimental information. The feature ranking and selection are used to remove the redundant information and increase the relation of indicators to the lifetime.

**KEYWORDS:** Railway turnout, structural health monitoring, rolling contact fatigue, track-side inertial measurement, feature extraction and selection.

### 1. INTRODUCTION

A turnout is an important part of railway superstructure that provides trains passing from one railway track to another without the run interruption. Turnouts are very expensive compared to ordinary tracks due to their constructive features. Because of their relative small length in overall railway network they share about 10% of the superstructure investment [1]. However, in terms of maintenance costs, this ratio of costs is substantially reversed. The track maintenance costs share about 50% of the overall infrastructure maintenance costs i.e. signal systems, catenary systems and engineering structures [1]. The turnout maintenance takes over a half of the track maintenance costs.

However, turnouts have an indirect influence on operational costs due to a traffic interruption and the following chain of delays in the overall network. The delay costs for the complex west European railway networks could be even more than the maintenance costs. In the study [2], the author estimates that 6% of delays due to unplanned turnout repairs cause 55% of train delays in the overall railway network of Netherlands. The influence of crossings' faults is 28% of the mentioned 55% of the delays.

Railway track health monitoring systems are a promising technique to reach lower operation costs, to prolong the lifecycle and to improve the reliability. The expected result of the systems is a state estimation, fault diagnosis and prediction, timing and extent of necessary maintenance. Modern railway track health monitoring systems can be divided into

on-board and track-side systems. In track-side SHM systems, sensors are mounted on the elements of tracks, usually on rails or sleepers and are used to estimate the track state. In the on-board SHM, sensors are mounted on the axle box or bogies or car bodies of operational trains or track measurement cars. The advantage of the on-board SHM over the track-side ones is that they allow to monitor the long track length and a big number of track devices (turnouts etc.) in a one line passing. The on-board SHM on operational trains have an additional advantage of low cost and almost real-time monitoring. The main disadvantage is the lower quality of information, e.g. for an operational on-board SHM, the measured accelerations are subjected to many external influences like wheel and bearing state and etc. [3–5]. Additionally, the collected information for some track parts can be extremely incomplete for the statistical processing. The track-side SHM provides much more qualitative and quantitative information that allows to take into account all trains passed, and to detect fault kinds.

The application of the SHM in civil engineering and transportation is an object of many studies. The detailed overview of big data analytic applications for a railway transportation system is presented in [8, 9]. The application of statistical methods for selection of ballast compaction features is presented in the study [10]. A railway track condition monitoring with an on-board axle-box inertial measurement and their interpretation is discussed in paper [11, 12]. The track-side wireless system for a sleeper vibrations measurement is proposed by [13] to monitor the track structural health. A systematical overview and anal-

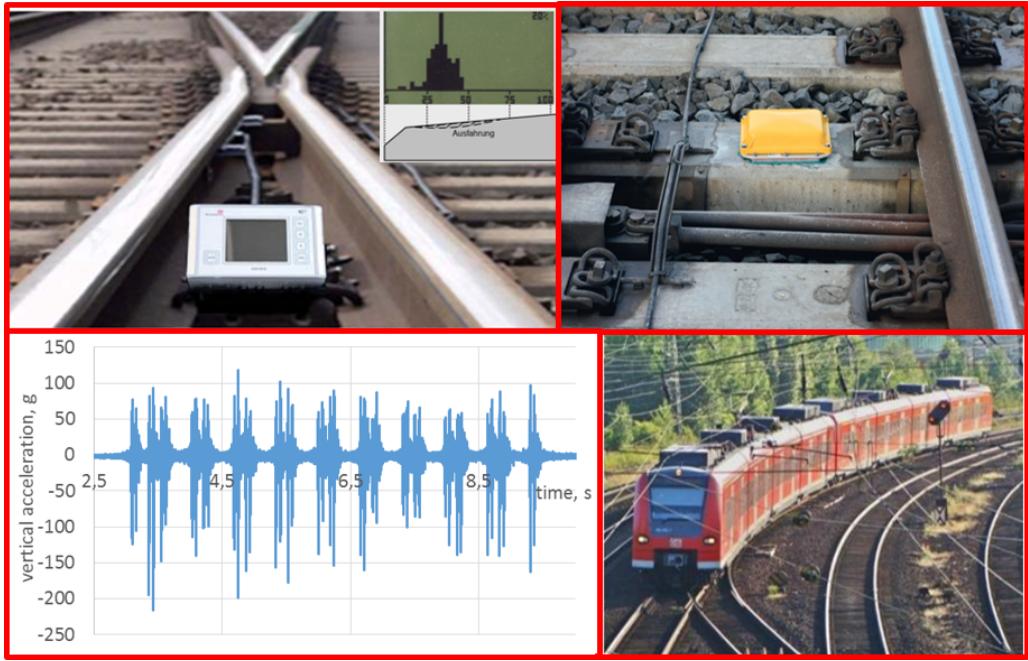


FIGURE 1. The track based inertial measurement systems (left above – ESAH-M/S [6], right above – WSHL [7], down – ESAH-M measurements).

ysis of the SHM system and possible ways of data interpretation are considered in [14]. A machine learning approach for track geometry big data analysis is presented in [15]. The application of reinforcement learning for the calibration of disturbance parameters in the railway operational simulation is proposed in [16].

An on-board inertial SHM is presented in [17–19], axle box ESAH-F measurement system, which is installed at regular trains, allows inspecting many turnouts with low expenses. The application of the system is limited to an existing fault or wear detection without any prediction of common surface damages. The influence of ballast settlements under the common crossing on vehicle based measurements is simulated in [20, 21].

The systems ESAH-M and ESAH-S are track-side inertial measurement SHM systems that are used for the assessment of common crossings loading state on German railways (DB AG). The system ESAH-M (Electronic Analysis System of Crossing - Portable) provides acceleration measurements complemented by positioning sensors on the track [6]. It is a mobile system that is installed on a common crossing after welding and grinding works to improve the rolling surface geometry. Additionally, the system can be used to control the loading state during the lifecycle. The system ESAH-S is a stationary modification that provides continuous acceleration measurements. The diagnostic system WSHL is similar to ESAH-S [7], and is also tested with the DB AG within the project “Intelligent Switch” (Figure 1).

The ESAH-M operation principle as well as a simple mechanical description of vertical wheel transition

while passing the common crossing is shown at Figure 2. During the passing of wheels over the wing rail, the contact point wanders along the conicity of the wheels outwards from point 1 to 2 (top view of Figure 2). Meanwhile, the contact point between wheel and rail drifts along the surface of the bent wing upwards, so that the wheelset moves downwards in the front view from point 1 to 2. Once the position of the wheel flange is in contact with the crossing, the contact point between the wheel and the rail jumps from point 2 on the wing rail to point 3 on the crossing. Afterwards, the wheel rolls again on the primary level of the frog rails from point 3 to 4.

