

## COMMON CROSSING FAULT PREDICTION WITH TRACK BASED INERTIAL MEASUREMENTS: STATISTICAL VS. MECHANICAL APPROACH

<sup>1</sup> Mykola SYSYN\*, <sup>2</sup> Ulf GERBER, <sup>3</sup> Olga NABOCHENKO  
<sup>4</sup> Vitalii KOVALCHUK

<sup>1,2</sup> Institute of Railway Systems and Public Transport, Technical University of Dresden  
Dresden, Germany, e-mail: <sup>1</sup>mykola.sysyn@tu-dresden.de, <sup>2</sup>ulf.gerber@tu-dresden.de  
<sup>3,4</sup> Department of the Rolling Stock and Track, Lviv Branch of Dnipropetrovsk National  
University of Railway Transport, Lviv, Ukraine  
e-mail: <sup>3</sup>olganabochenko@gmail.com, <sup>4</sup>kovalchuk.diit@gmail.com

Received 21 November 2018; accepted 19 April 2019

**Abstract:** The analysis of track based inertial measurements for common crossing fault detection and prediction is presented in the paper. The measurement of spatial acceleration in common crossing spike and impact position during overall lifecycle are studied regarding to rolling surface fatigue degradation. Two approaches for retrieving the relation of inertial parameters to common crossing lifetime are proposed. The first one is based on the statistical learning method - t-SNE algorithm that helps to find out similarities in measured dataset. The second one is a mechanical approach that handles the data with a fatigue and contact models. Both approaches allow the significant improvement of the common crossing fault detection as well as its early prediction

**Keywords:** Common crossing, Rolling surface fatigue, Inertial measurement systems, Fault detection, Predictive maintenance, t-SNE

### 1. Introduction

Railway turnout is a relatively expensive part of railway superstructure and at the same time has disproportionately short lifecycle [1]. The lifecycle of turnout is limited with those of its highly loaded elements like switch points and common crossings. A common crossing degradation process is characterized with mainly rolling surface fatigue damages that are difficult to predict with ordinary inspection methods [2]. The unexpectedly appeared faults in common crossing usually demand the train speed

---

\* Corresponding Author

limitation and often unplanned maintenance works with long time traffic interruption. The unpredictable delays of some trains due to turnout faults can cause far-reaching consequences to highly interrelated railway transportation system [1]. Therefore the common crossing is one of the limiting factors of railway infrastructure.

The rolling surface state of common crossing is inspected with many different systems [3], like profile and surface scan or video inspection, microstructure imaging, eddy current and ultrasound, vehicle based and track based inertial measurements. However, none of the systems cannot yet replace the conventional inspection way with expert judgment that is based on visual estimation and acoustic perception of train impacts.

Vehicle based inertial measurement, e.g. axle box (Electronic System-Analysis in Crossing - Vehicle) ESAH-F measurement system [4] that is installed at regular trains, allows inspecting many turnouts with low expenses. The application of the system is limited to existing fault or wear detection without prediction of common surface damages.

The systems Electronic System-Analysis in Crossings - Mobile (ESAH-M) (in Austrian) and Electronic System-Analysis in Crossings - Stationary (ESAH-S) (in Austrian) are track based inertial measurement that are used for the assessment of common crossings loading state on German Railways (DB AG). The system ESAH-M provides acceleration measurements complemented by positioning sensors on track [5]. It is a mobile system that is installed on 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 similar to ESAH-S diagnostic system is Sleeper Voids Measurement System (WSHL) (in Austrian) [6], which is also tested with DB AG within the project 'Intelligent switch'.

Many research papers as well as a long time measurement experience with inertial measurements have shown a significant uncertainty of the measured data. The key parameters for the turnout with the application of ESAH-M, were considered in the experimental study [7]. The study has shown that the calculated coefficient of variation of the measured accelerations can reach up to 63%. The main problem of track based inertial measurements is a very high random measurement variance that is much higher than the variance due to common crossing state change during its lifecycle. That is explained that ESAH-S system is subjected to a big number of unknown and unconsidered factors, like the wheel profile irregularities and wear, lateral wheel position, etc. One additional systematic error source was studied in [8], where the influence of impact position relative the sensor position was considered. It was experimentally proved that the varying impact position due to the properties of wave propagation could cause up to 72% variation of accelerations measured with one sensor system. This altogether leads to a high deviation of measurement results that makes it very difficult to estimate the changes of common crossing state.

