ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20240420001479
Original scientific paper

Innovative Approaches in Railway Management: Leveraging Big Data and Artificial Intelligence for Predictive Maintenance of Track Geometry

Richárd NAGY, Ferenc HORVÁT, Szabolcs FISCHER*

Abstract: This paper introduces and describes a method for extracting, processing, and analyzing large amounts of track geometrical data. It allows for a more accurate description of the orbital deterioration correlations than currently applied procedures, and it seems to be more valuable and efficient in practice. The initial data were the track geometry measurement and classification data for the whole national network provided by the Hungarian State Railways, i.e., the MÁV PLC. The MÁV provided data for the whole Hungarian railway network for 27 half-years, measured and recorded by the FMK-004 type special diesel locomotive (i.e., track geometry measuring car). The paper discusses the development of a procedure to automatically compute important condition ratings from the available data set of millions of units according to the algorithms created for railway industry colleagues, thus helping the maintenance and renewal decision-making process. Functions have been developed to classify the track geometry condition of a given railway line, to predict how long the service level can be maintained without intervention (i.e., e.g., lining, leveling, and tamping with a mechanized maintenance train), to determine the time of the necessary maintenance intervention, the time of the upgrade (rehabilitation or modernization), and to develop a track geometry prediction procedure that makes full use of the mathematical and computational possibilities of the present day.

Keywords: artificial neural networks (ANN); exponential predictive model; maintenance and renewal decision-making; mathematical and computational modeling; predictive maintenance; railway track geometry; track condition rating

1 INTRODUCTION

Railway transport is one of the most environmentally friendly modes of land transportation and the most energy-efficient in terms of the weight of goods transported [1]. Its "competitor", mainly because of journey/transportation times, is road transport [2, 3]. Unfortunately, for shorter distances within and between countries and continents, and for small volumes of specialized goods transport, railway transport is often unable to compare with road transport. This means that door-to-door transport by railway is more challenging. For significantly longer transport distances (i.e., e.g., more than 1000 km), air or water/sea transport is, in most cases, the appropriate choice. Unfortunately, these cannot easily compete with land transport due to several factors (which are not detailed in the current article due to the limited space).

The general design (superstructure set-up) of railway tracks can be divided into two main categories/types [4]: the traditional ballasted design with crushed stone ballast [5] and the ballastless solutions (e.g., slab tracks) [6]. Within these, track systems can be of either fishplated rail joints [7] or so-called CWR design (continuously welded rail) [7-9]. Regarding lifetime cost, the preferred solution is always the CWR one. For larger available resources, ballastless designs are the preferred option (construction costs can be significantly higher than for a ballasted superstructure, up to 2-3 times higher) [4]. In the case of ballastless tracks, the substituting of the elastic behavior of the track structure is needed by special elements, e.g., elastomers, damping solutions, etc. [10, 11]. Using modern, up-to-date solutions and elements in the railway tracks, the evolved mechanical stresses and strains can be reduced both in the superstructure and the substructure. Due to the physical phenomena of the deterioration (geometrical and structural) of the engineering structures (the authors involve the railway permanent ways and their tracks and the related elements, e.g., turnouts and crossings [12-15] in this category), maintenance is a vital procedure [16]. Parallel the straight sections, the relevancy of curved ones must be mentioned [17, 18]. If someone analyzes the behavior of the railway track and its deterioration, the stiffness [18] and the dynamic modeling [19] are also relevant factors.

The allocation of maintenance resources to the railway sector has always been a process that required a great deal of preparation and engineering judgment, and this is still the case today. Limited budgets always force the maintainer to prioritize and schedule network rail sections according to some reasoning system. These argumentation systems are typically based on engineering; even in this case, it is difficult to determine which model or intervention system is the most appropriate. Typically, the railway maintenance process can be carried out in three ways, such as predictive, preventive, and corrective maintenance [17, 18]. Corrective maintenance is the least efficient and expedient method of using material resources [20]. Preventive maintenance is the most typical and widespread of the three types of maintenance, but its main disadvantage is that it requires systemic intervention and touches parameters of the track condition that may not necessarily require intervention. The third and perhaps the most forward-looking is the predictive form of maintenance, which predicts the extent of some defect in a given section of the track as a function of time or train passage [20].

The timing of interventions should consider all factors that may have some cost implications [21]. Therefore, among other things, the track condition of a given railway section, the disruption to traffic flow caused by maintenance works, and the delays caused by this. The type and necessity of intervention will be significantly influenced by whether a condition problem occurs at a level that requires immediate action.

The authors of this study are aware that establishing a correct maintenance system, a sufficient quantity and quality of data chains, models to analyze them, knowledge of the elements of the respective intervention system, and finally, a decision support system based on their parameters are necessary. Among the factors listed here, this paper aims to present a set of procedures for collecting, filtering, and systematizing data, then compare path degradation models with each other and with reality and illustrate the consequences of decisions made using different models.

2 LITERATURE REVIEW

Generally speaking, the geometric degradation (or, in other words, deterioration) of railway tracks impacts track quality, running, comfort, and safety, as well as reliability, availability, allowable train speed, and overall railway performance [22]. A wide variety of methods and systems have, therefore, been developed worldwide to determine the timing of interventions. A predictive model creates a digital twin with deep learning systems, intending to optimize the existing problem for intervention response and cost [20]. Among the studies, one tried to optimize both objective functions: the total cost of planned maintenance and rehabilitation measures and the total number of trains caused by speed restrictions [23]. Swedish scholars also believe [22] that railway track maintenance should integrate proactive solutions to achieve the desired goals.

Following the example of Swedish researchers [22], it would be necessary to create a life cycle cost model that can compare the degradation process of a track in time with the intervention costs for the desired level of service and then analyze the systematic costs of interventions or non-interventions. In the Netherlands, there are two intervention-level groups for repairing defects: a group of lower and a group of higher-level interventions [21]. As other literature points out, maintenance costs can be reduced by implementing a long-term maintenance strategy and performing preventive maintenance activities. Furthermore, decision support systems in rail infrastructure can help reduce maintenance and renewal costs while mitigating the effects of train delays and enabling more informed decision-making on maintenance and renewal actions [23]. Studies and academic research show that maintenance costs can be reduced by implementing a long-term maintenance strategy and preventive maintenance activities [22]. It was the work of Pál Vaszary [24, 25] and his colleagues [26] that defined the theoretical exponential predictive model, which was later used and simplified into a linear model by Veit et al. [27, 28]. At present, the application and sub-analysis of these predictive models are in the hands of the railway section engineers, who determine what kind of degradation can be assumed for a given railway line in the coming period. Therefore, the authors of the current paper intended to create an analytical procedure highlighting the nature of these processes and assisting the experts in their decision-making.

The evolution (over time) of the geometry metrics and qualifiers without intervention typically shows a deteriorating character. Once they reach a limit value, various guidelines, instructions, and regulations will provide answers as to the timeliness and manner of intervention activities. The current regulation in Hungary is based on the European standard EN 13848-5, which defines three main limits [29]:

- *AL* - Alert Limit: the value beyond which the track geometry condition must be analyzed and taken into account in regularly scheduled maintenance operations.

- *IL* - Intervention limit: the value above which corrective maintenance must be carried out before reaching the immediate action limit.

- *IAL* - Immediate action limit: the value beyond which a speed limit or immediate track geometry correction is required.

It is always considered an essential question regarding what resources are available to complete the engineering job that has been assigned. Infrastructure maintenance and renewal costs on Europe's conventional main lines can be as high as \$150 000 per kilometer per year, and two-thirds of these costs are related to the track system [30]. Several studies have shown that existing maintenance costs can be expected to be reduced by 20-30% by developing a risk-based approach to infrastructure maintenance [31]. Track substructure maintenance costs account for 55% of the total maintenance costs for high-speed lines [32]. Several research groups [33] have also investigated the value of the relationship between planned and unplanned maintenance and its effectiveness. In their study [34], the authors conclude that a maintenance regime built with optimal maintenance interventions is associated with lower costs than regular preventive maintenance. They conclude that railway sections with the highest deterioration rate should be monitored to determine the optimal intervention time for track geometry correction. There are also studies [35] that do not test their theoretical models under real conditions but test the optimal intervention window on a fictitious railway infrastructure crossing several national borders. In [36], an Integrated Maintenance Index (IMI) was introduced in which the track geometry data is weighted by the number of defects, the magnitude of defects, and the impact of defects on safety. This is used to produce a track maintenance plan. In [37], a strategic approach to the design of existing maintenance systems is evaluated using various analytical and descriptive parameter tables, paying attention to the interdependencies between the type of intervention and the sensitivity of the section.

There are many ways of describing the deterioration of track geometry status in the literature. Some deterioration models treat the deterioration phenomena of a recently upgraded track section with a lognormal coefficient [30]. The literature divides statistical processing into four groups: open sections, main tracks of stations, secondary tracks of stations, and tracks on bridges. Their study analyzed only open sections, separating straight and curved sections. Several pieces of literature [30] highlight that there is no or very sparse data on the condition of existing railway lines in a given country, making the modeling process difficult for researchers and engineers. There are stochastic mathematical modeling techniques for predicting the correction of track geometry as preventive maintenance, where the change of the standard deviation of level is examined to determine the appropriate intervention time [38]. Indian researchers [39] also work on a system to help in decision Support, using AI and other advanced technological analysis software. In [40], a colored Petri net method is used to predict the degradation state of a given section of the track and its degradation rate by Monte Carlo simulation to estimate the type of interventions and optimize the costs. In [41], the problem of what logistical and tactical elements of the maintenance flow should be included in the decision chain so that maintenance capacity is available is addressed. In [42], they have demonstrated the effectiveness of their model on a numerical model based on sensitivity analysis. Using a model with two heuristic algorithms, they were able to quantify that maintenance and renovation costs can be reduced by up to 14%. Their model [43] proposes three output decisions: maintenance, refurbishment, and doing nothing. In [42], they present a system using a Monte-Carlo simulation technique with 14 components, with two levels of decision-making, one level acting on system-level interventions and the other recommending interventions that affect the components. Basically, it can be stated that many studies are being done on the rail maintenance and repair system, trying to optimize every element of the rail system from a cost-efficiency point of view. Such as [44] on optimizing the maintenance interventions for railway bridges in which the annual running costs are optimized, the data-driven decision Support model for railway signaling system failures that integrates different parameters of maintenance data, and the maintenance performance considering different parameters. In [45], two different scenarios are used to illustrate the high-speed line intervention model. One scenario is the traditional intervention model, while the other is a scenario with intervention options built with a narrower toolbox and narrower options. The study tried to draw attention to the need to take into account the consequences of the interactions between the operational system and the preventive maintenance requirements when making decisions. It can also be seen in the study reported in [46] that the impacts on the track are not characterized as a product of throughput and load or as a function of any number physically loading the track but rather as time intervals, typically in four increments. They created mathematical models that led to three results: do nothing, maintenance, and replacement. The inputs to their models are number of components, length of design phase, number of intervals, condition of the given component, repair factor of the given component, unexpected failure cost of the given component, maintenance cost of the given component, replacement cost of the given component, fixed cost of the system, reliability of the components in the system, maintenance budget available for the order. In the same study [46], two non-linear models that seek the minimum cost with maximum reliability were constructed. However, it also concludes that their model is limited in terms of the number of components that can be considered. As reported in [47], many studies consider the stochasticity of the faults measured in the railway track independently of the maintenance schedule. Detailed data are not necessarily needed if exact data are available. In the literature search, several studies were found where, similar to the present study, regression-based deterioration models were used to describe the state of the track over time, supplemented with artificial neural networks [48]. Similarly [49] used artificial neural networks (ANNs) to investigate the deterioration of the geometric condition of the railway track. Several others emphasized other data or parameters: while [46] focuses on the geometry and structural elements of the railway track, [48] focuses on the traffic passing through the railway track section. In [50], linear regression is used for the track deterioration model, while ordinal logistic regression is used to estimate the occurrence of local defects. The study in [51] showed that an Adaptive Network-based Fuzzy Inference System can handle the prediction of the gauge values of straight lines and curves with an approximation of 0.6 and 0.78 R2, respectively, which has been validated by real measurements. The ANN built in [52] had 12 input variables (ballast age, maintenance history, rail type, sleeper type, sleeper age, level crossing, bridge, track speed class, average annual frequency of trains, degradation value after tamping, curvature) while experimenting with the number of hidden neurons where it started with 3 neurons and terminated with 20 neurons, observing the R^2 value in both learning and testing phase. Finally, they concluded that 10 hidden neurons were the ones that yielded the appropriate accuracy, namely R^2=0.85, which is sufficient to state that the model can predict the decay rate correctly.