The ESAH-M system consists of two proximity sensors at the wing rail and one 3D acceleration sensor at the spike of a common crossing. The proximity sensors detect the wheel flange movement and time. The longitudinal wheel velocity is determined with 2 proximity sensors. The wheel impact position is found with an analysis of information from the 3D acceleration sensor. The measurement result of the system is the impact density distribution IDD (Figure 2, left above).

However, all the advantages of the SHM application bring new challenges of the measured information utilization. The main problem of track-side inertial measurements is a very big random measurement variance that is much higher than the systematic variance during a crossing’s lifecycle. A lot of studies of inertial measurements have shown a significant uncertainty of the measured data and the necessity of a proper result interpretation. The performance study of ESAH-M in [22] has shown that the calculated coefficient of variation of the measured accelerations can reach up

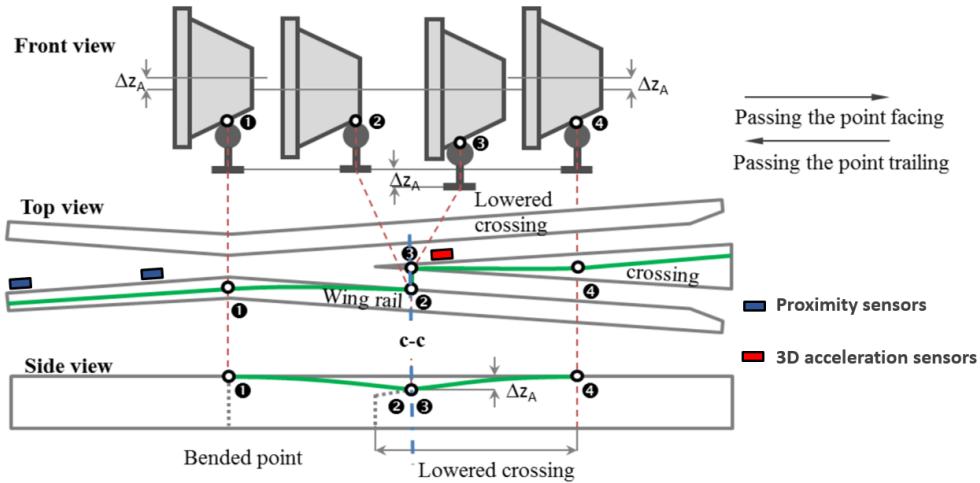


FIGURE 2. Wheel trajectory while passing the common crossing.

to 63%. The studies in [23] have shown one additional systematic error source due to an influence of the impact position relative to the sensor position that can reach up to 72% for a one sensor measurement system. One way to cope with the problem is an application of modern statistics and machine learning methods. The methods are widely used in civil engineering and transportation research. The basics of statistical data processing for railway common crossings as well as optimization techniques were developed in studies [24]. The machine learning based predictive detection approach with an application to the transportation problem, is proposed in the study [25]. An early detection of the common crossing faults with on-board measurements and machine learning approach is presented in [26]. A multiple sensor approach is used in [27] for the beam damages identification based on inertial measurements. The application of principal component analysis for the assessment of bridge's structural health is described in [28]. The feature extraction approach with a statistical classification and machine learning methods for the transportation infrastructure was proposed in [29]. The mechanical simulation is used for an unsupervised feature detection and data fusion in [30], for rail-infrastructure monitoring that is based on operational trains.

Most of the recent studies in statistics and machine learning are focused on non-linear methods [31–33]. However, linear classification and regression methods often have many advantages over the non-linear methods concerning interpretability. For relatively small datasets, they can also be more accurate since they are less subjected to overestimation. A multivariate regression analysis with variable selection and regularization techniques contains many popular methods like stepwise regression, ridge and lasso regression, principal components analysis (PCA) and partial least squares (PLS) regression [34, 35].

The aim of this paper is to develop simple structural health indicators that are based on track side inertial measurements and are able to describe the

state change during the lifecycle of a common crossing. A wavelet analysis is used to extract important information features from inertial measurements. PCA and PLS techniques are used to recover the relation to the lifetime.

## 2. FEATURE EXTRACTION AND PRELIMINARY DATA ANALYSIS

The acceleration measurements with the track-side ESAH-M system are used as a source of statistical information. The spatial accelerations, train speeds and impact point longitudinal position at turnout of steel HB350 with 1/12 crossing angle were measured during 11 time moments of the lifecycle. The inertial dynamic responses from wheels of 4-6 trains, each up to 50 axles, were measured each time. The first visible surface damages that appeared at 28 Mt are shown at Figure 3. The total frog lifecycle was limited to 29 Mt. The overall statistics contains 2263 observations or predictors.

The measured acceleration time series during a common crossing and wheel interaction have a complex form. After the biggest first impact, lower loading peaks that are called second impact (Figure 4, above) follow. The first impact that is usually taken into account is subjected to rather big variation and cannot be considered as the single reliable feature describing the interaction. The “morlet” wavelet transform was used to extract additional spectral features from the random enough time series of acceleration measurements (Figure 4, down).

Four spectral features are found for each axle in 4 frequency ranges 0-23 Hz, 23-160 Hz, 160-1280 Hz and 1280-5000 Hz (Figure 4, down). The frequency range selection corresponds to the own vibration frequencies of wheel and crossing elements. The wavelet coefficients are averaged in frequency-time window for each axle. With the exception of the features, the derivative ones are used like the relation of high to low spectral features. The abbreviations of all col-

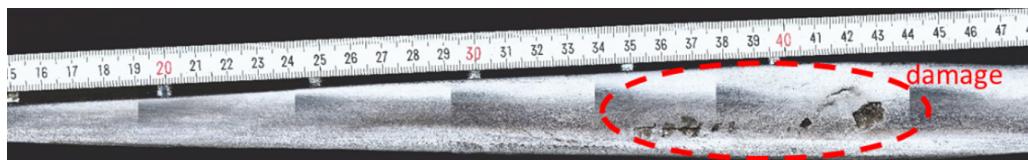


FIGURE 3. Rail contact faults on frog nose rolling surface at 28 Mt.