The problem is a typical application field for data mining methods. The methods are widely used in transportation research. The assessment of bridges' conditions with the principal component analysis of data collected during visual inspections is described in [9]. The promising technique for application in transportation histogram based image segmentation method that is presented in [10]. The extraction of objects of interest from

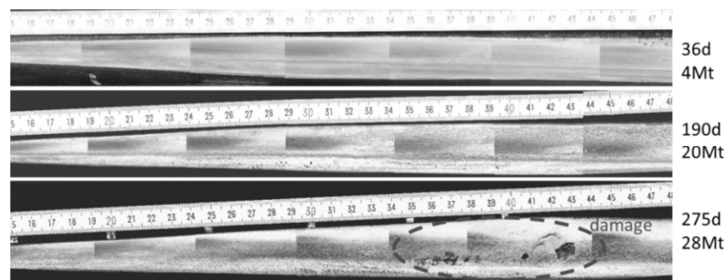
point clouds and their automatic classification are used in [11] for highway image data processing. The machine learning based predictive detection approach that could be used for many similar transportation problems, is introduced in the study of [12]. The statistics based feature selection for evaluation of railway ballast consolidation is proposed in [13]. A data fusion approach is proposed in Lederman et al. [14], to combine acceleration measurements from multiple trains and thus enables to detect track changes earlier and more reliable. A data mining approach with sequential feature selection is proposed in the study [15] to retrieve the statistical correlation between the inertial measurements of operational trains and railway degradation.

Except of purely statistical approaches, the combined ones with mechanical model approaches are recently used. The analysis carried out in [16] has shown, that despite of all power of big data and deep learning methods that are based on enormous datasets, the best practical results can be reached in combination with mechanical engineering approach. Thus, the mechanical simulation is used for unsupervised change and feature detection in [14], while rail-infrastructure monitoring is based on operational trains.

In the present paper, the statistical and mechanical approaches are used to recover the statistical correlation between the inertial measurement parameters and common crossing lifetime and state.

## 2. Preliminary data analysis

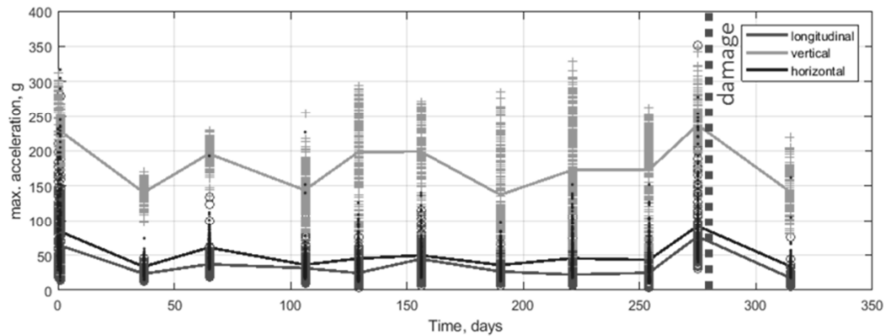
The statistical analysis is based on the acceleration measurements of ESAH-M system of DB System-technik together with High Resolution Photo Imaging (HRPI) inspection. The measurements were carried out at turnout with 1/10 crossing angle, built up common crossing from rail steel R350. The inertial parameters were measured in 11 lifetime days during common crossing lifecycle 315 days or 29 Mt. Each lifetime days included 4-6 measurements of train passing with up to 50 axles. The measurement for 1 axle consists of 3 acceleration components and the impact position. The overall statistics contains 2263 observations. HRPI inspection was carried out until the first visible surface damages appeared at 275 day (or 28 Mt (million tons)). *Fig. 1* shows the images from the lifecycle beginning, and the first damages.