3 MATERIALS AND METHODS

3.1 The Source of Data

Assessing the condition of railway tracks is in the fundamental interest of all maintainers. Swedish scholars measure more than 30 railway parameters, pointing out the need to improve the accuracy of the GNSS identification [22]. They also face the problem of a lack of a long-term maintenance strategy and, therefore, use only arbitrary judgment and historical data to plan maintenance activities [22]. Overseas, the TQI (track quality index) is typically used to represent track conditions. TQI is traditionally derived from the measured values of the rail track geometry per foot [53]. In Hungary, as in other countries, the rail track geometry is characterized by its three most characteristic properties.

Figure 1 FMK-004 Track measuring car

As described by MÁV KFV Ltd., the FMK-004 measuring car performs track geometry measurements in several countries of Central Europe in addition to Hungary. The measuring car is diesel-powered and self-propelled. The track measurement results are available in printed and electronic form in real-time on the measuring car and for the customer immediately after the measurement is completed. The geometric measurement results are viewed, evaluated, and analyzed by the office system.

The track geometry measuring system is contact-based, i.e., the measuring system calculates most of the measurement results from the movements of the wheels rolling on the rails (a total of 18 wheels, larger or smaller, on nine axles). The measuring system can also provide alignment and longitudinal level diagrams on distortion-free and original chord bases. The cross-level is determined using a gyroscope. The twist can be calculated for different bases.

The track geometric parameters that can be provided are:

- gauge,
- cross-level,
- twist (on five different basis),
- longitudinal level (on original chord and D1 or D2 wavelength range),
- alignment (on the original chord and D1 or D2 wavelength range (see Fig. 2).

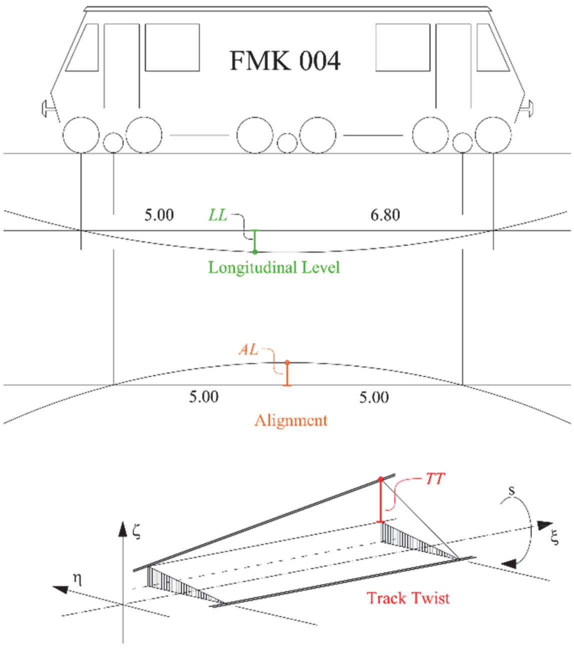

Figure 2 Explanation of the three types of track geometric characteristics measured by the FMK-004 measuring car

The relative geometry of the rail track is described by three measuring numbers:

- alignment as *AL*,
- longitudinal level as *LL*,
- track twist as *TT*.

The Qualification number is the arithmetic mean of *AL*, *LL*, *TT* on 2.5 m basis and *TT* on 6.0 m basis values, which describes the general geometric condition of the track. It is calculated to a given qualifying length (earlier 500 m, currently 200 m) [29]. Other outputs that require processing are measurement graphs (different measuring limits with objects), a list of local defects (taking into account different measuring limits), and general track condition judgment (evaluation) by measuring and qualification numbers (at any length).

It is worth mentioning that in Hungary, two measuring wagons can make geometric measurements of railway lines. They can provide the exact data for 500 m by 200 m and 25 cm section lengths. In our research, very long data series availability was more important than the segmentation of the qualifying section, so we stuck to the 500 m data. However, in further research, we can analyse other data using the process developed in this paper.

3.2 The Procedure of Data Analysis

It was already probable during the literature search, and the amount of data the authors received for the studies ensured that only a computer procedure could process data of the order of millions of units. Since the data was obtained in Excel and the industry uses this program, it was obvious to use the built-in Visual Basic for Applications, an object-oriented, event-driven, structured programming language developed by Microsoft with an integrated development environment. MATLAB was used, and a specialized program system was developed to perform numerical calculations and use a programming language. The programming system developed by MathWorks can perform matrix calculations, represent functions and data, implement algorithms, and create user interfaces. Although the software is purely numerical, it can display mathematical expressions graphically by adding the MuPAD package. (Matlab, 2021) Furthermore, it was helpful in the artificial neural network process, as it is well-programmable to handle and process databases prepared in Excel. Finally, a significant part of the statistical analyses was performed using a software tool called RStudio.

A computer procedure was made to automatically process the measurement and classification data provided by the FMK-004 track geometry measuring car, determine the change in geometric condition based on time series values, provide a forecast of the change with its accuracy, determine the place and time of the necessary maintenance intervention, help the engineer in track condition analysis and the preparation of more complex professional decisions. Due to the length of the program's original flowchart, only the extract is available here in Figure 3. Still, a more precise and detailed description is also available [54], illustrating and explaining the flow data's context in more detail.

The program:

- collects a specified number of consecutive half-year data series without any additional work,
- fits linear, exponential, logarithmic, and power-based regression descriptors to the collected data series,
- collects the function parameter and deterministic coefficient of each regression equation, classifies them, and writes an equation of each of the four types for the line in question, which is plotted on a graph,
- plots the parameters of the equations separately on histograms and, according to the distribution, further ranges can be selected for which the parameters of the given type of function are repeatedly calculated,
- calculates the correlation between the desired measures and the desired rating, i.e., the magnitude of the linear relationship between two values,

- calculates the deviation of missing or outliers from the regression-fit equation of the data series,
- creates some statistical analysis of the geometrical deterioration values of three railway lines, like Kolmogorov-Smirnov, Welch two-sample t-test, Descriptive Statistics, Comparative Distribution Analysis, and Paired t-test,
- builds up seven models and creates a comparative analysis to describe the accuracy of the models.

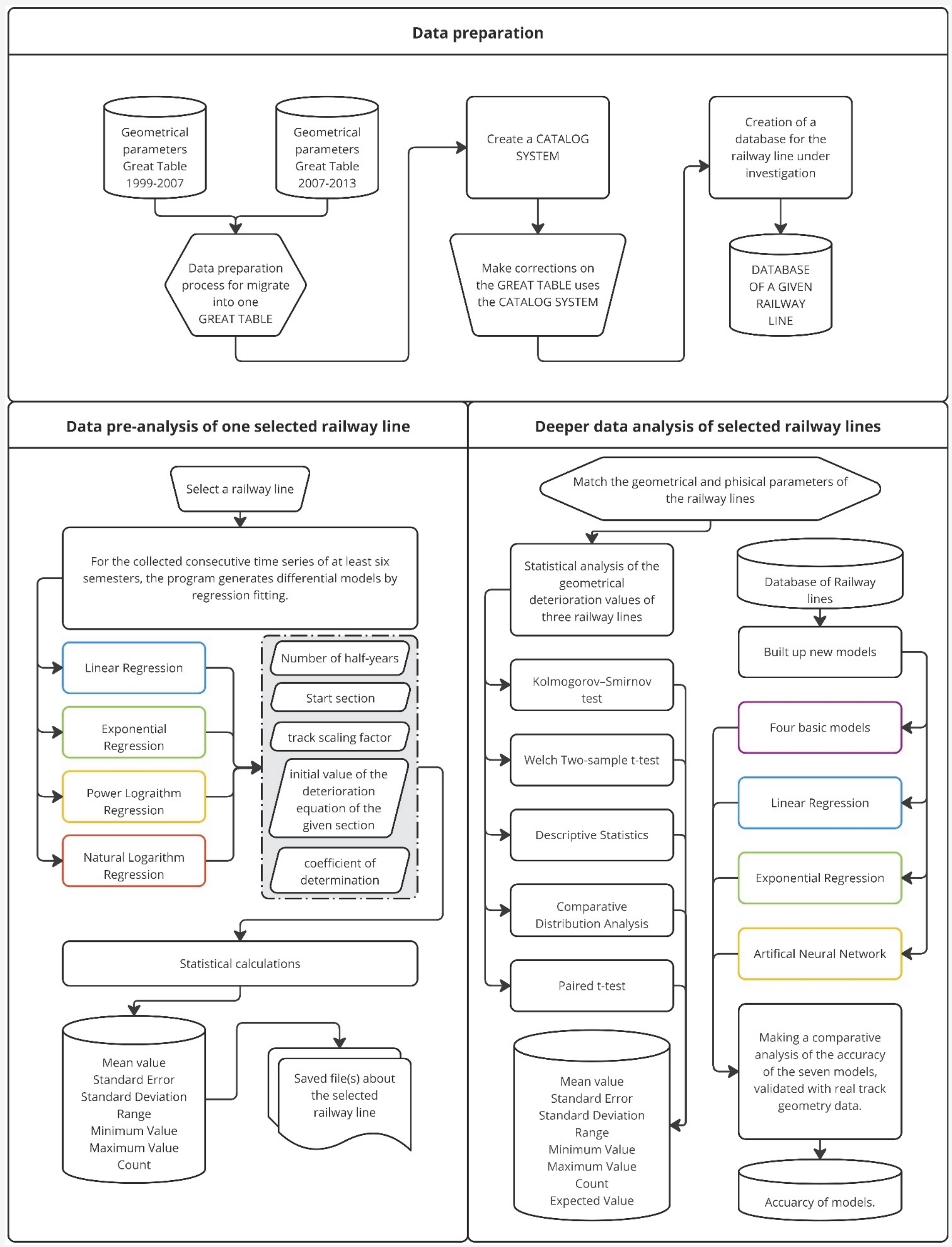

Figure 3 Flowchart of the procedure

3.3 Statistical Comparison of Data

Two railway lines with the most divergent maintenance cycles were selected, and a third line was halfway between the two extremes. Thus, line #1 (Budapest - Hegyeshalom - Rajka) was selected as the line with the highest maintenance quality, line #60 as the line (Gyékényes - Pécs) with average quality and line #140 (Cegléd - Szeged) the worst one.