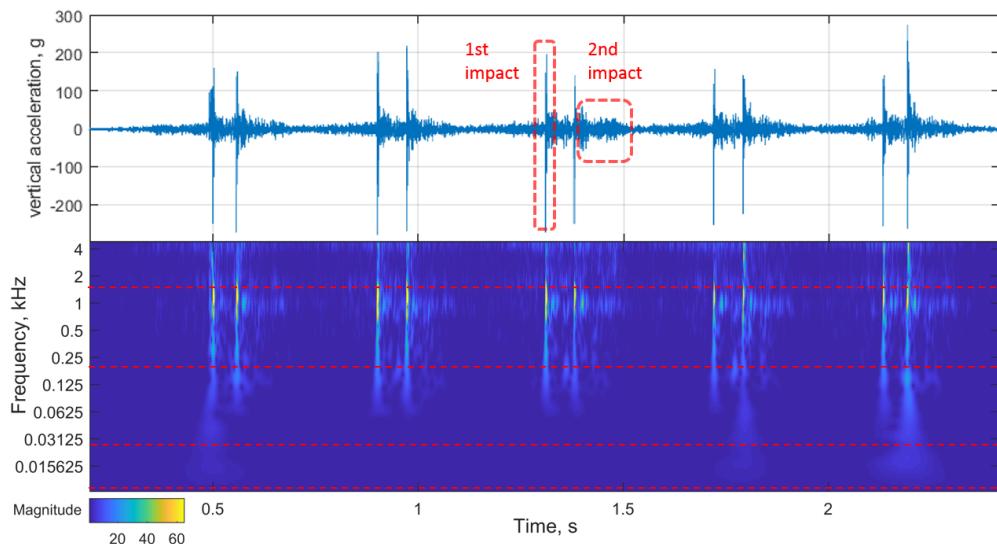


FIGURE 4. Vertical accelerations (above) and their wavelet image (down) for one train.

lected feature set as well as their description are the following:

- maxX** - maximal lateral acceleration;
- maxY** - maximal vertical acceleration;
- maxZ** - maximal longitudinal acceleration;
- IP** - impact point position on a frog nose;
- mx20, my20, mz20** - lateral, vertical and longitudinal acceleration spectral features for the frequency range 1280-5000 Hz;
- mx50, my50, mz50** - lateral, vertical and longitudinal acceleration spectral features for the frequency range 160-1280 Hz;
- mx78, my78, mz78** - lateral, vertical and longitudinal acceleration spectral features for the frequency range 23-160 Hz;
- mx95, my95, mz95** - lateral, vertical and longitudinal acceleration spectral features for the frequency range 0-23 Hz;
- Vi** - train velocity, km/h;
- FRx** -  $(\text{mx20} + \text{mx50}) / (\text{mx78} + \text{mx95})$ ;
- FRy** -  $(\text{my20} + \text{my50}) / (\text{my78} + \text{my95})$ ;
- FRz** -  $(\text{mz20} + \text{mz50}) / (\text{mz78} + \text{mz95})$ .

The preliminary data analysis is carried out before the main statistical analysis. The variation of measured maximal first impact accelerations and their mean values over the lifecycle is shown at Figure 5. The vertical accelerations are 4-5 times bigger than

lateral and longitudinal ones. The variation of measured points is as big as their mean value and much bigger than their lateral and longitudinal measurements. The systematic relation of the acceleration mean values to the lifetime cannot be observed. However, the variation range of measured data or their maximal values can be clearly observed for vertical and lateral accelerations. The relative big accelerations at the initial operation time would be explained with the stabilization process after the crossing renewal. The rolling contact fatigue (RCF) damage (Figure 3) was initially registered at 28 Mt, but it could be supposed from the acceleration change that the fault could already appear at 25 Mt.

The analysis of spectral features variation (Figure 6) shows a similar process. The spectral features groups m20 and m50 have the biggest wavelet coefficients values. The feature m95 has the lowest values and, at the same time, is subjected to the highest variation, so it could be considered as a low informative. The weight of lateral and longitudinal accelerations is the highest in high frequency range. The vertical accelerations mostly differ from two others in the frequency range 160-1280 Hz. The difference between the mean values of longitudinal and lateral accelerations can clearly be observed for the feature m50. None of the spectral features can indicate the unambiguous relation to the lifetime.

With the exception of inertial measurements, the train velocity is measured with the system ESAH-M. The velocity, together with a known initial wheel

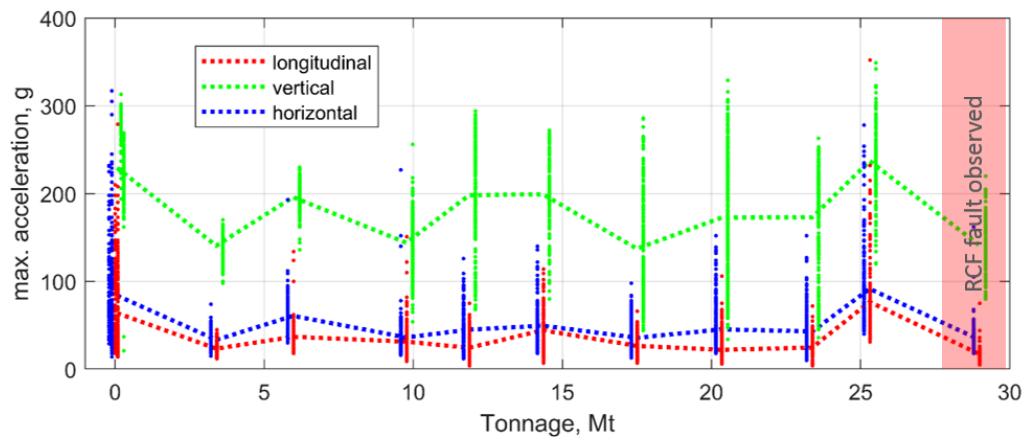


FIGURE 5. The variation of the maximal impact accelerations and their mean values over the common crossing lifecycle.