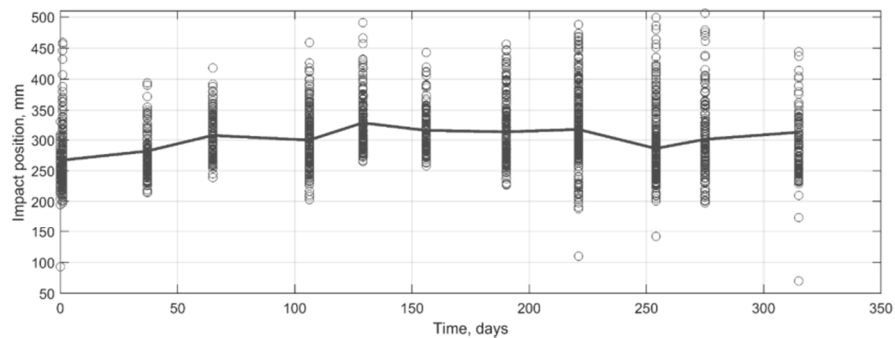
*Fig. 1.* The rolling surface state inspection images

The change of maximal impact accelerations and their mean values over the switch lifecycle is depicted in *Fig. 2*. All three acceleration components are characterized with

a high random variation of measured values. No systematic change of acceleration mean values can be observed. The increased accelerations after 250 days of the lifecycle can be explained with the appeared damages itself. The similar progress can be observed at *Fig. 3* of impact longitudinal position. The insignificant systematic mean value variation appears in first 100 days and after that is almost constant in range 290-320 mm. The mean value variation is many times covered with the random variance of each measurement.



*Fig. 2.* Maximal impact accelerations and their mean values over the common crossing lifecycle



*Fig. 3.* Impact longitudinal positions over the common crossing lifecycle

Apparently, no conclusions about the common crossing state could be done with only mean value analysis of acceleration components and impacts position. In the following chapters two approaches are proposed that enable us to recover the hidden relations in the statistic and use them to estimate common crossing state.

### 3. Statistical approach to recover the relation to life time

The algorithm t-distributed Stochastic Neighbor Embedding (t-SNE) is used for the statistics structure analysis. This is a machine learning algorithm for dimensionality reduction that is well-suited to visualizing high-dimensional data [17]. High

dimensional data are embedded in low two or three dimensions in a way that respects similarities between points. The result of embedding is a plot of original data points that are grouped in clusters with same similarity. Therefore, the result is only qualitative expression of similarities in the statistic without considering their reasons.

The t-SNE algorithm is based on two calculation steps. In the first step the similarity matrix is determined on construction of a distance probability distribution over pairs of high-dimensional data. The different distance metrics can be used, such as ‘Chebychev’, ‘Mahalanobis’ metrics etc., the usual is ‘Euclidean’. The value of similarity is controlled with a parameter of the effective number of local neighbors of each point that is called perplexity. In the second step a similar probability distribution over the points is projected in the low-dimensional plot. The iterative optimization procedure is used to minimize the Kullback-Leibler divergence between a  $t$ -distribution in the low-dimensional space and a Gaussian distribution in the high-dimensional space [18].

The results of t-SNE analysis for the common crossing statistics are shown in Fig. 4. All the observations are plotted as points with 3 different colors according to the lifetime. Remark that there are no clear groups belonging to the same lifetime. Some low prevailing concentration of red points and blue points is visible, but the zones are relatively small and cannot be exactly separated from overall statistics. The green points that correspond to the middle lifecycle are almost uniformly distributed.

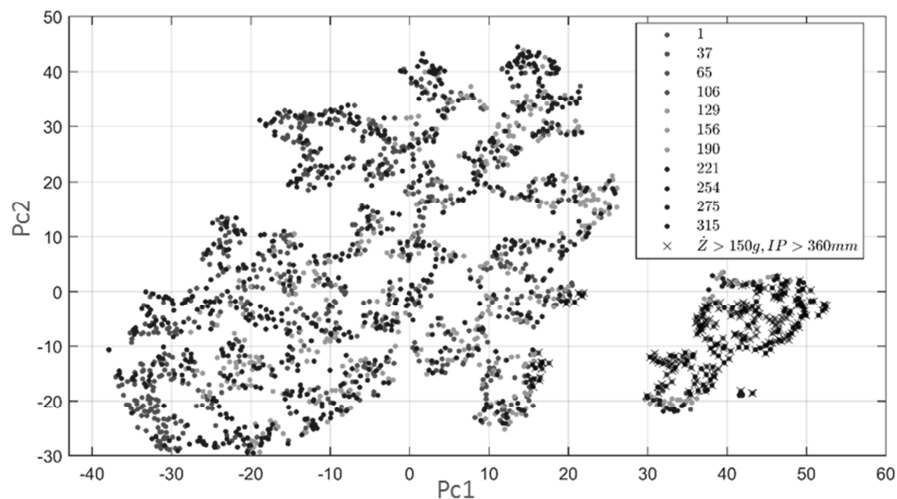


Fig. 4. The results of t-SNE analysis (Pc1-first principal component, Pc2-second principal component)