The series of measuring numbers, which continuously deteriorated for at least six semesters, were collected for each 500-m-long section of all three lines. These are henceforth referred to as six semi-annual deterioration series $SSDS_i$, where index i means the individual characteristic: $i = LL$ as Longitudinal Level, $i = AL$ as Alignment, and $i = TT$ as Track Twist). The extent of deterioration in these series is represented by the difference in the measuring numbers of consecutive semesters labeled dLL_i, dAL_i, and dTT_i.

All statistical tests were performed at a significance level of 0.05 (p-value = 0.05).

3.3.1 Kolmogorov-Smirnov Test

A non-parametric test is used to compare the distribution of two samples. The null hypothesis s is as follows:

H_0: $F_X(x) = F_0(x)$, the two samples are from the same distribution.

The counterhypothesis is

H_1: $F_X(x) \neq F_0(x)$, the two samples are not from the same distribution.

In this test, the distributions of *dLL*, *dAL*, and *dTT* metrics for Highest Frequency of Rejections Statistic (HFRS) sections of lines maintained at different maintenance frequencies were compared.

Table 1 p-value of Kolmogorov-Smirnov test

Differences of measurement numbers	No. of lines		
	#1-#60	#60-#140	#140-#1
dLL	2.20E^{-16}	4.32E^{-11}	2.20E^{-16}
dAL	2.20E^{-16}	0.21681	2.20E^{-16}
dTT	2.20E^{-16}	0.000139	2.20E^{-16}

In light of the results in Tab. 1, comparing the *dLL*, *dAL*, and *dTT* values of line #1 with those of line #60 and line #140, it can be seen that there is a significant difference in the distribution of the deterioration rates of all three measures. However, between the values for *dAL* at line #60 and line #140, there is no significant difference between the two distributions.

3.3.2 Mean Values

The *dLL*, *dAL,* and *dTT* mean values of *SSDS* for the three lines were also calculated and can be seen in Tab. 2.

These pseudo-mean values show the increase in deterioration of a given line for a given characteristic in the first six half-years after the intervention. Extrapolating this to the whole line makes it easy to calculate the average degradation rate values for the same line under the traffic loadings without intervention. The average values refer to the average of the whole line and not to a specific 500-m section.

Table 2 Mean-values

Differences of measurement numbers	No. of lines		
	#1	#60	#140
dLL	6.02	13.22	20.03
dAL	3.71	7.20	7.17
dTT	3.11	6.18	6.96

3.3.3 Welch Test No. 1.

Significant differences in means were tested using Welch's version of a two-sample t-test. The results are reported in the Tab. 3 below. The null hypothesis in all cases is that the means are equal.

Table 3 p-values of Welch test no.1

Measurement numbers	No. of lines		
	#1-#60	#60-#140	#140-#1
LL	2.20E^{-16}	3.26E^{-08}	2.20E^{-16}
AL	3.15E^{-11}	0.9613	4.51E^{-14}
TT	4.42E^{-08}	0.2574	2.20E^{-16}

Statistical tests show no significant difference between the means of *dAL* and *dTT* of line #60 and line #140, while in the other cases, there is (at 0.05 significance level).

The Welch test confirmed what was already suspected from the means, that the Alignment and Track Twist measuring numbers for Line #60 and Line #140 deteriorated at a similar rate over the six months following the intervention. However, in all other cases, the rate of deterioration is different.

So far, all the values of *dLL*, *dAL*, and *dTT* associated with *SSDS* have been examined using statistical methods. In the following section, the same values broken down by post-intervention half-year were analyzed. Histograms of these disaggregated values are shown in Fig. 4, and the distribution diagram is shown in Fig. 5.

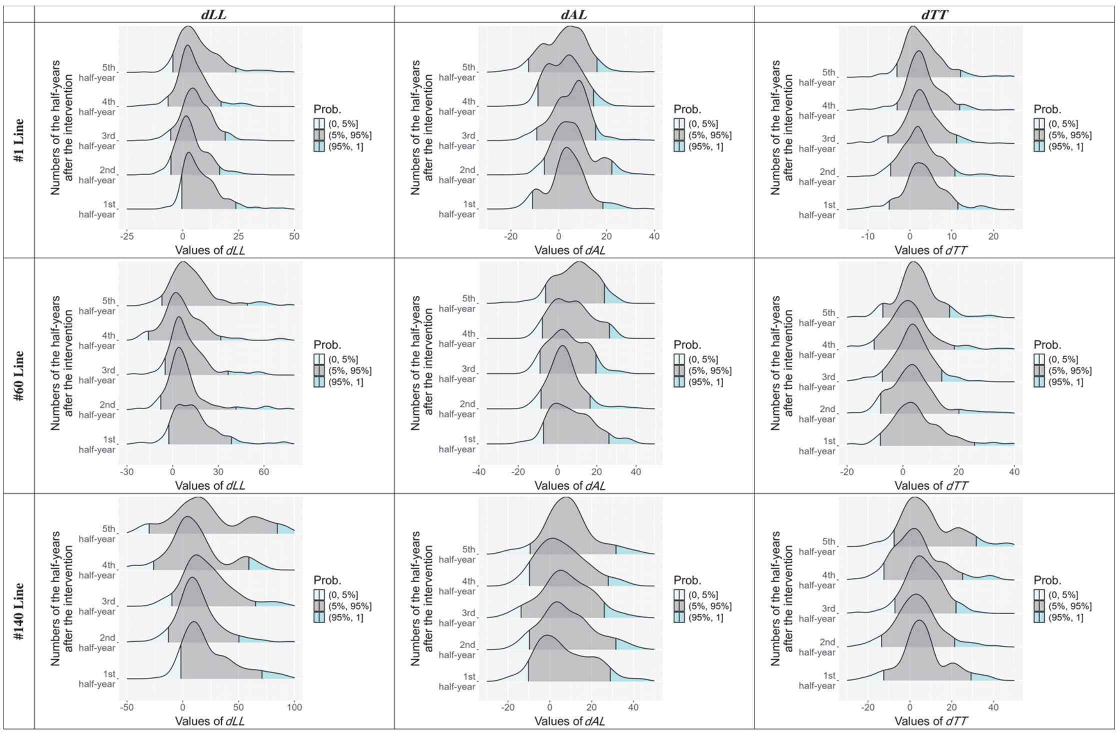

Figure 4 Change in the increment of geometric characteristics from half-year to half-year plotted on histograms

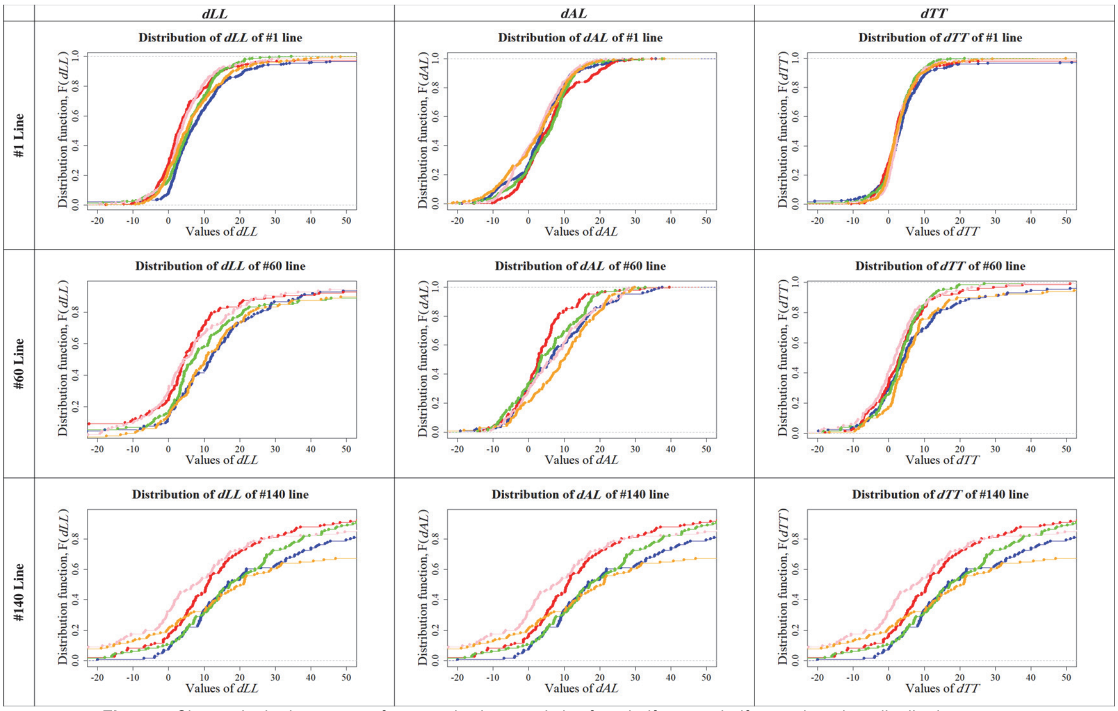

Figure 5 Change in the increment of geometric characteristics from half-year to half-year plotted on distribution curves

It can be seen that there is a significant difference between the histograms among the lines, but no significant change in the values of the same characteristics for the same line can be detected. Further statistical tests were therefore carried out. The expected values of *dLL*, *dAL*, and *dTT* for the *SSDS* in the series of the three lines are summarized in Tab. 4.