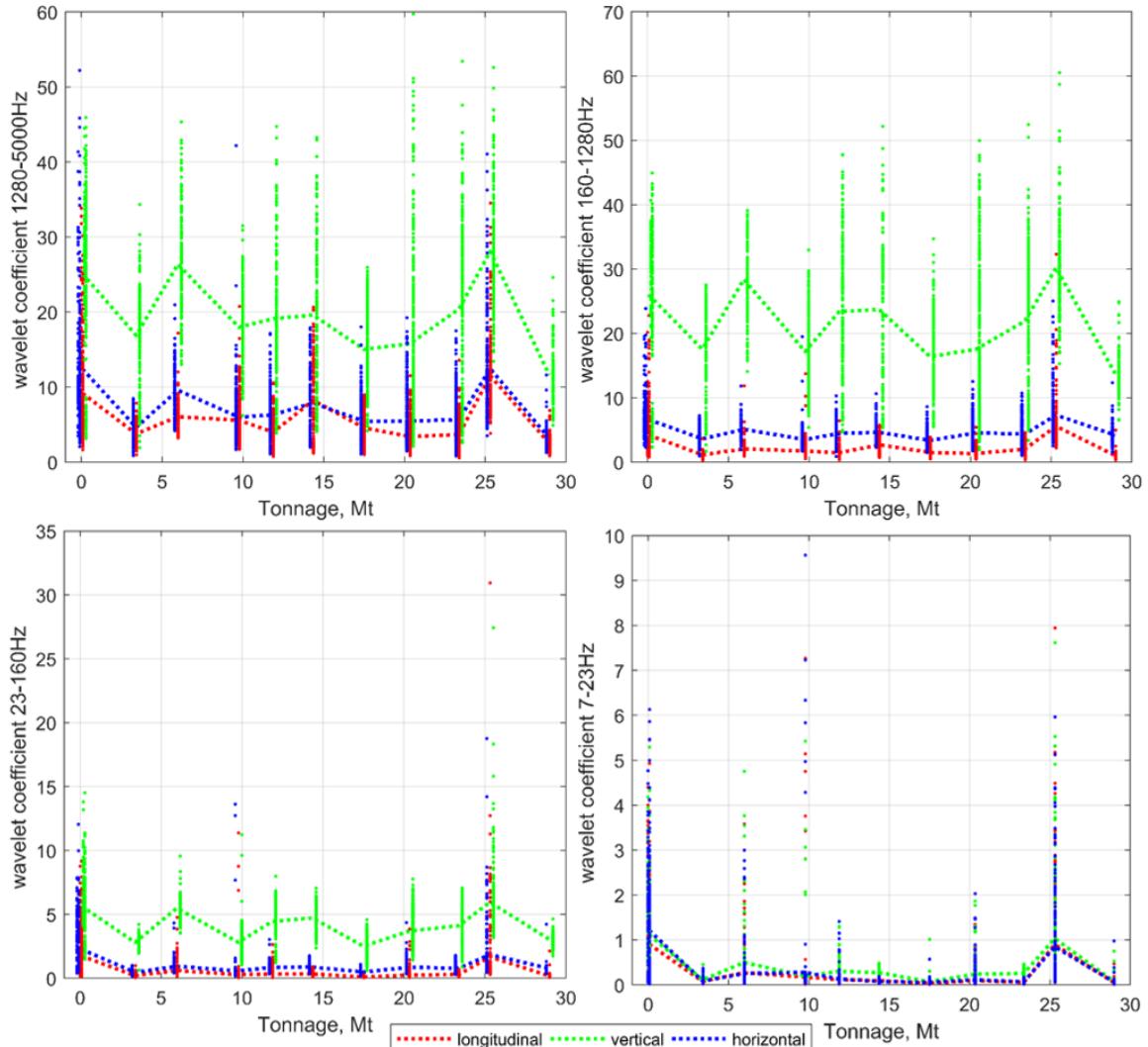


FIGURE 6. The variation of the spectral features m20 (above, left), m50 (above, right), m78 (down, left), m95 (down, right) and their mean values over the crossing lifecycle.

position and maximal acceleration time moment, allows to calculate the impact position point feature IP. Figure 7 shows the distribution of impact positions together with their vertical accelerations in a form of scatter plot of observations. Point colours correspond to common crossing lifetime in passed megatons of trains. The diagram shows no clear prevailing relation of the lifetime to the impact position and vertical acceleration. Nevertheless, some small clusters of red points (early lifetime) with big accelerations, more than 270 g, at the beginning of the impact zone, up to 250 mm, can be visually estimated.

The distribution histograms of the vertical acceleration and impacts position density are depicted in one axes with the scatter plot at Figure 7. The maximal impact density 400-580 impacts/25mm corresponds to the frog nose zone 220-320 mm. It is noteworthy that the RCF damages (Figure 3) have appeared out of the zone in range 350-430 mm where the impact density is more than 5 times lower. The vertical acceleration distribution is more flat with the most number of impacts in range 100-250 g. None of the diagrams indicate some special impact loading distribution in the zone where the RCF fault occurred.

Obviously, neither impact position feature, nor acceleration values, nor their spectral features could be considered as reliable indicators for a common crossing structural health estimation. In the following sections, two statistical learning approaches are considered that allow to reveal the hidden relations in the statistics and to use them to develop the SHM indicators.

### 3. METHODS

#### 3.1. PRINCIPAL COMPONENT ANALYSIS AND FEATURE SELECTION

The principal component analysis (PCA) is used to find out the internal relations in the feature's data set. The idea behind the PCA consists in constructing the linear combinations, or directions, that represent the predictors the best [31, 36, 37]. The combinations are linearly uncorrelated variables, which are called principal components (PC). The  $i$ -principal component can be described as:

$$z_i = W_{11}x_1 + W_{21}x_2 + \dots + W_{p1}x_p$$

where  $z_i$  is the  $i$ -score or  $i$ -principal component;  $W_{11}, \dots, W_{p1}$  - is the loading vector consisting of weights of the first principal component;  $x_1$  - predictors or observations in normalized form.

Figure 8 shows the scatter plot of observation points of the first two principal components space together with the magnitude and direction of each feature's contribution. The points' colours correspond to the same legend as in Figure 7. In contradistinction to Figure 7, where the predictors were plotted in an acceleration-impact position space, the obvious 2 groups can be observed in the PC1-PC2 space. The red point group corresponds to the beginning of the lifecycle and the

blue one to the end. The middle lifetime points or the green group cannot be clearly distinguished from the end lifecycle group.

The two groups with a different lifetime have different PC1 and PC2 values. Both principal components have a relation to the lifetime. The diagram in Figure 8 shows the weight and direction of each feature. The direction can be compared with the direction lifetime change: the features Vi, my50, my20 and maxY are the best correspondent to the state change. The bottom group of features mz95, mx95, my95, mz78 and mx78 are the least related to the state change, even though they have a big weight. That means that the features are low informative for the prediction of the lifetime change. To rank the influence of each feature, their weights are plotted for the first and the second component in Figure 9. The feature selection is based on the feature weights, 13 significant features are used for the development of a structural health indicator. The features mx78, mz78, mx95, my95, mz95, FRy and FRz are not taken into account.

Figure 9 shows that both principal components influence the lifetime or the response value. The response value, tonnage in this case, is not used in the PCA to determine the principal component directions, i.e. the directions are determined in an unsupervised way. Therefore, the PCA has a significant limitation: the directions that explain the predictors the best could not be the best directions for predicting the response. Therefore, each principal component has some influence weight depending on the direction of variation between the two groups with different lifetimes. The optimal direction or the optimal influence part of each principal component should correspond to the maximal coefficient of determination:

$$\max_{k \in [0,1]} R^2((1-k) \cdot PC1 + k \cdot PC2) \rightarrow k$$

Figure 10 shows the relation of two principal components to the common crossing lifetime as well as their combination with the maximal coefficient of determination. The combination is considered as an indicator of a common crossing lifetime.

The regression analysis with the proposed indicator is carried out to estimate its quality to describe a common crossing structural health change during its lifecycle. Figure 11 shows the measured points calculated in form of the indicator parameter, their mean values and the linear regression line with 95% prediction bounds. Different to the measured values (Figure 5, 6), the proposed indicator has an obvious though not high relation to the lifetime. The reason is a high variation of indicators for each measurement time, which causes wide function confidence bounds. At the same time, the indicator points show a high growth of the maximal indicator values with a longer lifetime. However, the points are relatively seldom and have a low influence on the regression line.

