Transportation Research Part C 95 (2018) 185–206

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

A decision support approach for condition-based maintenance of rails based on big data analysis☆

Ali Jamshidi[a], Siamak Hajizadeh[a], Zhou Su[b], Meysam Naeimi[a], Alfredo Núñez[a,*], Rolf Dollevoet[a], Bart De Schutter[b], Zili Li[a]

[a] Section of Railway Engineering, Delft University of Technology, Delft, The Netherlands
[b] Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

Keywords:
Decision support system
Condition-based maintenance
Rail surface defects
Fuzzy inference system
Axle Box Acceleration (ABA) system

ABSTRACT

In this paper, a decision support approach is proposed for condition-based maintenance of rails relying on expert-based systems. The methodology takes into account both the actual conditions of the rails (using axle box acceleration measurements and rail video images) and the prior knowledge of the railway track. The approach provides an integrated estimation of the rail health conditions to support the maintenance decisions for a given time period. An expert-based system is defined to analyse interdependency between the prior knowledge of the track (defined by influential factors) and the surface defect measurements over the rail. When the rail health conditions is computed, the different track segments are prioritized, in order to facilitate grinding planning of those segments of rail that are prone to critical conditions. In this paper, real-life rail conditions measurements from the track Amersfoort-Weert in the Dutch railway network are used to show the benefits of the proposed methodology. The results support infrastructure managers to analyse the problems in their rail infrastructure and to efficiently perform a condition-based maintenance decision making.

1. Introduction

The increase in train traffic and axle loads affect the health conditions of railway infrastructure. Hence, efficient infrastructure monitoring and maintenance is among the major concerns of infrastructure managers in order to improve the performance of railway operations (Åhrén and Parida, 2009). As such, infrastructure health conditions should be monitored and considered in the maintenance decision making process. Effective management of infrastructure health conditions is crucial to guarantee the desired asset quality level (Parida and Chattopadhyay, 2007; Gandomi and Haider, 2015; Zywiel and Oberlechner, 2001). It also plays an important role in meeting the demands for the whole system performance when the infrastructure is upgraded e.g. when increasing traffic capacity, the maintenance regime should be adapted to avoid compromising safety and infrastructure health requirements. To keep the infrastructure system working at an effective level, a conditions-based maintenance system is required not only to consider the actual heath conditions but also evolution during the maintenance decision horizon (Jamshidi et al., 2017b; Li et al., 2014).

Condition-based monitoring is used in railway infrastructures to estimate the actual health conditions of the assets, so that degradation processes can be effectively controlled. It helps to keep the infrastructure manager continually informed of the estimated health of the railway infrastructure. Condition-based monitoring is supposed to collect information that will allow an effective

☆ This article belongs to the Virtual Special Issue on "Big Data Railway".

* Corresponding author.
E-mail address: a.a.nunezvicencio@tudelft.nl (A. Núñez).

https://doi.org/10.1016/j.trc.2018.07.007
Received 24 August 2017; Received in revised form 29 June 2018; Accepted 14 July 2018

operation by reducing maintenance cost, eliminating unnecessary operations, and focusing on places where the problems are located and where they will be in the coming period. Furthermore, the enhancement in usage of the railway infrastructure needs a systematic monitoring plan to keep the trains running safely by considering all related data influencing the health conditions. The data for a typical railway infrastructure includes a large amount of frequent measurements from the monitoring systems of the assets involved in the railway operations. To ensure the required performance level, a huge amount of data should be collected, transmitted, processed, and properly stored so that it can be used as historical information. This whole process reflects a transition from raw infrastructure data into actionable maintenance knowledge. Therefore, the database constituted from continuous monitoring will become larger and larger over time and applying big data analysis approaches is inevitable (Fumeo et al., 2015). In order to design proper maintenance plans in railways, it is necessary to explore and analyse the growing amount of data and to extract useful information. To do so, different sensors can be used to collect the data obtained in railway track monitoring at different times, environmental conditions, and frequencies. These data can exhibit different characteristics: (1) discrete or continuous, (2) spatial or temporal, (3) signal and images among others (Lasisi and Attoh-Okine, 2018; Attoh-Okine, 2017; He et al., 2013; Liu and Dick, 2016; Ghofrani et al., 2018).