Two main clusters can be clear observed: the smaller one contains about 10% observations. The perplexity parameter is used 50 points. The right cluster is almost completely assigned to the observations with the vertical accelerations more than 150 g and impact points with coordinates more than 360 mm. The points are marked with cross.

There are some points that are not marked in the right cluster and also some few crossed points in the big cluster. The outlier points could be explained with the fact that the longitudinal and horizontal accelerations are not taken into account. For simplicity only the vertical acceleration and impact position parameters are taken into account.

The data points selected in t-SNE analysis are used to study the correlation to lifetime. Fig. 5 shows the accelerations and their mean values over the lifecycle for observations with vertical acceleration more than 150 g and impacts zone further than 360 mm. The apparent increase of vertical and horizontal accelerations until the damage can be observed. The vertical acceleration dataset is fitted with cubic relation. The cubic fit does not show the best results for all lifecycle comparing to the first or second order polynomials. The reason of the choice is the lowest polynomial order that can describe the inherent, typical technical system deterioration process. The fit shows the monotonous increase of vertical accelerations until about 200 days. After that, the growth stops and even falls down.

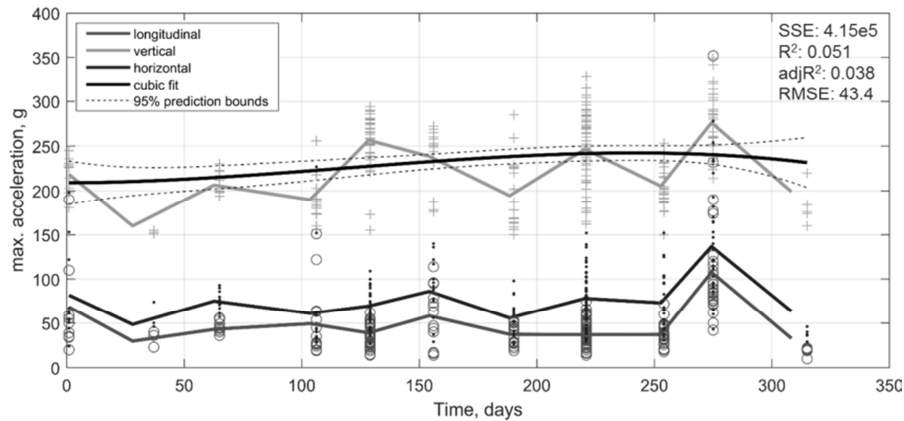


Fig. 5. Maximal impact accelerations and their mean values over the switch lifecycle for observations with vertical acceleration more than 150 g and impacts zone further than 360 mm

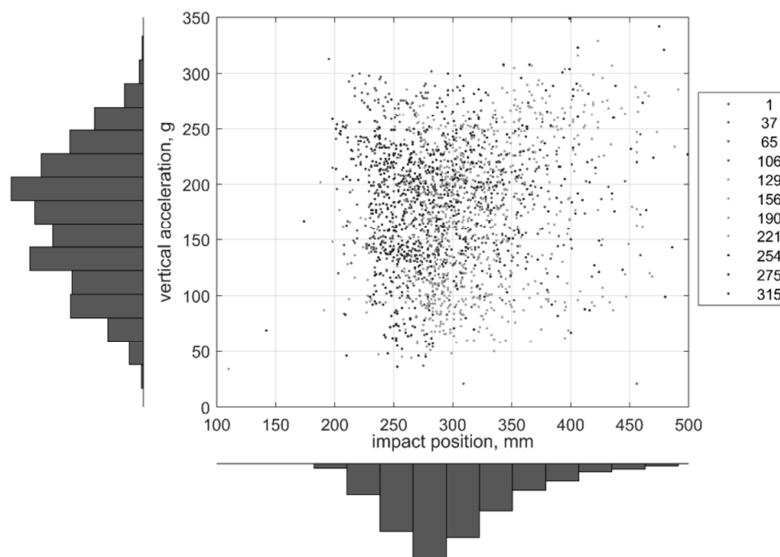
Nevertheless the evident improvement of the relation, comparing this to the initial data (Fig. 2), the reliability of the prediction is relatively low. The systematic increase of predicted acceleration over the lifetime is about the same as the uncertainty range with 95% function prediction reliability. The main reason is a low number of observations. Additionally, the relative high variation between some measurements could be caused by the short time of measurements (4-6 trains) that do not take into account all train types as well as many other factors.