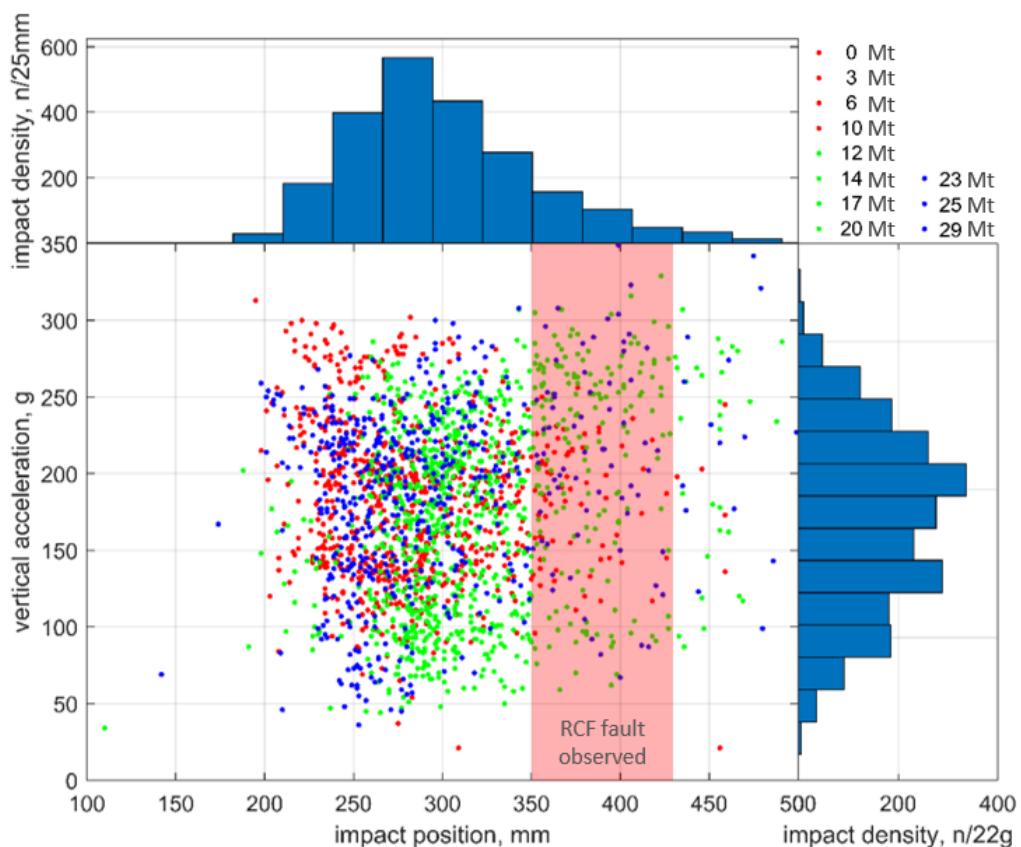


FIGURE 7. The impact longitudinal position and vertical acceleration distributions.

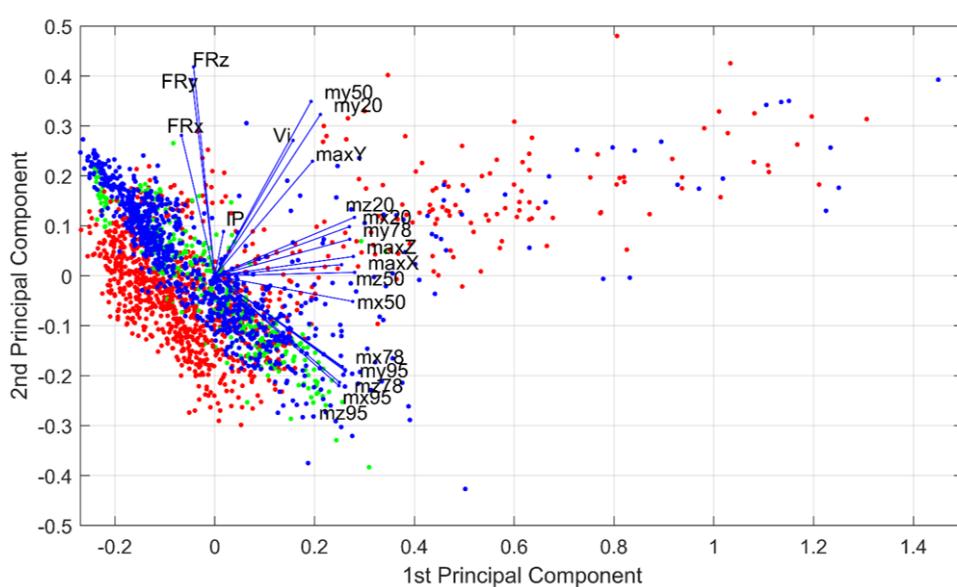


FIGURE 8. The first two principal components and feature vectors.

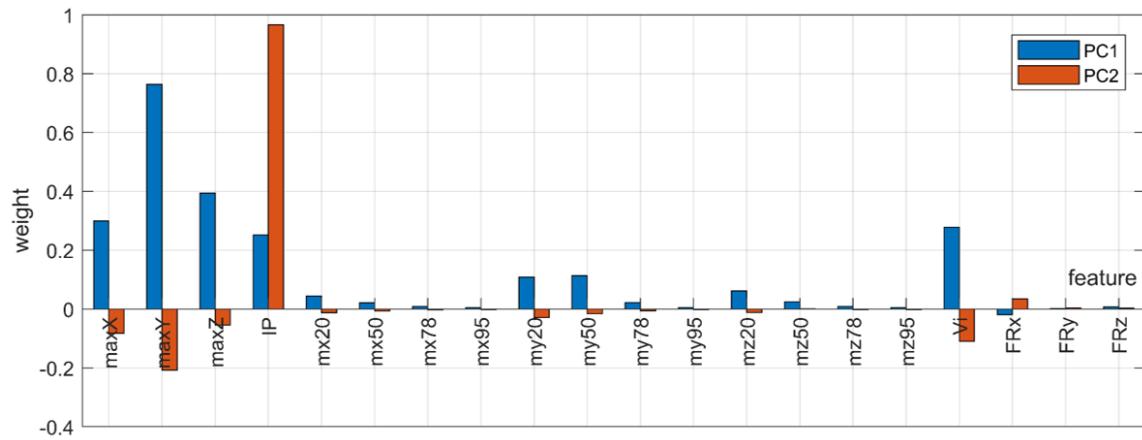


FIGURE 9. The features and weights of PC1 and PC2.