In condition-based maintenance for railways, the monitoring data are mostly collected periodically with regular sampling intervals. For some critical assets, the monitoring can be adapted to other possible needs including continuous measurement. The essential concept for the monitoring data is to take the degradation of the infrastructure into account, in particular for critical infrastructure like rails. This paper focuses on rail conditions monitoring, which has a critical role in the network performance (He et al., 2015, 2013). In an intensively used network, a considerable amount of the maintenance budget has to be allocated for the rail, e.g., in the Dutch railway network, this amounts to almost half of the annual maintenance budget (approximately 60.000 euro/km) (Zoeteman et al., 2014). As a high percentage of failures are directly related to the rail, it is important to assess the rail conditions in order to obtain a proper condition-based maintenance approach. More specifically, the health conditions analysis involves detecting the rail surface defects that can potentially result in rail breaks and derailment in extreme cases (Liu et al., 2011, 2012; Islam et al., 2016; Xu and Zhai, 2017).

Rolling contact fatigue (RCF) affects the health conditions of the rail due to the contact in the interface between wheel and rail (Makino et al., 2016). RCF is a generic term describing a range of rail surface defects and has been an interesting challenging research topic, in particular the influence of RCF on maintenance decision making (Sciammarella et al., 2016). Moreover, its influence is related to other factors including traffic type, train speed, traffic load, rail/wheel profile, train characteristics, and maintenance policy (Popović et al., 2013). Once RCF appears, it induces considerable dynamical forces on the rail surface, and subsequently cracks are initiated and propagated from the surface (Zhuang et al., 2018; Makino et al., 2012). The most important cause of the appearance of defects is the large number of trains passing over rail critical components, most significantly at welds, joints, and switches (Molodova et al., 2014). Early detection of surface defects is important to mitigate induced maintenance costs as well as unforeseeable consequences of rail breaks. There are different methods to diagnose the conditions of rail defects, including ultrasonic measurements (Fan et al., 2007), eddy current testing (Song et al., 2011), and guided-wave based monitoring (Mariani et al., 2013). In this paper, the focus is on a type of rail surface defect called squat. The costs for treating these defects in the Dutch railway network are considerably high (more than 5000 euro/km per year) (Molodova et al., 2014). The maintenance of squats should be different according to their severity. For late-stage squats, a rail replacement plan is a proper decision while for the light squats, grinding a thin layer from the rail surface is the most effective solution. Hence, when all residual damages are removed, grinding is effective and the rail will be turned to a healthy condition. To optimally plan grinding operations, condition-based maintenance relying on early detection of the squats is required. Although a defect detection method could give an indication of the health of the rail, the infrastructure manager requires prior knowledge to (1) be aware of all influential factors, (2) analyse interdependency between the rail observations and the influential factors, and (3) obtain a future view of the track conditions. In this paper, we relate influential factors to the rail health conditions to show the effect of the track characteristics in the rail observations. Hence, by having knowledge about the track characteristics, potential risks about the rail can be anticipated due to the effect of the influential factors on the appearance of defects and consequently on the rail health conditions. Therefore, an analysis of influential factors should be taken into account to give at the most a proper prospect of the infrastructure health conditions.

Mixed Integer Linear Programming (MILP) is a common approach for track maintenance scheduling. An MILP model is developed in Wen et al. (2016) for optimal condition-based preventive maintenance for a single track divided into multiple segments, considering various economic and technical factors such as train speed limits and track quality. The optimal planning of routine maintenance activities and projects like grinding to minimize maintenance costs and track possession time for a single track is formulated as an MILP in Budai et al. (2006). The optimal scheduling of rail, sleeper, and ballast renewal at a network level is formulated as an MILP problem in Caetano and Teixeira (2016) to minimize the expected life-cycle cost and track unavailability. In Peng and Ouyang (2014), the optimal clustering of track maintenance jobs into projects to minimize total maintenance costs for the network of track is recast as a Vehicle Routing Problem (VRP). The track maintenance problem considering different priorities for each section in the railway network is formulated as an VRP with customer costs in Heinicke et al. (2015). A time-space network model is developed in Peng and Ouyang (2012) for the optimal scheduling of capital maintenance projects like rail replacement. A metaheuristic based on simulated annealing is developed in Santos and Teixeira (2011) to determine the optimal tamping length of a tamping machine, minimizing the associated logistic costs and fixed machine costs. In this paper, we use a simplified MILP model to optimize the rail grinding decision plan into clusters that can be related to the actual conditions of the rail. The proposed MILP model not only uses different clusters for determining the most critical pieces of tracks, but also simultaneously takes time and budget constraints into account. Moreover, the model benefits a new method for estimating rail health conditions as an input data. This eases implementing the condition-based maintenance strategy, reaching an effective maintenance plan in terms of rail health conditions