Remark that the maximal values of the accelerations show much higher increase than their mean values. That means that the prediction could be further improved, if more observations could be provided. Interesting is the position of impacts points found from t-SNE analysis (more than 360 mm) corresponds to those with damage appeared near the lifetime end (Fig. 1).

#### 4. Mechanical approach to recover the relation to lifetime

The analysis has shown that the strong statistical correlation can be found in the relative small subgroup of observations. This subgroup differs from the overall statistic with the upper vertical acceleration range and the small impact zone, that near coincide with the damage zone. However, the relation is found, the statistical approach cannot explain the reason of damage appearance. For that reason, the simple mechanical approach is used to recover the relation to the lifetime of the common crossing. The approach is based on the well-known relations between loadings and cycles in fracture mechanics, surface curvature influence from contact mechanics, wear influences.

The distributions of impact positions and accelerations are studied at first. *Fig. 6* shows a scatter plot of observations of the vertical acceleration and impacts position together with their distribution histograms. The scatter plot depicts the measurement time during the lifecycle.



*Fig. 6.* The vertical acceleration and impact position distribution

Visual estimation of the diagram shows no evident prevailing relation of measurement time to impact position and vertical acceleration. Except of some small zones with big loading ( $>270$  g) and the beginning of impact zone ( $< 250$  mm) where all measurements correspond to early lifetime. The impact zone ( $> 300$  mm) in the same loading range differs in contrast to the first one in almost absent early lifetime observations.

The impact distribution demonstrates the maximal number of impacts in range 220-320 mm. The vertical acceleration distribution is broader with the most number of impacts in range 100-250 g.

Fig. 7 shows the detail analysis of impact position distribution divided in different acceleration classes. The 3 lower histograms for loading classes up to 300 g have similar form as the upper histogram for the overall statistics. The histogram for loading class more than 300 g is quite different with maximal number of loadings in zone 350-425 mm. The zone corresponds to the inspected damage (Fig. 1) at the end of frog lifecycle.

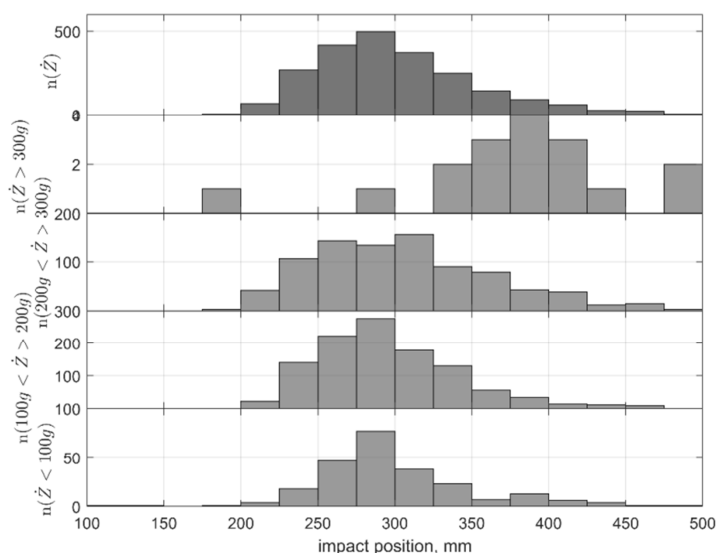


Fig. 7. Impact position distribution for loading classes

From the analysis it is evident that the damages are not caused with major loading set but with the small number of very big loadings. The damage zone 350-450 mm is subjected to 10 measured impacts with loading more than 300 g, 160 impacts of loading class 200-300 g and 120 impacts of lower loading. To take into account the influence of all loading classes the equivalent loading impacts are calculated with the relation:

$$N = N_G \left( \frac{\sigma_{-1}}{\sigma_a} \right)^m, \quad (1)$$

where  $\sigma_{-1}$  is the base fatigue stress limit corresponding to the  $N_G$  loading cycles;  $m$  is the material dependent fatigue parameter;  $\sigma_a > \sigma_{-1}$  is the stress for calculation of  $N$  loading cycles.