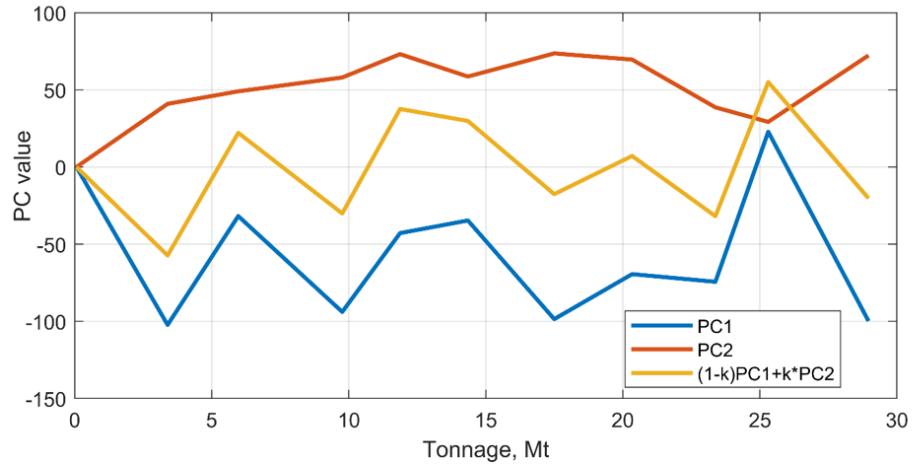


FIGURE 10. The mean PC1 and PC2 values vs. the common crossing lifetime.

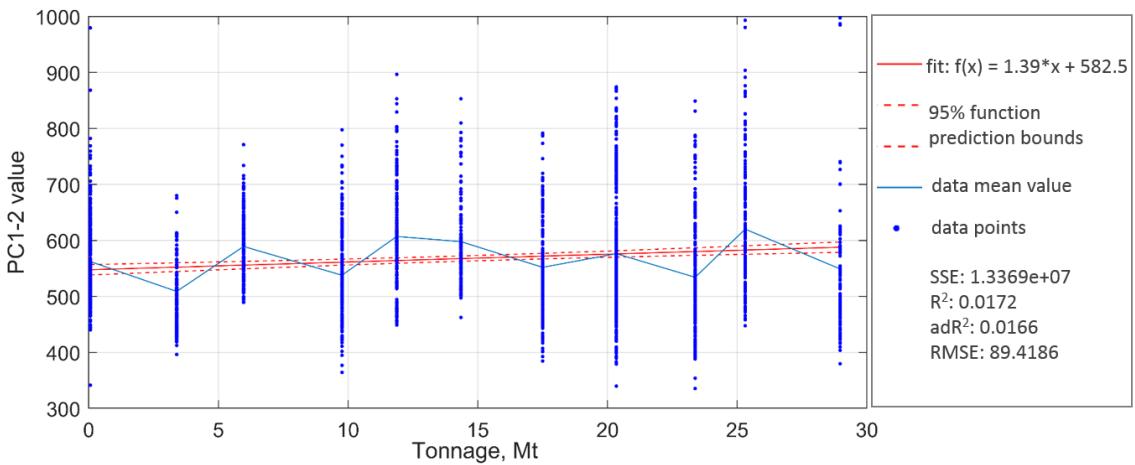


FIGURE 11. The regression of PC1-2 values vs. the frog lifetime.

### 3.2. PARTIAL LEAST SQUARE REGRESSION

The main drawback of the PCA is that it does not take into account the response variable, and therefore there is no guarantee that the principal components found will describe the variation of the response variable. The partial least squares regression (PLSR) [35, 36] is a supervised alternative to the PCA. The method combines the advantages of multiple linear regression and feature transformation methods. The PLS regression is, similar to the PCA, a dimension reduction method. The basic idea of the technique consists in searching the combinations of the predictors that have a large covariance with the response values. This is done in two steps: the identification of the squares of a new set of features that are linear combinations of the original features, and then fitting a linear model via least squares using these new features.

Different to the PCA, the PLS regression determines new features in a supervised way. The new features in the PLSR not only approximate the initial features well as in the PCA, but are also related to the response. The result of the PLS regression is a simple linear model:

$$Y = T \cdot Q + E$$

where  $Y$  - response vector;  $E$  - an error term;  $Q$  - a matrix of regression coefficients;  $T = X \cdot W$  - factor score matrix;  $W$  - weight matrix;  $X$  - vector of predictor variables.

Different to the PCA approach vector, where the first two components were taken into account, for the PLS regression, up to 20 components could be used. But with the high number of components, the risk of overestimation grows. To find out the optimal number of components, the 10-fold cross validation is carried out. Within the validation, the observations are randomly divided into 10 groups, the first of them is used as a validation set while the others are used for the PLSR model fitting. Figure 12 shows the relation of the estimated mean square error for the cross validation set and the set with a resubstitution, i.e. the same data are used for the validation and fitting. In the case of resubstitution, the mean square error (MSE) decreases with the number of components. Whereas for cross-validation, if the number of components is more than 13, the MSE estimation does not significantly decrease, or even increase.

Figure 13 shows the results of the PLS regression with the optimal number of components. The observation points, similarly to Fig. 11, are recalculated to a new indicator. The linear regression fit shows an evident relation to the lifetime. The variation of the indicator points is much lower than the one for the principal component regression. This causes a much lower uncertainty of the function prediction bounds. The coefficient of the determination for the PLS indicator is more than 10 times higher than for the PCA.

## 4. DISCUSSION

The proposed two indicators for a common crossing structural health monitoring with track-side inertial measurements have a significantly better relation to the crossing's lifetime comparing to that of the measured values. The much better results of the PLS indicator than the PCA one could be explained with more components used. The PLS also uses all initial features, whereas the PCA uses only 13 most significant features. Therefore, the PLS indicator provides a better prediction due to the model complexity. The application of complex non-linear regression models would probably give better results, but at the same time, it would bring the problem of interpretation. Additionally, the PCA analysis has shown that some features are redundant, like all the spectral features up to 23 Hz and some of them up to 16 Hz. This is explained with a low sensor sensitivity and measurement error in the frequency range.

The drawback of the indicators is that they describe a state change without any indication of reasons. This means that the appearance of the RCF fault shown on the Figure 3 is not directly related to the indicator and the other unconsidered changes of sleepers or fastenings are more related to the indicator. The problem cannot yet be solved with statistical approach only. The mechanical degradation models could be a promising complement of the statistical ones.