Table 4 Expected values

dLL					
No. of half-years	1-2	2-3	3-4	4-5	5-6
line #1	9.34	6.73	4.25	5.44	7.04
line #60	16.25	13.11	18.25	13.79	18.33
line #140	34.04	18.70	21.30	18.54	41.81
dAL					
No. of half-years	1-2	2-3	3-4	4-5	5-6
line #1	3.53	6.00	4.72	2.47	2.44
line #60	7.52	8.69	5.08	7.23	9.05
line #140	11.85	9.48	9.03	8.95	21.35
dTT					
No. of half-years	1-2	2-3	3-4	4-5	5-6
line #1	4.73	4.37	2.17	3.51	3.00
line #60	11.22	5.32	4.43	2.46	10.79
line #140	10.52	4.64	9.11	7.62	17.52

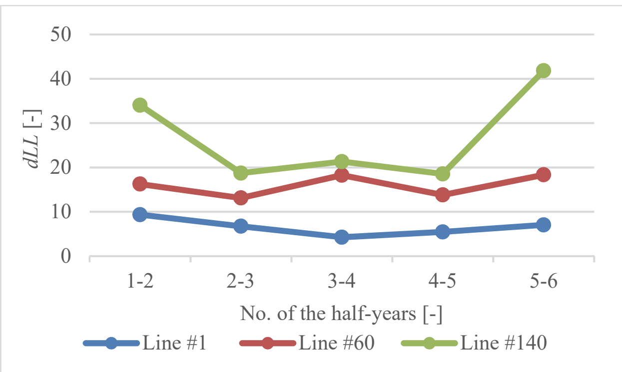

Figure 6 Expected values of *dLL*

For the three lines, it was determined that the more frequently a line is maintained, the lower the expected values of the characteristics in semesters after the intervention, and the smaller the variance of these values and the uncertainty of the prediction, which can be seen in Figs. 6 to 8. At the same time, it is also observed that longitudinal level is the characteristic that is most affected by the number of interventions.

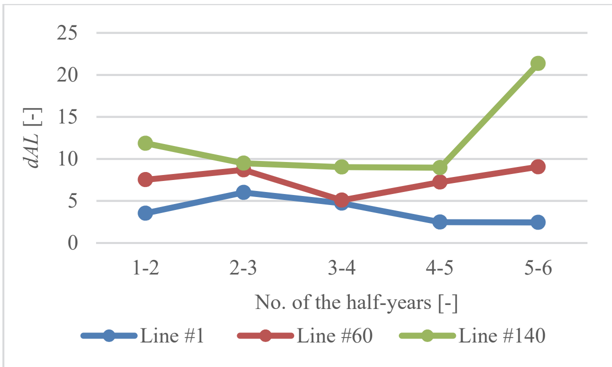

Figure 7 Expected values of *dAL*

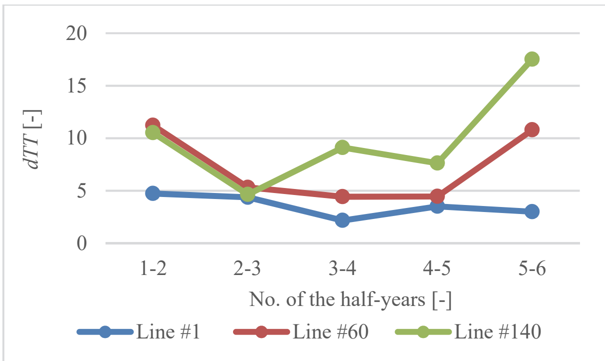

Figure 8 Expected values of *dTT*

3.3.4 Welch Test No. 2

For each of the three railway lines, a Welch version of a two-sample t-test was used to determine whether the mean of the rate of deterioration for one line and the rate of deterioration for another line in the same period had the same or different distributions between the two half-years.

Table 5 p-values of Welch test no. 2

LL				
No. of lines	$ddLL_{12\text{-}23}$	$ddLL_{23\text{-}34}$	$ddLL_{34\text{-}45}$	$ddLL_{45\text{-}56}$
#1-#60	0.2462	0.1469	0.04947	0.03416
#60-#140	0.01373	0.00312	0.6841	0.4717
#140-#1	1.6E^{-07}	0.00015	1.3E^{-08}	0.01818
AL				
No. of lines	$ddAL_{12\text{-}23}$	$ddAL_{23\text{-}34}$	$ddAL_{34\text{-}45}$	$ddAL_{45\text{-}56}$
#1-#60	0.00046	0.01024	0.7063	4.3E^{-06}
#60-#140	0.1273	0.00245	0.03631	0.687
#140-#1	0.00228	0.05238	0.01325	0.1227
TT				
No. of lines	$ddTT_{12\text{-}23}$	$ddTT_{23\text{-}34}$	$ddTT_{34\text{-}45}$	$ddTT_{45\text{-}56}$
#1-#60	0.0946	0.6207	0.1331	0.3204
#60-#140	0.8689	0.7751	0.02877	0.08928
#140-#1	0.01394	0.8775	4.7E^{-05}	0.1615

After comparison, it can be seen that *LL* is the measuring number that has the most different distribution, while the *TT* increment is the one that has the most similar distribution regardless of the number of interventions.

3.3.5 Paired t-Test No. 1

A paired t-test was used to examine whether or not the increment of a given measure on a given railway line is the same in consecutive half-years at a significance value of 0.05. This test examines the impact of the intervention on measuring numbers (*LL*, *AL*, *TT*) in the six months following the intervention, comparing the increase in measuring numbers between two semesters. That is, whether or not the phenomenon of a higher rate of deterioration in the post-intervention period is present for these three lines. Alternatively, if not, the half-year when a significant difference in the distribution of deterioration starts to emerge.

Table 6 p-values of Paired t-test no. 1

LL				
No. of lines	$ddLL_{12\text{-}23}$	$ddLL_{23\text{-}34}$	$ddLL_{34\text{-}45}$	$ddLL_{45\text{-}56}$
line #1	0.943	0.9635	0.1672	0.07784
line #60	0.9781	0.05914	0.6774	0.1761
line #140	0.9981	0.2547	0.6772	0.005676
AL				
No. of lines	$ddAL_{12\text{-}23}$	$ddAL_{23\text{-}34}$	$ddAL_{34\text{-}45}$	$ddAL_{45\text{-}56}$
line #1	0.01544	1	1	0.6118
line #60	0.9981	0.1064	0.04943	0.0677
line #140	0.7728	0.5694	0.507	0.03147
TT				
No. of lines	$ddTT_{12\text{-}23}$	$ddTT_{23\text{-}34}$	$ddTT_{34\text{-}45}$	$ddTT_{45\text{-}56}$
line #1	0.617	0.9924	0.02797	0.7666
line #60	0.902	0.6516	0.8718	0.000176
line #140	0.9832	0.02571	0.675	0.01544

The test suggests no significant difference exists between the increases in the *LL* measuring number in successive half-years as the number of half-years after intervention increases. The only difference is between the fifth and sixth half-year of line #140. The claim is weaker for the alignment and twist values, but they can also be stated. Overall, it can be said that the rate of deterioration is gradual, without large jumps for all three lines, but cannot be said to be uniform.

3.3.6 Paired t-Test No. 2

A paired t-test was used to examine from which half-year onward the increase in a given metric in the first half-year and subsequent half-years for a given line will not be the same at a significance value of 0.05.

Table 7 p-values of Paired t-test no. 2

LL				
No. lines	$ddLL_{12\text{-}23}$	$ddLL_{12\text{-}34}$	$ddLL_{12\text{-}45}$	$ddLL_{12\text{-}56}$
line #1	0.943	0.1953	0.0005063	2.07E–07
line #60	0.9781	0.548	0.06804	0.2321
line #140	0.9981	0.1575	0.001176	6.21E-06
AL				
No. lines	$ddAL_{12\text{-}23}$	$ddAL_{12\text{-}34}$	$ddAL_{12\text{-}45}$	$ddAL_{12\text{-}56}$
line #1	2.20E^{-16}	2.20E^{-16}	2.20E^{-16}	2.20E^{-16}
line #60	0.9981	0.2218	6.85E^{-08}	6.30E^{-14}
line #140	0.7728	0.0279	0.002088	1.58E^{-06}
TT				
No. lines	$ddTT_{12\text{-}23}$	$ddTT_{12\text{-}34}$	$ddTT_{12\text{-}45}$	$ddTT_{12\text{-}56}$
line #1	0.617	0.07815	5.37E–^{05}	3.37E^{-09}
line #60	0.902	0.6212	0.4197	0.01846
line #140	0.9832	0.1514	0.007647	1.22E^{-06}

After the analysis, it can be stated that of the three railway lines, line #60 is the one whose characteristic increments are least affected by half-year changes. For line #1, the incremental difference between the *LL* and the *TT* measuring numbers compared to the post-intervention half-year starts to differ significantly from the fourth half-year at the latest.

Welch's two-sample t-tests on all three lines showed that the rate of deterioration is constant, thus demonstrating that the deterioration of the scores is linear in the investigated period. This is somewhat tempered by the expected values of the measures obtained from the descriptive statistics analysis, which shows that the abundance of maintenance interventions pushes the nature of the deterioration patterns towards linearity. In contrast, the absence of maintenance patterns leads to more excellent dispersion and uncertainty, which pushes the deterioration patterns away from linearity.

3.4 Seven Models to Predict the Degradation

Predictive models can only work well if they can predict the state of the railway line as early as possible and as late as possible. After an intervention, it is possible to make empirical predictions based on the first six months of measured data, and maintenance professionals have this capability. With the traffic passing through, the structural design, and the wealth of experience, it is possible to predict the condition for one or two half-years within certain limits, but it is difficult to predict it for a more extended period. At the same time, the question arises as to what procedures exist that, knowing the physical parameters (track, curve, speed, ballast), can provide forecasts for several half-years from a single piece of data within a reasonable margin of error.

The deterioration patterns were modeled using linear, exponential, power logarithm, and natural logarithm regression approximations. As a result of further investigations and studies, the linear and the exponential regression models are the two regression models that can describe the deterioration processes with sufficient confidence. In previous research [55], a detailed description of four simple models was constructed, followed by two regression approximations and one artificial neural network model. Comparisons with lower-order models are worthwhile because the new higher-order model may produce less than the results intended by its creators.

3.4.1 Basic Models

All four basic models are linear and low-order, but they are perfectly adequate for control purposes, and there will be one model out of the four that is particularly good at predicting track geometric degradation.