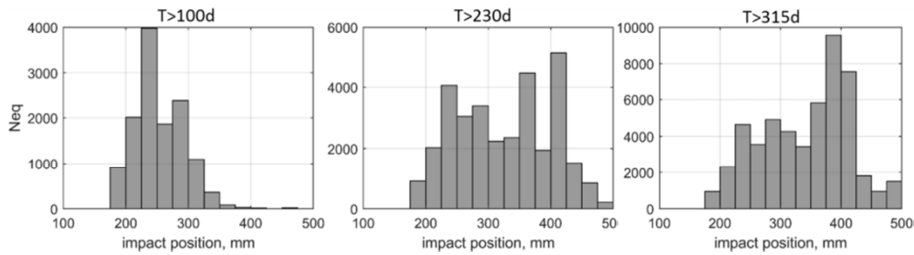
According to the relation (1), under the simplified assumption of dynamic, contact and wear influences, the following equation of equivalent impacts number  $N_{eq}$  is derived [19]:



$$N_{eq} = N_{ref} \left( \frac{\sqrt{P_{st} + m_{in} \cdot \ddot{z}_{ref}}}{f_1(x) \cdot f_2(x) \cdot f_3(x) \cdot \sqrt{P_{st} + m_{in} \cdot \ddot{z}}} \right)^m, \quad (2)$$

where  $N_{ref}$  is the reference maximal load cycles for ordinary track;  $\ddot{z}_{ref}$  is the reference acceleration that corresponds to  $N_{ref}$ ;  $m_{in}$  is the inertial mass;  $P_{st}$  is the statistical wheel loading;  $\ddot{z}$  is the impact acceleration, to which  $N_{eq}$  number of impacts is calculated;  $f_1(x)$ ,  $f_2(x)$ ,  $f_3(x)$  are the influence parameters of contact and wear and tension depending on longitudinal coordinate of impact.

The results of equivalent impact calculation are presented in *Fig. 8* in form of accumulated impact distribution up to some time. The diagram shows the movement of maximal often loaded rolling surface from 220 mm at the beginning to 400 mm at the end of lifecycle.



*Fig. 8.* Equivalent impact position distribution (left: up to 100 days, middle: up to 230 days, right: overall lifecycle)

The maximal value of equivalent impacts per 25 mm bar depending on the common crossing lifetime is depicted in *Fig. 9* together with power and exponential fits. The diagram shows the initial step of  $N_{eq}$  parameter due to high loading impacts at the lifetime beginning. After that the parameter stays almost the same until 220 days of operation. After that the parameter grows quickly.

The stepwise growth of equivalent impact curve is explained with the fragmental and short experiment measurement and insufficient amount of trains a day. The real relation with more statistic information should be more fluent as shown with fitted exponential and power curves.

The depicted *Fig. 5* cubic fit of vertical acceleration can be significantly improved with taking into account the mechanical consideration. According to it, each data point becomes different weight of influence, depending on the equivalent impact number. *Fig. 10* shows the cubic fit with the new approach. The data points with high acceleration have usually a high weight and thus pull the regression line. The influence of data points with low accelerations is reduced. The approach enables us to increase

significantly the coefficient of determination of the prediction, though the deviations predicted are the same as for the initial data.