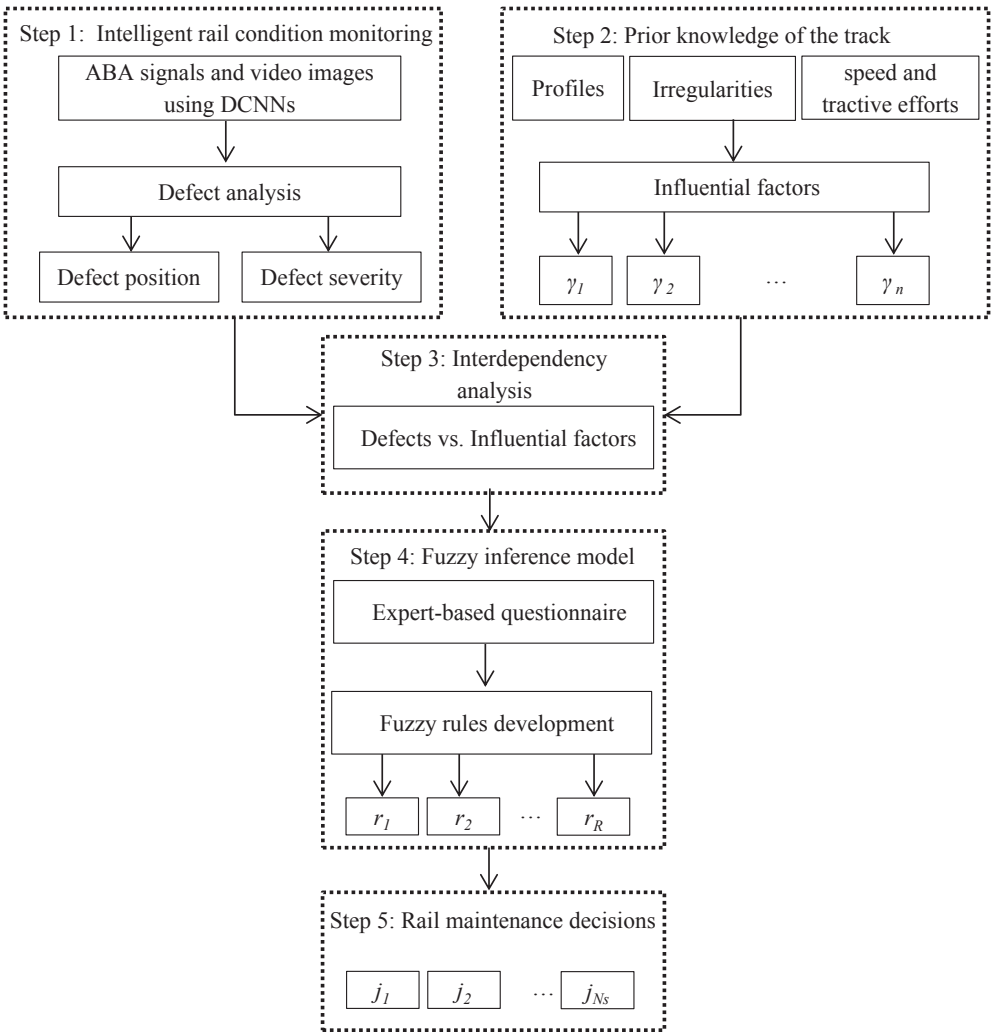

Fig. 1. Flowchart of the proposed methodology.

and also reduces the high cost of track maintenance activities. In this paper, we propose a condition-based maintenance methodology taking both the observations and the prior knowledge of the track into account. The idea is to find the interdependency between the defect status and all major influential factors of the track prior knowledge. The defect status is defined in terms of number and severity of the defects. We investigate the interdependencies between the influential factors and the defect appearance by studying the track characteristics. Once the interdependency is studied, a set of rules is generated to connect rail conditions to the influential factors. The results then indicate which pieces of the rail are prone to be defective. The infrastructure manager is then able to propose maintenance planning according to the critical pieces of railway track. The methodology uses big data analytics, with real-life data measured from a Dutch railway track, using Axle Box Acceleration (ABA) measurements and rail video images (Molodova et al., 2014; Li et al., 2015; Hajizadeh et al., 2014).

Fig. 1 shows the flowchart of the proposed methodology in five steps. The major contribution of the paper is to propose a methodology for rail maintenance decision making that is a combination of new methods and also uses already existing models. Particularly, in Step 5 we make use of the model proposed in Su et al. (2017). Moreover, the proposed methodology is presented in an integrated framework to keep simplicity and coherence between steps. This helps not only to guarantee real-life implementation, but also to keep the infrastructure manager updated of the infrastructure health conditions. In Step 1, the rail defects are detected by using two sources containing the ABA signals and rail video images. A list of critical defects is then provided to represent rail observation. In Step 2, track influential factors, γ_j, are presented to give context on the prior knowledge of the track for each segment j. Step 3 explains the interdependency analysis between the influential factors and the rail observations obtained from Step 1. The aim is to investigate how the influential factors are related to the rail observations.