Basic model #1 (see Eq. (1)):

$$SAD_{n+1} = SAD_n + \frac{(SAD_n - SAD_1)}{n} . \tag{1}$$

Basic model #2 (see Eq. (2)):

$$SAD_{n+1} = SAD_n + (SAD_n - SAD_{n-1}) . \tag{2}$$

Basic model #3 (see Eq. (3)):

$$SAD_{n+1} = SAD_n + \frac{(SAD_n - SAD_{n-1})}{2} + \frac{(SAD_{n-1} - SAD_{n-2})}{2} . \tag{3}$$

Basic model #4 (see Eq. (4)):

$$SAD_{n+1} = SAD_n + \frac{(SAD_n - SAD_{n-1})}{4} + \frac{(SAD_{n-1} - SAD_{n-2})}{4} + \frac{(SAD_{n-2} - SAD_{n-3})}{4} + \frac{(SAD_{n-3} - SAD_{n-4})}{4} \tag{4}$$

where the SAD_t - is the value of the predicted grade point average in the third half-year, n - the index of the calculating half-year.

3.4.2 Regression Models

Of the regression fit tests presented in the previous chapters, the linear and exponential regression equations presented here all have a coefficient of determination of at least $R^2 = 0.75$.

Linear regression model general equation (see Eq. (5)):

$$y_i = \alpha + \beta x_i + u_i \tag{5}$$

Equation currently used (see Eq. (6)):

$$SAD_t = SAD_0 + bt \tag{6}$$

where the SAD_t - is the value of the predicted grade point average in the t^{th} half-year, SAD_0 - the value of the rating number in the first half-year after the employment, b - the track-specific superstructure scaling factor, t - number of half-years since the last job.

Exponential regression model general equation (see Eq. (7)):

$$y_i = \alpha + e^{\beta x_i} + u_i \tag{7}$$

Currently used equation (see Eq. (8)):

$$SAD_t = SAD_0 + e^{bt} \tag{8}$$

3.4.3 Artificial Neural Network model

The neurons of the artificial neural network (ANN) model were constructed using data from a railway with the most consecutive half-years without intervention, and the 500 m sections were assigned the following parameters:

- thickness of ballast bed (40 cm, 50 cm): $\boldsymbol{X}_{1i}$ and $\boldsymbol{X}_{2i}$
- rail syst (MAV48.5, 54E1, 60E1): $\boldsymbol{X}_{3i}$, $\boldsymbol{X}_{4i}$ and $\boldsymbol{X}_{5i}$
- curvature: $\boldsymbol{X}_{6i}$
- straight or curved: $\boldsymbol{X}_{7i}$
- track age or year and half year of SAD number: $\boldsymbol{X}_{8i}$
- from SAD_1 to SAD_6: from $\boldsymbol{X}_{9i}$ to $\boldsymbol{X}_{14i}$

Each index represents the number of columns in the $\boldsymbol{X}$ matrix, so number of neurons in the input layer is 14.

Table 8 Content of the ***X*** matrix

Number of the input neurons	Physical meaning	Type of the input values
1.	Ballast thickness = 40 cm	0 or 1
2.	Ballast thickness = 50 cm	0 or 1
3.	Track system 48.5	0 or 1
4.	Rail system 54	0 or 1
5.	Rail system 60	0 or 1
6.	Curvature	value
7.	Straight or Curved	0 or 1
8.	Age of the track	value
9.	Track geometry measuring number in the 1st half-year	value
10.	Track geometry measuring number in the 2nd half-year	value
11.	Track geometry measuring number in the 3rd half-year	value
12.	Track geometry measuring number in the 4th half-year	value
13.	Track geometry measuring number in the 5th half-year	value
14.	Track geometry measuring number in the 6th half-year	value

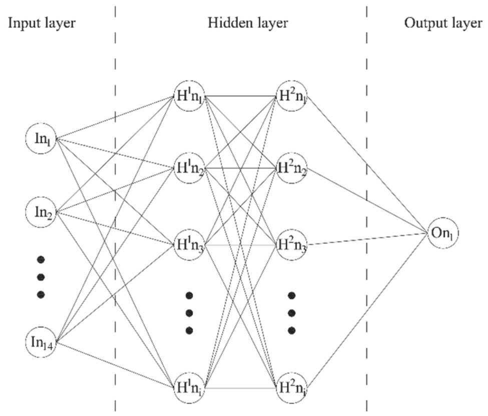

Figure 9 Structure of the ANN

The $\boldsymbol{X}$ matrix is built from the classifier numbers of the given line, the database itself on which the neural network learns. In total, 4985 lines were built into the $\boldsymbol{X}$ matrix. Then, the $\boldsymbol{y}$ target vector had to be defined from the first *SAD* numbers outside the learning timeline. Fig. 9 clearly shows the $\boldsymbol{X}$ matrix's structure and the neural network's topology. In this work, the Levenberg-Marquardt method was used.

For each of the seven models, the root mean square error (RMSE) between the predicted and real values was summed for the respective forecast half-year, shown in Tab. 9 and Fig. 10.

Table 9 RMSE values of the prediction models

	t^{th} half-year after an intervention								
	8	9	10	11	12	13	14	15	16
Basic #1	10.67	15.94	17.28	24.3	33.51	17.95	28.41	34.24	43.22
Basic #2	14.53	21.27	26.87	38.55	40.58	60.89	55.55	66.74	79.25
Basic #3	12.64	15.71	18.82	24.56	22.1	30.81	33.88	30.3	52.11
Basic #4	11.84	14.83	17.9	21.47	20.99	28.53	34.04	33.96	45.25
Lin. Reg.	10.49	14.83	15.49	20.66	26.28	19.05	28.75	36.13	48.98
Exp. Reg.	17.11	29.48	39.71	52.42	85.28	44.21	70.72	85.1	47.88
ANN	10.86	11.54	12.71	17.17	17.29	19.95	22.47	18.95	26.46

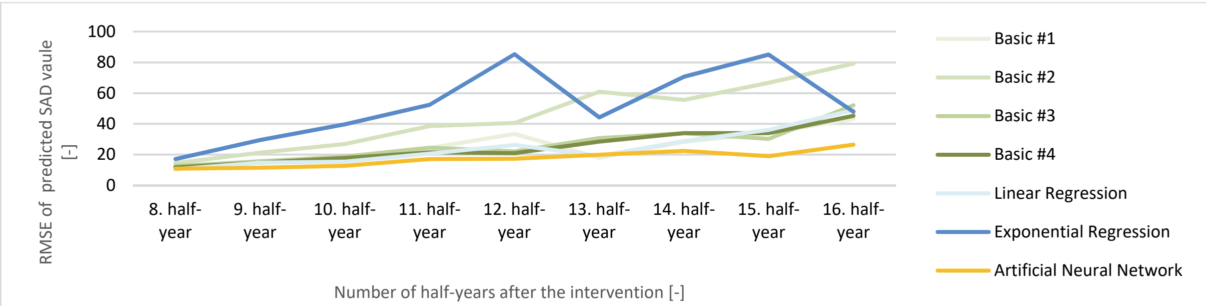

Figure 10 RMSE of the prediction models

The results of the comparison of the models are shown in Tab. 10 that the ANN's forecast for nine half-years is, on average, 22% better than the linear regression model, 27% better than the basic model #4, and 62% better than the exponential regression model.

Table 10 RMSE values differences to ANN

Artificial Neural Network										
	t^{th} half-year after an intervention									Mean Error differences to ANN
	8	9	10	11	12	13	14	15	16	
Basic #4	8%	22%	29%	20%	18%	30%	34%	44%	42%	27%
Linear Regression	–4%	22%	18%	17%	34%	–5%	22%	48%	46%	22%
Exponential Regression	37%	61%	68%	67%	80%	55%	68%	78%	45%	62%

The choice of models for further investigation follows directly from the results, which show that the linear regression and the ANN models have been chosen. In contrast, in the case of the exponential regression model, it is a common idea in the domestic literature that exponentiality is the best descriptor of deterioration behavior. This paper uses the linear, exponential, and ANN models to introduce the deterioration time difference function and the concept of reserved time (see Section 3.5). Both definitions can be applied to any other model.

3.5 Difference between Deterioration Time Function and Reserved Time

After the models (linear regression, exponential regression, and ANN) were built for line #1 and line #140, the *SAD* qualifying number values were predetermined for 20 half-years. As an additional point, it should be noted here that there are no validation half-years of this number for these railway lines, and 20 half-years seems to be an extended period, but the existence and performance of this type of study are essential for a long-term forecast.

After compiling the data for the evaluation, it can be concluded that line #1 and line #140 under study are also characterized by the fact that the superstructure of the line along the length varies. These structural differences and the parameters of the linear and exponential deterioration model are shown in Tab. 11.

Table 11 Functions of the selected models

No. of the line	Track system	Function	SAD_0	b
#1	54	$SAD_{lin54} = 84.48 + 6.024 \cdot t$	84.48	6.024
		$SAD_{exp54} = 85.8 \cdot e^{(0.06785 \cdot t)}$	85.80	0.06785
	60	$SAD_{lin60} = 74.08 + 4.05 \cdot t$	74.08	4.05
		$SAD_{exp60} = 75.28696 \cdot e^{(0.0548 \cdot t)}$	75.29	0.0548
#140	48.5	$SAD_{lin48.5} = 109.80 + 15.66 \cdot t$	109.80	15.66
		$SAD_{exp48.5} = 112.511 \cdot e^{(0.0915 \cdot t)}$	112.51	0.0915
	60	$SAD_{lin60} = 81.94 + 4.23 \cdot t$	81.94	4.23
		$SAD_{exp60} = 83.28 \cdot e^{(0.0584 \cdot t)}$	83.28	0.0584

Tab. 11 contains exponential equations $SAD(t)$, where t means the number of elapsed terms (half a year) since the last maintenance. The equations in Tab. 11 characterize the deterioration of the geometrical condition on track sections of line #1 and #140 built up with rail system MAV48.5, 54E1, and 60E1, shown in Figs. 11 to 14.

Further research might be worthwhile to include a threshold value that prevents the *SAD* number from decreasing in the case of no intervention. However, there is also a pulsation in the real data where a decrease in *SAD* number is observed between consecutive semesters. There could be several reasons for this. The first and well known internationally is that the timing of measurements can be affected by variations in the weather conditions affecting the substructure. So, for example, in the even half of the year the measurement was made in a warmer environment, while in the odd half of the year the measurement was made in a cold environment with a frozen substructure. However, it should also be noted that *SAD* is a value to be calculated from three measurement numbers (*LL*, *AL*, *TT*), which may have the consequence that one of these numbers may be geometrically improved due to certain coexisting conditions, and thus *SAD* may also be improved. Finally, the nature of the ANN's learning flow matrix may result in the prediction behaving in the same way using the pattern of the sample, which may be stunted by this pulsating decay pattern in the real measurement results.