## 5. CONCLUSIONS

This paper presented the statistical interpretation of the measured data on a crossing with the track-side SHM system. The following main results are formulated:

- (1.) Contemporary techniques of the track side SHM systems and modern statistical processing methods are presented.
- (2.) The extraction of time domain and spectral feature set from the measured spatial acceleration is described.
- (3.) A preliminary analysis has shown a low relation of the extracted features to the crossing lifetime.
- (4.) The principal component analysis and partial least square regression methods are introduced.
- (5.) The statistical methods have indicated a significant relation of the measured dataset to the lifetime.
- (6.) Redundant features that are not necessary for the crossing's structural health estimation were found.
- (7.) Indicators that can easily be used for the interpretation of the structural health and are based on linear models are proposed.

## LIST OF SYMBOLS

**RUL** remaining useful life

**SHM** structural health monitoring

**ESAH-F** Electronic Analysis System of Crossing-Vehicle

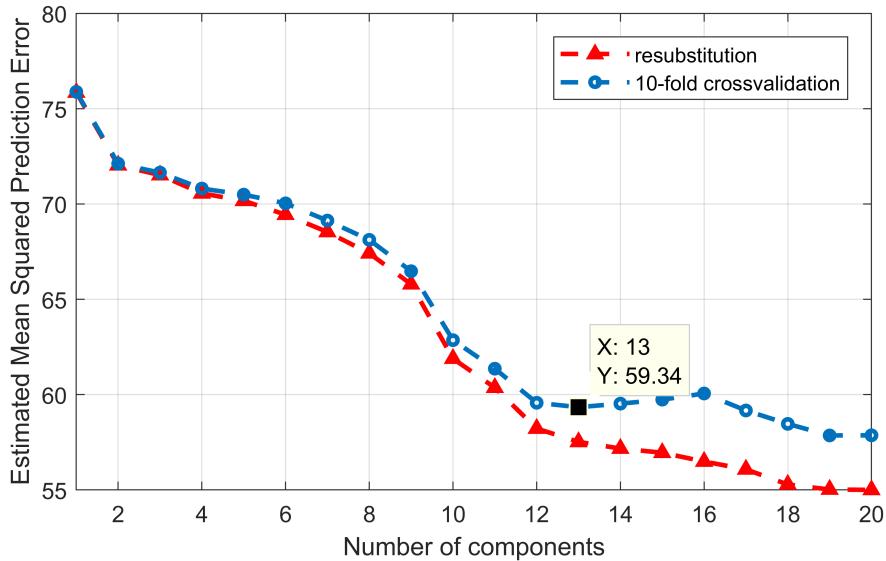


FIGURE 12. Results of two cross-validations with different feature sets.

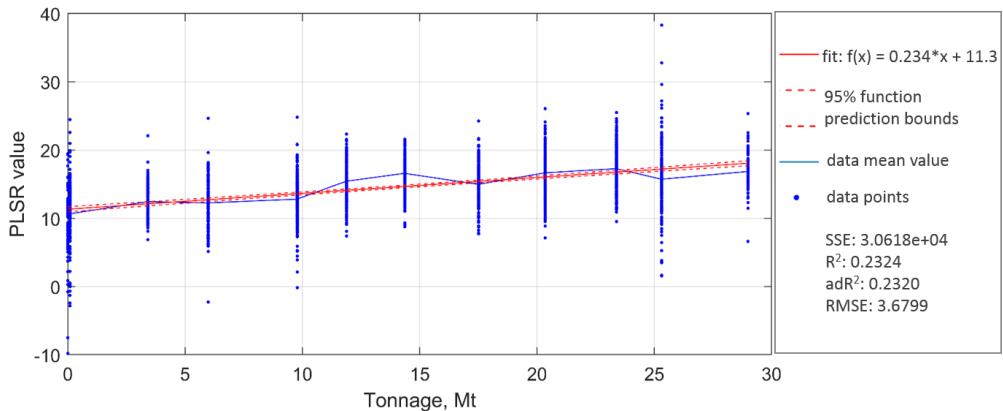


FIGURE 13. The fit of PLSR values vs. the frog lifetime.

**ESAH-M/S** Electronic Analysis System of Crossing-Portable/ Stationary

**WHSL** Turnout sleeper monitoring system

**IDD** impact density distribution

**PCA** principal components analysis

**PLS** partial least squares

**PLSR** partial least squares regression

**RCF** rolling contact fatigue

**MSE** mean square error

**maxX** maximal lateral acceleration

**maxY** maximal vertical acceleration

**maxZ** maximal longitudinal acceleration

**IP** impact point position on a frog nose

**mx20, my20, mz20** lateral, vertical and longitudinal acceleration spectral features for the frequency range 1280-5000 Hz

**mx50, my50, mz50** lateral, vertical and longitudinal acceleration spectral features for the frequency range 160-1280 Hz

**mx78, my78, mz78** lateral, vertical and longitudinal acceleration spectral features for the frequency range 23-160 Hz

**mx95, my95, mz95** lateral, vertical and longitudinal acceleration spectral features for the frequency range 0-23 Hz

**Vi** train velocity

**FRx** lateral spectral feature ratio

**FRx** vertical spectral feature ratio

**FRx** longitudinal spectral feature ratio

#### ACKNOWLEDGEMENTS

The authors greatly appreciate the financial and technical support of Germany Railway Company (DB Systemtechnik GmbH) and WITT Elektronik GmbH.

#### REFERENCES

- [1] L. Fendrich, W. Fengler. *Handbuch Eisenbahninfrastruktur*. Second edition. Springer-Verlag Berlin Heidelberg, 2013. doi:10.1007/978-3-642-30021-9.
- [2] L. Xin. *Long-Term Behaviour of Railway Crossings: Wheel-Rail Interaction and Rail Fatigue Life Prediction*.