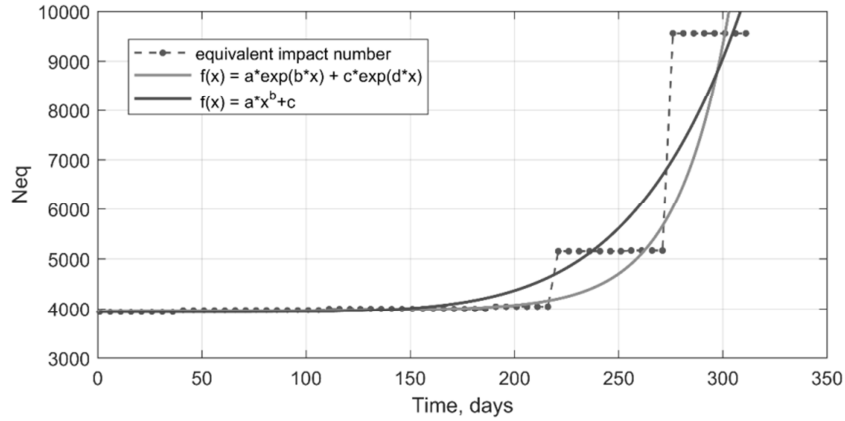


Fig. 9. The change of the maximal equivalent impact density over the common crossing lifecycle

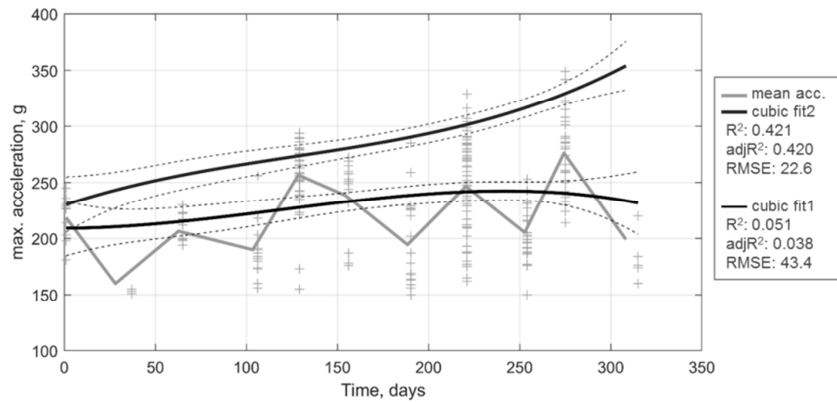


Fig. 10. Improvement of the statistical prediction with the mechanical approach

## 5. Discussion

The presented study shows the comparing results of statistical and mechanical approaches for recovering the relation of measured dynamical parameters to the lifetime of the common crossing. Both statistical and mechanical approaches are able to find clear relations. Moreover, the mechanical approach, distinct to statistical, can determine the accelerated changes of parameters, and thus can be used for prediction of rolling surface damages. The significant changes could be observed up to 50-100 days before the first damages appear. Statistical approach could produce the better result if more data were present.

Nevertheless, the mechanical approach has a significant drawback - it is based at accumulated statistic data. This approach could be used with a stationary monitoring system that collects data continuously. The portable inspection system has the measurement information from relatively short span of lifetime without the information before. The fusion of statistical and mechanical approaches is a most promising way to solve the problem.

## 6. Conclusion and subsequent studies

The early prediction of turnout faults is crucial for the reliability of railway transportation networks. The inertial measurements are one of the perspective techniques for creating cost efficient predictive and prescriptive maintenance system. This paper shows how the use of modern statistical tool together with simple mechanical ones helps to recover the statistical relationship between the inertial measurement and the common crossing lifetime. The relationship could be used for common crossing early fault prediction. The analysis has shown that the faults appear not accidentally and are caused with some small set of critical wheel impacts. The reason of the impact formation is still unclear and is the aim of the subsequent studies as well as the improvement steps of the measurement system:

1. Taking into account the information not only about the maximal acceleration and its position, but also other 2 components together with spectral parameters. Derivation of one parameter that represents common crossing state with principal component analysis and supervised learning methods;
2. Exclusion of systematical measurement errors, such as error due to different wave decay value for different distance from impact to sensor;
3. Estimation of the necessary configuration for the measurement system, the acceleration sensors position, number of sensors;
4. Gathering the additional information from new data sources, and more useful utilizing the existing data sources, e.g. estimation of analogous information from proximity sensors to determine the cross position of wheel flange in a flange way gap.

## Acknowledgements

The Authors would like to express their thanks to Germany Railway Company (DB Systemtechnik GmbH) and WITT Elektronik GmbH for their data and financial supports.