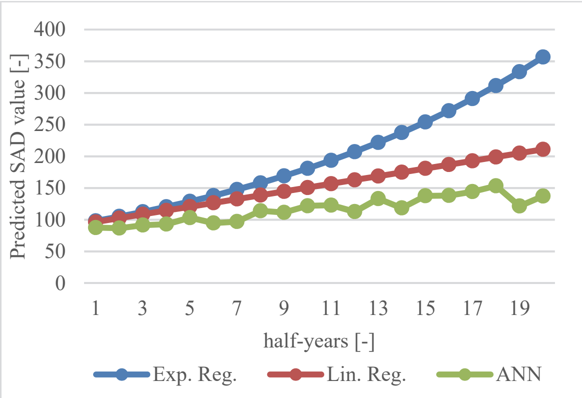

Figure 11 Predicted *SAD* values on line #1 with 54-rail system

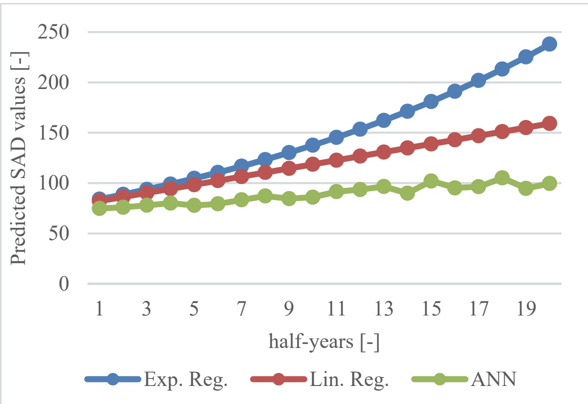

Figure 12 Predicted *SAD* values online #1 with 60 system

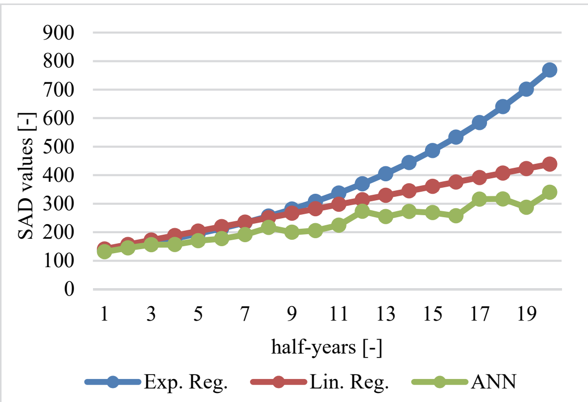

Figure 13 Predicted *SAD* values on line #140 with 48.5 system

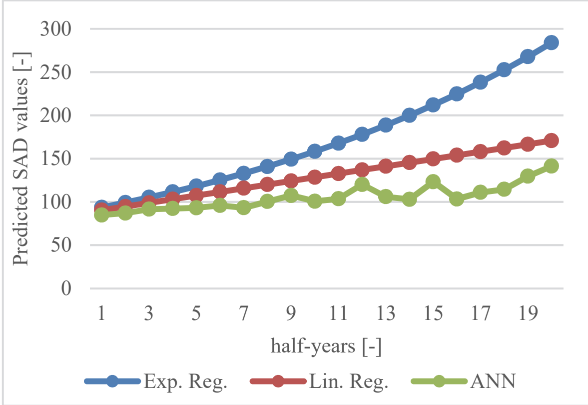

Figure 14 Predicted *SAD* values on line #140 with 60 system

Based on the result above, the deterioration condition of a track built with a given rail system can be estimated as a function of the terms if there is no regulation work in the future. Furthermore, with the current *SAD* qualifying number of the 500 m section with the rail system, it can be determined which *SAD* values would have evolved if they had been built with superstructure system 60.

In a previous article [55], the procedure is explained in detail, which, using the exponential regression model, shows what the *SAD* numbers of the 500 m sections of the line would take on the number *t* of terms after the last maintenance if the line had a 60-system superstructure instead of a 54-rail system superstructure. It should be noted that the *SAD* is related to the quality of the superstructure and the ballast and substructure. Consequently, tracks with 60 kg/m rails are newer, and the ballast and substructure are of better quality.

If, therefore, sections of an existing line with the same superstructure, in the case of line #1 with a 54 system, were to be rebuilt with a 60 rail system superstructure, it can be calculated how much more economical the operation would be due to the absence of regulatory interventions, if the number of interventions is only taken into account assuming the same intervention efficiency.

Table 12 Functions of difference of deterioration time

No. of line	Function of difference of deterioration time	Steepness	t_0	Type of the model
#1	$t_{DTD,exp} = 1.2381 \cdot t + 1.3853$	1.2381	1.3853	exp
	$t_{DTD,lin} = 0.4874 \cdot t + 2.5679$	0.4874	2.5679	lin
#140	$t_{DTD,exp} = 1.5668 \cdot t + 4.1511$	1.5668	4.1511	exp
	$t_{DTD,lin} = 2.8021 \cdot t + 6.5863$	2.8021	6.5863	lin

In the function (t_{DTD}), describing the deterioration time difference, *t* is the number of half-years after the maintenance, and t_0 shows how many terms of time difference there are between the bigger and the smaller weighed rail systems on the given line after the moment of maintenance. So, on line #1, the line sections constructed with 60 superstructure system have a deterioration time difference of 2.57 (one term is equal to a half year), terms with the linear model and 1.4 with the exponential model compared to structure system 54, and on line #140, the reserve is at least 6.5 terms with the linear model and 4.1511 with the exponential model compared to system 48.5 superstructure. The equations in Tab. 11 also show how many t_{DTD} terms will be elapsed until the sections with track system 60 will be in the same condition as those with track system 54 in the actual half-year.

In the previous sections, the linear, exponential, and ANN-modelled degradation processes of the sections of lines #1 and #140 constructed with different track systems have been determined. In the following, the authors present an analysis of how the deterioration rate and the deterioration speed of the 500 m sections of line #1, which can be characterized by a given scale, compared with the standards in force in the country.

According to Hungarian regulation D. 54, chapter 51, [56], the qualifying categories for continuously welded rail tracks (500-meter qualifying length, v = 160 km/h track speed) are the following (Tab. 12).

The geometric deterioration functions of track sections are separated and listed in Tab. 13 according to track systems on line #1 with the linear, exponential, and ANN models.

Table 13 Value of *SAD* number limit in Q_i categories 160 km/h speed

Q_1	Q_2	Q_3	Q_4	Q_5	Q_6
67.3	75.9	86.8	105.4	118.6	148.4

It is important to note that the ANN prediction model behaves as a quasi-black box, i.e., it does not have an exact formulation. As a result, it can only be of practical use if the ANN forecasting model is made available to engineers for each line by centralized control and used by engineers for each forecast. The other way is to use linear regression modeling of the forecasting models generated by the ANN with central control built up over time and to share this exact formula. This latter solution may seem feasible, given the authors' current capabilities. Thus, the ANN model approximation has been made, and the descriptive equation shown in the Table below is also the result of this process.

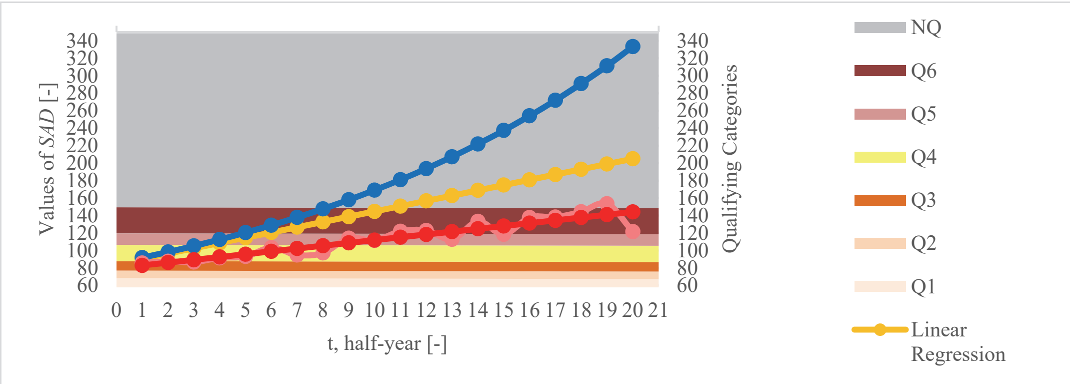

Figure 15 Predicted *SAD* values of 54 superstructure system with qualifying categories

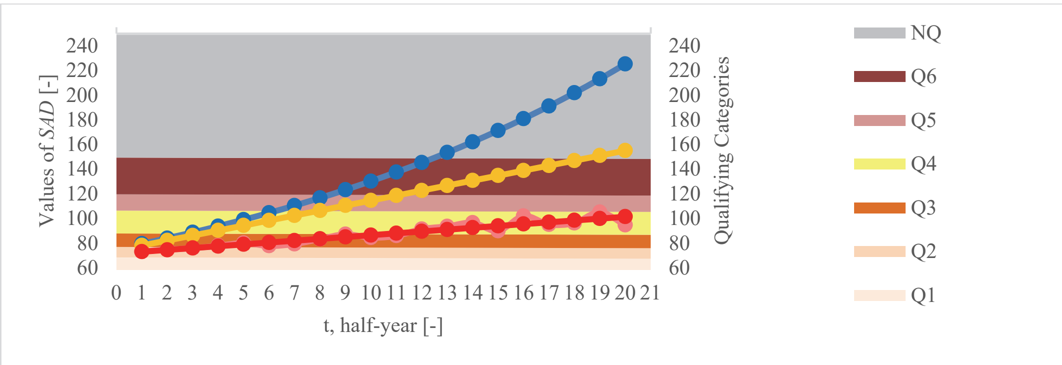

Figure 16 Predicted *SAD* values of 60 superstructure system with qualifying categories

Fig. 15 and Fig. 16 were created using $SAD(t)$ deterioration functions according to Tab. 13 and regulation D. 54, chapter 51. All of them can describe the change in the $SAD(t)$ deterioration function, but the differences are clearly visible in Figs. 15 and 16, illustrating the time after the last intervention until the limit of the given rating category is reached. This t_{Qi} is the reserved time for the given rating category calculated using the formula below for the exponential models (see Eq. (9)):

$$t_{Q_i} = \frac{\ln\left(Q_i / SAD_0^j\right)}{b^j}, \tag{9}$$

and for linear models (see Eq. (10)):

$$t_{Q_i} = \left(Q_i - SAD_0^j\right) / b^j . \tag{10}$$

Q_i is the *SAD* value of the given qualifying category from Tab. 12; i is the given qualifying categories; j is the system of the superstructures and numbers 54 and 60; b comes from Tab. 13.