- PhD Thesis.: TU Delft, 2017. doi:10.4233/uuid:7ee5405a-85f1-4bd2-b776-2013715c8783.
- [3] U. Gerber, A. Zoll, W. Fengler. Wear and rolling contact fatigue on common crossings of railway turnouts (in german). *Eisenbahntechnische Rundschau* **2**:36–41, 2015.
- [4] V. Kovalchuk, M. Sysyn, J. Sobolevska, et al. Theoretical study into efficiency of the improved longitudinal profile of frogs at railroad switches. *Eastern-European Journal of Enterprise Technologies* **94**(4):27–36, 2018. doi:10.15587/1729-4061.2018.139502.
- [5] V. Kovalchuk, M. Sysyn, Y. Hnativ, et al. Development of a promising system for diagnosing the frogs of railroad switches using the transverse profile measurement method. *Eastern-European Journal of Enterprise Technologies* **92**(2):33–42, 2018. doi:10.15587/1729-4061.2018.125699.
- [6] A. Zoll, U. Gerber, W. Fengler. The measuring system esah-m (in german). *Eisenbahningenieur Kalender 2016* pp. 49–62, 2016.
- [7] T. Böhm, N. Weiss. Turnout analytics - smart sensors and artificial intelligence for the all-round healthy turnout (in german). *Eisenbahntechnische Rundschau* **67**(5):42–45, 2017.
- [8] F. Ghofrani, Q. He, R. Goverde, X. Liu. Recent applications of big data analytics in railway transportation systems: A survey. *Transportation Research Part C: Emerging Technologies* **90**:226–246, 2018. doi:10.1016/j.trc.2018.03.010.
- [9] N. O. Attoh-Okine. *Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering*. Wiley Series in Operations Research and Management Science. John Wiley & Sons, Inc., 2014.
- [10] M. Sysyn, U. Gerber, O. Nabochenko, V. Kovalchuk. Evaluation of railway ballast layer consolidation after maintenance works. *Acta Polytechnica* **59**(1):77–78, 2019. doi:10.14311/AP.2019.59.0077.
- [11] A. Chudzikiewicz, R. Bogacz, M. Kostrzewski, R. Konowrocki. Condition monitoring of railway track systems by using acceleration signals on wheelset axle-boxes. *Transport* **33**(2):555–566, 2018. doi:10.2478/v10174-010-0030-1.
- [12] P. Salvador, V. Naranjo, R. Insa, P. Teixeira. Axlebox accelerations: Their acquisition and time-frequency characterisation for railway track monitoring purposes. *Measurement* **82**:301–312, 2016. doi:10.1016/j.measurement.2016.01.012.
- [13] L. Brajovic, M. Malovic, Z. Popovic, L. Lazarevic. Wireless system for sleeper vibrations measurement. *Communications - Scientific Letters of the University of Zilina* **16**(4):21–26, 2014.
- [14] C. Li, S. Luo, C. Cole, M. Spiriyagin. An overview: modern techniques for railway vehicle on-board health monitoring systems. *Vehicle System Dynamics* **55**(7):1045–1070, 2017. doi:10.2478/v10174-010-0030-1.
- [15] E. Martey, L. Ahmed, N. Attoh-Okine. Track geometry big data analysis: A machine learning approach. *Proceedings - 2017 IEEE International Conference on Big Data* pp. 3800–3809, 2017. doi:10.1109/BigData.2017.8258381.
- [16] S. Rapp, U. Martin, M. Strähle, M. Scheffbuch. Track-vehicle scale model for evaluating local track defects detection methods. *Transportation Geotechnics* **19**:9–18, 2019. doi:10.1016/j.jrtpm.2016.03.001.
- [17] U. Gerber, A. Zoll, W. Fengler. Vehicle-based assessment of wear on common crossings (in german). *EI Eisenbahningenieur* **5**:26–30, 2013.
- [18] S. Scholz, R. Lommock. Models for onboard train diagnostics data to improve condition-based maintenance. *Proceedings of the 16th International Conference on Automated People Movers and Automated Transit Systems* pp. 85–93, 2018. doi:10.1061/9780784481318.010.
- [19] Y. Cui, U. Martin, W. Zhao. Calibration of disturbance parameters in railway operational simulation based on reinforcement learning. *Journal of Rail Transport Planning and Management* **6**(1):1–12, 2016. doi:10.1016/j.jrtpm.2016.03.001.
- [20] M. Sysyn, U. Gerber, D. Gruen, et al. Modelling and vehicle based measurements of ballast settlements under the common crossing. *European Transport - Trasporti Europei* **71**:1–25, 2019.
- [21] O. Nabochenko, M. Sysyn, V. Kovalchuk, et al. Studying the railroad track deterioration impairment as a result of an uneven subsidence of the ballast layer. *Eastern-European Journal of Enterprise Technologies* **97**(1):50–59, 2019. doi:10.15587/1729-4061.2019.154864.
- [22] X. Liu, V. Markiene, I. Shevtsov, R. Dollevoet. Experiment study of key parameters in turnout crossing degradation process. *Proceedings of the 10th International Conference on Contact MechanicsColorado, USA 1-3 September* pp. 122–127, 2015.
- [23] M. Sysyn, V. Kovalchuk, D. Jiang. Performance study of the inertial monitoring method for railway turnouts. *International Journal of Rail Transportation* **7**(2):103–116, 2019. doi:10.1080/23248378.2018.1514282.
- [24] B. E. Glyusberg. *Methods for optimizing the basic elements of turnouts and their application to common crossing of mass structures (in Russian)*. Dr. Thesis.: Railway Research Institute, Moscow, 1989.
- [25] I. Bolkeny, V. Fuvesi. Ai based predictive detection system. *Pollack Periodica* **13**(2):137–146, 2018. doi:10.1556/606.2018.13.2.14.
- [26] M. Sysyn, D. Gruen, U. Gerber, et al. Turnout monitoring with vehicle based inertial measurements of operational trains: a machine learning approach. *COMMUNICATIONS Scientific Letters of the University of Zilina* **21**(1):42–48, 2018.
- [27] M. Sokol, M. Venglar. System identification of a composite beam. *Pollack Periodica* **12**(3):43–54, 2017. doi:10.1556/606.2017.12.3.5.
- [28] S. Sein, J. Matos, J. Idnurm. Statistical analysis of reinforced concrete bridges in estonia. *Baltic Journal of Road and Bridge Engineering* **12**(4):225–233, 2017.
- [29] M. Hulkova, K. Pavelka, E. Matouskova. Automatic classification of point clouds for highway documentation. *Acta Polytechnica* **58**(3):165–170, 2018. doi:10.14311/AP.2018.58.0165.

- [30] G. Lederman, S. Chen, J. Garrett, et al. A data fusion approach for track monitoring from multiple in-service trains. *Mechanical Systems and Signal Processing* **95**:363–379, 2017. doi:10.1016/j.ymssp.2016.06.041.
- [31] T. Hastie, R. Tibshirani, J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second edition. Springer-Verlag New York, 2009. doi:10.1007/978-0-387-84858-7.
- [32] X. Kong, C. Hu, Z. Duan. *Principal component analysis networks and algorithms*. Springer Singapore, 2017. doi:10.1007/978-981-10-2915-8.
- [33] S. Shai, B. Shai. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014. doi:10.1017/CBO9781107298019.
- [34] G. James, D. Witten, T. Hastie, R. Tibshirani. *An Introduction to Statistical Learning with Applications in R*. Springer Verlag New York, 2013. doi:10.1007/978-1-4614-7138-7.
- [35] R. Rosipal, Krämer. *Overview and Recent Advances in Partial Least Squares*. Springer Berlin Heidelberg, 2006. doi:10.1007/11752790\_2.
- [36] M. Kuhn, K. Johnson. *Applied Predictive Modeling*. Springer-Verlag New York, 2013. doi:10.1007/978-1-4614-6849-3.
- [37] N. Attoh-Okine, A. Lasisi. Principal components analysis and track quality index: A machine learning approach. *Transportation Research Part C: Emerging Technologies* **91**:230–248, 2018. doi:10.1016/j.trc.2018.04.001.