## References

- [1] Fendrich L., Fengler W. *Field manual railway infrastructure*, (in German), Springer-Verlag 2013.
- [2] Gerber, U., Zoll A., Fengler W. Wear and rolling contact fatigue on common crossings of railway turnouts (in German), *Eisenbahntechnische Rundschau*, Vol. 1-2, 2015, pp. 36–41.

- [3] Kovalchuk V., Sysyn M., Hnativ Y., Bal O., Parneta B., Pentsak A. Development of a promising system for diagnosing the frogs of railroad switches using the transverse profile measurement method, *Eastern European Journal of Enterprise Technologies*, Vol. 2, No. 1, 2018, pp. 33–42.
- [4] Gerber U., Zoll A., Fengler W. Vehicle-based assessment of wear on common crossings, (in German) *El-Eisenbahningenieur*, Vol. 5, 2013, pp. 26–30.
- [5] Zoll A., Gerber U., Fengler W. The measuring system ESAH-M, (in German), *Eisenbahningenieur Kalender*, 2016, pp. 49–62.
- [6] Böhm T., Weiss N. Turnout analytics - smart sensors and artificial intelligence for the all-round healthy turnout, (in German), *Eisenbahntechnische Rundschau*, No. 5, 2017, pp. 42–45.
- [7] Liu X, Markiene V., Shevtsov I. Dollevoet R. Experiment study of key parameters in turnout crossing degradation process, *Proceedings of the 10<sup>th</sup> International Conference on Contact Mechanics*, Colorado, USA, 1-3 September 2015, pp. 122–127.
- [8] Sysyn M. P., Kovalchuk V. V., Jiang D. Performance study of the inertial monitoring method for railway turnouts, *International Journal of Rail Transportation*, Vol. 4, No. 4, 2018.
- [9] Sein S., Matos J. C., Idnurm J. Statistical analysis of reinforced concrete bridges in Estonia, *Baltic Journal of Road and Bridge Engineering*, Vol. 12, No. 4, 2017, pp. 225–233.
- [10] Storz T., Ercsey Z., Várady G. Histogram based segmentation of shadowed leaf images, *Pollack Periodica*, Vol. 13, No. 1, 2018, pp. 22–32.
- [11] Hůlková M., Pavelka K., Matoušková E. Automatic classification of point clouds for highway documentation, *Acta Polytechnica*, Vol. 58, No. 3, 2018, pp. 165–170.
- [12] Bolkeny I., Fuvesi V. Ai based predictive detection system, *Pollack Periodica*, Vol. 13, No. 2, 2018, pp. 137–146.
- [13] Rybkin V., Gerber U., Nabochenko O., Sysyn M., Determination of the ballast layer degree compaction with dynamic and kinematic analysis of the acoustic waves impacts, *Sborník Pednáek Zelezniní Dopravní Cesta VO a SP Stavební*, 2010, pp. 123–130.
- [14] Lederman G., Chen S., Garrett J., Kovačević J., Noh H. Y., Bielak J. Track-monitoring from the dynamic response of an operational train, *Mechanical Systems and Signal Processing*, Vol. 87, Part A, 2017, pp. 1–16.
- [15] Jamshidi A., Hajizadeh S., Su Z., Naeimi M., Núñez A., Dollevoet R., Schutter B., Li Z. A decision support approach for condition-based maintenance of rails based on big data analysis, *Transportation Research, Part C, Engineering Technologies*, Vol. 95, 2018, pp. 185–206.
- [16] Fengler W. On the way to the Railway 4.0 - Predictive maintenance, (in German) *Eisenbahntechnische Rundschau*, Vol. 5, No. 3, 2017, pp. 3–5.
- [17] Wattenberg M., Viégas F., Johnson I. How to use t-SNE effectively, *Distill*, Vol. 1, No. 10, 2016, Paper e2.
- [18] Van Der Maaten L., Hinton, G. Visualizing data using t-SNE, *Journal of Machine Learning Research*, Vol. 9, 2008, pp. 2579–2605.
- [19] Glyusberg B. Methods for optimizing the basic elements of turnouts and their application to common crossing of mass structures, (in Russian) *Doctoral Thesis*, Railway Research Institute, Moscow, 1989.