Table 14 Geometric deterioration functions

No. Line	Track system	Function	SAD_0	b
#1	54	$SAD_{\text{lin54}} = 84.48 + 6.024 \cdot t$	84.48	6.02400
		$SAD_{\text{exp54}} = 85.8 \cdot e^{(0.06785 \cdot t)}$	85.80	0.06785
		$SAD_{\text{ANN54 with Lin. Reg.}} = 79.67 + 3.224 \cdot t$	79.67	3.22400
	60	$SAD_{\text{lin60}} = 74.08 + 4.05 \cdot t$	74.08	4.05000
		$SAD_{\text{exp60}} = 75.28696 \cdot e^{(0 \cdot 0548 \cdot t)}$	75.29	0.05480
		$SAD_{\text{ANN60 with Lin. Reg.}} = 71.55 + 1.496 \cdot t$	71.55	1.49600

The reserve times belonging to the different track systems and qualifying categories on line #1 are in Tab. 14 and Tab. 15.

Table 15 The reserve time in half-year of the 54 superstructure system

	t_{M1}	t_{M2}	t_{M3}	t_{M4}	t_{M5}	t_{M6}
Lin. Reg. 54 rail system	-	-	0.39	3.47	5.66	10.61
Exp. Reg. 54 rail system	-	-	0.17	3.03	4.77	8.08
ANN Lin. Reg. 54 r. s.	-	-	2.21	7.98	12.08	21.32

Table 16 The reserve time in half-year of the 60 superstructure system

	t_{M1}	t_{M2}	t_{M3}	t_{M4}	t_{M5}	t_{M6}
Lin. Reg. 54 r. s.	-	0.45	3.14	7.73	10.99	18.35
Exp. Reg. 54 r. s.	-	0.15	2.60	6.14	8.29	12.38
ANN Lin. Reg. 54 rail system	-	2.91	10.20	22.65	31.48	51.42

Following the example above, this procedure can determine the reserve time (t_{Mi}) of any selected line, helping to manage the track geometry quality control in time.

In addition to using one model to determine the reserve time, the deterioration process provided by the three models in this paper can be compared. Based on validated tests, it can be stated with good confidence that the exponential model is the least accurate in terms of prediction, while the ANN is the most accurate. Suppose the authors accept the reserve time generated by the ANN model. In that case, the time spent in a given classification category for line sections with different superstructures for the whole line can be used to make longer-term intervention plans, thus helping the national railway maintenance system. The reserve time function shows the difference in geometric condition between track sections with different superstructure systems six months after the work has been carried out. This function provides essential data for correctly determining the date of interceptions.

It is clear that on a railway line or in a part of the network where maintenance resources are insufficient to prevent the introduction of speed-restriction signals, it is essential to keep railway lines in the best possible condition and carry out interventions in good time.

4 CONCLUSIONS

This paper presents a comprehensive approach to understanding and predicting the geometric deterioration of railway tracks using a novel computer procedure using data produced by FMK-004 track geometry measuring car. The study provides valuable insights into the nature of track deterioration through statistical analysis and the development of predictive models, including an advanced Artificial Neural Network (ANN) model. Introducing the deterioration time difference function and the concept of reserve time further enhances the ability to plan maintenance interventions strategically, offering a significant step forward in railway maintenance and safety management. The findings underscore the importance of adopting sophisticated analytical tools and models to improve the predictability and efficiency of railway track maintenance. Ultimately, this research contributes to the ongoing efforts to enhance railway infrastructure reliability, ensuring safer and more efficient railway operations. In the following paragraphs, the main conclusions are written in detail.

A computer procedure has been created that automatically processes the measurement and qualification data provided by the FMK-004 track geometry measuring car, determines the change in geometric condition based on time series values, provides a forecast of the change with its accuracy, determines the place and time of the necessary maintenance intervention, helps the engineer in track condition analysis and the preparation of more complex professional decisions. The procedure is innovative in processing and integrating a vast amount of complex data, aiding in efficient track maintenance planning and execution.

Welch's two-sample t-test on three railway lines (#1, #60, and #140) showed that the rate of deterioration is constant in the investigated period. This is somewhat tempered by the expected values of the measures obtained from the descriptive statistical analysis, which shows that the abundance of maintenance interventions pushes the nature of the deterioration patterns towards linearity. In contrast, the absence of maintenance patterns leads to higher dispersion and uncertainty, which pushes the deterioration patterns away from linearity. This finding suggests that a linear degradation model is applicable for these lines, but variability in maintenance activities can influence deterioration behavior.

The deterioration time difference function has been defined, which can be used to characterize the effect of the implementation or non-implementation of a superstructure upgrade on the change in geometric condition numerically, and A concept has been introduced with the reserve time and defined its calculation method, which expresses the length of time needed for the railway track to deteriorate from its repaired condition after the upgrade to the limit value of the given classification category. This methodological approach allows for quantifying the benefits of superstructure upgrades and planning maintenance activities more strategically.

In the example of three railway lines: #1, #60, #140 the Vaszary deterioration equation was shown to be suitable for describing the geometric deterioration of the railway track but not for predicting the deterioration. It is thus demonstrated that a method must be sought that models the prediction of track condition changes more reliably. The example of the same lines showed that the new model (ANN) gives significantly better results than the previous models. As a result, the average prediction error of the geometric deterioration of the line reduced to 38% of the exponential prediction error, while the absolute error of the model remains below 27% even for the ninth forecast half-year. This advancement in predictive modeling using ANN over traditional methods underscores the potential for more accurate and reliable forecasts of track conditions, significantly impacting maintenance planning and safety assessments.

5 REFERENCES

[1] Jia, R., Shao, S., & Yang, L. (2021). High-speed rail and CO_2 emissions in urban China: A spatial difference-in-differences approach. *Energy Econonomics*, *99*, 105271. https://doi.org/10.1016/j.eneco.2021.105271

[2] Ramazan, B., Mussaliyeva, R., Bitileuova, Z., Naumov, V., & Taran, I. (2021). Choosing the logistics chain structure for deliveries of bulk loads: Case study of the Republic

Kazakhstan. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, *2021*(3), 142-147. https://doi.org/10.33271/nvngu/2021-3/142

[3] Nugymanova, G., Nurgaliyeva, M., Zhanbirov, Z., Naumov, V., & Taran I. (2021). Choosing a servicing company's strategy while interacting with freight owners at the road transport market. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, *2021*(1), 204-210. https://doi.org/10.33271/nvngu/2021-1/204

[4] Lichtberger, B. (2005). *Track compendium*. Eurail press, Hamburg, Germany, 634.

[5] Ézsiás, L., Tompa, R., & Fischer, S. (2024). Investigation of the possible correlations between specific characteristics of crushed stone aggregates. *Spectrum of Mechanical Engineering and Operational Research*, *1*(1), 10-26. https://doi.org/10.31181/smeor1120242

[6] Bian, X., Jiang, H., Cheng, C., Chen, Y., Chen, R., & Jiang, J. (2014). Full-scale model testing on a ballast less high-speed railway under simulated train moving loads. *Soil Dynamics and Earthquake Engineering*, *66*, 368-384. https://doi.org/10.1016/j.soildyn.2014.08.003

[7] Brautigam, A., Szalai, S., & Fischer, S. (2023). Investigation of the application of austenitic filler metals in paved tracks for the repair of the running surface defects of rails considering field tests. *Facta Universitatis, Series: Mechanical Engineering*, 12081. https://doi.org/10.22190/FUME230828032B

[8] Németh, A. & Fischer, S. (2021). Investigation of the glued insulated rail joints applied to CWR tracks. *Facta Universitatis, Series: Mechanical Engineering*, *19*(4), 681-704. https://doi.org/10.22190/FUME210331040N

[9] Fischer, S., Harangozó, D., Németh, D., Kocsis, B., Sysyn, M., Kurhan, D., & Brautigam, A. (2023). Investigation of heat-affected zones of the material weldings. *Facta Universitatis, Series: Mechanical Engineering*, 11420. https://doi.org/10.22190/FUME221217008F

[10] Kuchak, A.-T.-J., Marinkovic, D., & Zehn, M. (2021). Parametric investigation of a rail damper design based on a lab-scaled model. *Journal of Vibration Engineering & Technologies*, *9*, 51-60. https://doi.org/10.1007/s42417-020-00209-2

[11] Kuchak, A.-T.-J., Marinkovic, D., & Zehn, M. (2020). Finite element model updating - Case study of a rail damper. *Structural Engineering and Mechanics*, *73*(1), 27-35. https://doi.org/10.12989/sem.2020.73.1.027

[12] Kou, L., Sysyn, M., & Liu, J. (2024). Influence of crossing wear on rolling contact fatigue damage of frog rail. *Facta Universitatis, Series: Mechanical Engineering*, *22*(1), 25-44. https://doi.org/10.22190/FUME220106024K

[13] Sysyn, M., Gerber, U., Kluge, F., Nabochenko, O., & Kovalchuk, V. (2020). Turn out remaining use full life prognosis by means of on-board inertial measurements on operational trains. *International Journal of Rail Transportation*, *8*(4), 347-369. https://doi.org/10.1080/23248378.2019.1685918

[14] Kovalchuk, V., Sysyn, M., Gerber, U., Nabochenko, O., Zarour, J., & Dehne, S. (2019). Experimental investigation of the influence of train velocity and travel direction on the dynamic behavior of stiff common crossings. *Facta Universitatis, Series: Mechanical Engineering*, *17*(3), 345-356. https://doi.org/10.22190/FUME190514042K

[15] Sysyn, M., Nabochenko, O., Kovalchuk, V., Gruen, D., & Pentsak, A. (2019). Improvement of inspection system for common crossings by track side monitoring and prognostics. *Structural Monitoring and Maintenance*, *6*(3), 219-235. https://doi.org/10.12989/SMM.2019.6.3.219

[16] Suiker, A.-S. & de Borst, R. (2003). A numerical model for the cyclic deterioration of railway tracks. *International Journal for Numerical Methods in Engineering*, *57*(4), 441-470. https://doi.org/10.1002/nme.683

[17] Kurhan, M., Kurhan, D., & Hmelevska, N. (2024). Innovative approaches to railway track alignment optimization, in curved sections. *Acta Polytechnica Hungarica*, *21*(1), 207-220. https://doi.org/10.12700/APH.21.1.2024.1.13

[18] Kurhan, D., Kurhan, M., & Husak, M. (2020). Impact of the variable stiffness section on the conditions of track and rolling stock interaction. *IOP Conference Series: Materials Science and Engineering*, *985*(1), 012005. https://doi.org/10.1088/1757-899X/985/1/012005

[19] Kurhan, D. & Kurhan, M. (2019). Modeling the dynamic response of railway track. *IOP Conference Series: Materials Science and Engineering*, *708*(1), 012013. https://doi.org/10.1088/1757-899X/708/1/012013

[20] Sresakoolchai, J. & Kaewunruen, S. (2023). Railway infrastructure maintenance efficiency improvement using deep reinforcement learning integrated with digital twin based on track geometry and component defects. *Scientific Reports*, *13*(1), 2439. https://doi.org/10.1038/s41598-023-29526-8

[21] Su, Z., Jamshidi, A., Núñez, A., Baldi, S., & De Schutter, B. (2017). Multi-level condition-based maintenance planning for railway infrastructures - a scenario-based chance-constrained approach. *Transportation Research Part C: Emerging Technologies*, *84*, 92-123. https://doi.org/10.1016/j.trc.2017.08.018

[22] Al-Douri, Y.-K., Tretten, P., & Karim, R. (2016). Improvement of railway performance: a study of Swedish railway infrastructure. *Journal of Modern Transportation*, *24*, 22-37. https://doi.org/10.1007/S40534-015-0092-0

[23] Andrade, A.-R. & Teixeira, P.-F. (2011). Biobjective optimization model for maintenance and renewal decisions related to rail track geometry. *Transportation Research Record*, *2261*(1), 163-170. https://doi.org/10.3141/2261-19

[24] Vaszary, P. (1992). *Theoretical basis for the assessment of the geometric condition of railway track.* C.Sc. dissertation, Hungarian Academy of Sciences, 118.

[25] Vaszary, P., Kiss, F., Koren, C.-E.., Stadler, T., & Horváth, F. (1986). *Geometrical tolerances for railway superstructure I*. Technical report ID. 91-113-80, Széchenyi István KTMF, 38.

[26] Vaszary, P., Koren, C.-E., Kiss, F., Kontra, G., & Stadler, T. (1987). *Geometrical tolerances for railway superstructure II*. Technical report ID. 91-113-84, Széchenyi István KTMF, 31.

[27] Veit, P. & Marschnig, S. (2010). Sustainability in track: a precondition for high speed traffic. *Proceedings of the Joint Rail Conference 2010*, 349-355. https://doi.org/10.1115/JRC2010-36225

[28] Veit, P. (2007). Track quality-luxury or necessity? *Railway Technical Review - RTR Special*, *July*, 8-12.

[29] European Committee for Standardization (2017). EN 13848-5:2017. Railway applications - Track - Track geometry quality - Part 5: Geometric quality levels - Plain line, switches and crossings, 30.

[30] Andrade, A.-R. & Teixeira, P.-F. (2011). Uncertainty in rail-track geometry degradation: Lisbon-Oporto line case study. *Journal of Transportation Engineering*, *137*(3), 193-200. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000206

[31] Andrade, A.-R. & Teixeira, P.-F. (2012). A Bayesian model to assess rail track geometry degradation through its life-cycle. *Research in Transportation Economics*, *36*(1), 1-8. https://doi.org/10.1016/j.retrec.2012.03.011

[32] López-Pita, A., Teixeira, P.-F., Casas, C., Ubalde, L., & Robusté, F. (2007). Evolution of track geometric quality in high-speed lines: Ten years experience of the Madrid-Seville line. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *221*(2), 147-155. https://doi.org/10.1243/0954409JRRT62

[33] Andrade, A.-R. & Teixeira, P.-F. (2014). Unplanned-maintenance needs related to rail track geometry. *Proceedings of the Institution of Civil Engineers-Transport*, *167*(6), 400-410. https://doi.org/10.1680/TRAN.11.00060

[34] Letot, C., Soleimanmeigouni, I., Ahmadi, A., & Dehombreux, P. (2016). An adaptive opportunistic maintenance model based on railway track condition prediction. *IFAC-Papers on Line*, *49*(28), 120-125. https://doi.org/10.1016/j.ifacol.2016.11.021

[35] Lethanh, N. & Adey, B.-T. (2016). A real option approach to determine optimal intervention windows for multi-national rail corridors. *Journal of Civil Engineering and Management*, *22*(1), 38-46. https://doi.org/10.3846/13923730.2014.994030

[36] Li, H. & Xu, Y. (2009). Railway track integral maintenance index and its application. *Proceedings of the International Conference on Transportation Engineering 2009*, 2514-2519. https://doi.org/10.1061/41039(345)415

[37] Lautala, P. & Pouryousef, H. (2011). Sensitivity analysis of track maintenance strategies for the high speed rail (HSR) services. *Proceedings of the Joint Rail Conference 2011*, 141-150. https://doi.org/10.1115/JRC2011-56112

[38] Vale, C. & Ribeiro, I.-M. (2014). Railway condition-based maintenance model with stochastic deterioration. *Journal of Civil Engineering and Management*, *20*(5), 686-692. https://doi.org/10.3846/13923730.2013.802711

[39] Verma, R., Koul, S., & Prasad, B. (2020). Expert System - A Case of Indian Railway's Track Maintenance and Renewal Operations. *Proceedings of 3rd International Conference on Innovative Computing and Communication (ICICC-2020)*, 3563091. https://doi.org/10.2139/ssrn.3563091

[40] Shang, H. & Bеŕenguer, C. (2015). Delayed maintenance model for deteriorating track using colored Petri Nets. *IFAC-Papers on Line*, *48*(21), 464-469. https://doi.org/10.1016/j.ifacol.2015.09.570

[41] Santos, R., Teixeira, P.-F., & Antunes, A.-P. (2015). Planning and scheduling efficient heavy rail track maintenance through a Decision Rules Model. *Research in Transportation Economics*, *54*, 20-32. https://doi.org/10.1016/j.retrec.2015.10.022

[42] Reddy, V., Chattopadhyay, G., Larsson-Kråik, P.-O., & Hargreaves, D.-J. (2007). Modelling and analysis of rail maintenance cost. *International Journal of Production Economics*, *105*(2), 475-482. https://doi.org/10.1016/j.ijpe.2006.03.008

[43] Pargar, F., Kauppila, O., & Kujala, J. (2017). Integrated scheduling of preventive maintenance and renewal projects for multi-unit systems with grouping and balancing. *Computers & Industrial Engineering*, *110*, 43-58. https://doi.org/10.1016/j.cie.2017.05.024

[44] Moreu, F., Spencer, B.-F., Foutch, D.-A., & Scola, S. (2017). Consequence-based management of railroad bridge networks. *Structure and Infrastructure Engineering*, *13*(2), 273-286. https://doi.org/10.1080/15732479.2016.1162817

[45] Pouryousef, H., Teixeira, P., & Sussman, J. (2010). Track maintenance scheduling and its interactions with operations: Dedicated and mixed high-speed rail (HSR) scenarios. *Proceedings of the Joint Rail Conference 2010*, 317-326.

[46] Moghaddam, K.-S. & Usher, J.-S. (2011). Preventive maintenance and replacement scheduling for repairable and maintainable systems using dynamic programming. *Computers & Industrial Engineering*, *60*(4), 654-665. https://doi.org/10.1016/j.cie.2010.12.021

[47] Gerum, P.-C.-L., Altay, A., & Baykal-Gürsoy, M. (2019). Data-driven predictive maintenance scheduling policies for railways. *Transportation Research Part C: Emerging Technologies*, *107*, 137-154. https://doi.org/10.1016/j.trc.2019.07.020

[48] Moridpour, S., Mazloumi, E., & Hesami, R. (2017). Application of artificial neural networks in predicting the degradation of tram tracks using maintenance data. *Applied big data analytics in operations management*, *IGI Global*, 30-54. https://doi.org/10.4018/978-1-5225-0886-1.ch002

[49] Guler, H. (2014). Prediction of railway track geometry deterioration using artificial neural networks: a case study for Turkish state railways. *Structure and Infrastructure Engineering*, *10*(5), 614-626. https://doi.org/10.1080/15732479.2012.757791

[50] Khajehei, H., Ahmadi, A., Soleimanmeigouni, I., & Nissen, A. (2019). Allocation of effective maintenance limit for railway track geometry. *Structure and Infrastructure Engineering*, *15*(12), 1597-1612. https://doi.org/10.1080/15732479.2019.1629464

[51] Karimpour, M., Hitihamillage, L., Elkhoury, N., Moridpour, S., & Hesami, R. (2018). Fuzzy approach in rail track degradation prediction. *Journal of Advanced Transportation*, *2018*, 3096190. https://doi.org/10.1155/2018/3096190

[52] Khajehei, H., Ahmadi, A., Soleimanmeigouni, I., Haddadzade, M., Nissen, A., & Latifi Jebelli, M.-J. (2022). Prediction of track geometry degradation using artificial neural network: a case study. *International Journal of Rail Transportation*, *10*(1), 24-43. https://doi.org/10.1080/23248378.2021.1875065

[53] Sharma, S., Cui, Y., He, Q., Mohammadi, R., & Li, Z. (2018). Data-driven optimization of railway maintenance for track geometry. *Transportation Research Part C: Emerging Technologies*, *90*, 34-58. https://doi.org/10.1016/j.trc.2018.02.019

[54] Nagy, R. (2024). Detailed procedure of the analysis of railway deterioration. *Online graphical dataset.* https://doi.org/10.13140/RG.2.2.10522.66248

[55] Nagy, R. (2017). Analytical differences between seven prediction models and the description of the rail track deterioration process through these methods. *Intersectii/Intersections*, *14*(1), 14-32.

[56] Hungarian State Railways (2014). 14978/2014/MÁV, D.54. *Technical specifications for construction and track maintenance I.* Chapter 51, Budapest, 54.

Contact information:

Richárd NAGY
Department of Transport Infrastructure,
Faculty of Architecture, Civil Engineering and Transport Sciences,
Széchenyi István University,
Gyor, 9026, Hungary
Vehicle Industry Research Center,
Széchenyi István University,
Gyor, 9026, Hungary

Ferenc HORVÁT
Department of Transport Infrastructure,
Faculty of Architecture, Civil Engineering and Transport Sciences,
Széchenyi István University,
Gyor, 9026, Hungary
Vehicle Industry Research Center,
Széchenyi István University,
Gyor, 9026, Hungary

Szabolcs FISCHER
(Corresponding author)
Department of Transport Infrastructure,
Faculty of Architecture, Civil Engineering and Transport Sciences,
Széchenyi István University,
Gyor, 9026, Hungary
Vehicle Industry Research Center,
Széchenyi István University,
Gyor, 9026, Hungary
E-mail: fischersz@sze.hu