The analysis of the interdependency between track characteristics and the rail observations is to support expert judgments in order to develop health condition rules as proposed in Step 4. In Step 4, s expert decision system is proposed using an inference system. To do so, a fuzzy approach is used including two steps: (1) a questionnaire filled out by experts and (2) a set of fuzzy health

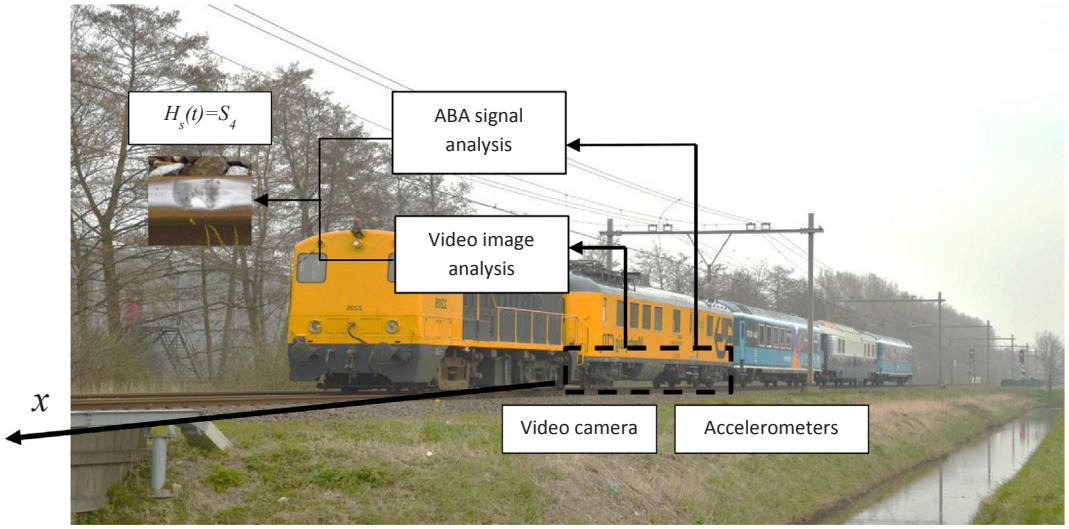

Fig. 2. Defect severity analysis via ABA signals and image data using on-board train measurement. In the scheme, a severe squat is shown (S_4). The actual measurements were obtained from two different trains: the CTO Train for the ABA signals, and the Inspectation measurement train for the video images.

condition rules, r_1, r_2, ...,r_R. The fuzzy rules help to link the influential factors to the health condition. The rules are generated according to expert judgment through a questionnaire. Thus, by generating the rule set, the inference system is built using Mamdani inference model. The aim for the Step 4 is to find the most critical segments that require maintenance among all rail segment. Therefore, the infrastructure manager will have the information of the critical segments. To include operational considerations for the maintenance decisions e.g. time slot limits, logistic concerns, etc., Step 5 is proposed. This results in optimal suggestions for the maintenance decisions. A real-life case from the Dutch railway network is provided to apply the framework and show the applicability of the framework.

2. Step 1: Intelligent rail conditions monitoring

In this paper, we require a technology that can detect defects in an early stage. Hence, we consider to use ABA measurements (Li et al., 2008). To enhance the visualization, ABA measurements are combined with rail image videos (Jamshidi et al., 2017a,b; Faghih-Roohi et al., 2016). In our case study, the ABA measurement and rail video images are used to study rail surface defects; specifically squats, as they are costly for railway networks. A global scheme of the measurement systems is given in Fig. 2.

Li et al. (2015) show the feasibility of early-stage squat detection using an ABA system. The ABA system can be employed to detect a range of surface defects, most importantly, squats, corrugations, and damaged welds, insulated joints, and switches. The ABA system can be embedded in in-service operational trains. Four channels are assigned for the ABA measurement including left rail and right rail, and horizontal and vertical accelerations to capture all the ABA signals.

The image data is collected by a set of high frame rate cameras that are mounted on a specialized measurement train. A top view camera is aimed at the rail surface defects, with each frame covering a length of 15 cm of the track along the longitudinal rolling direction. The recordings are pre-processed into video compilations where consecutive frames have a few millimetres of overlap and the effects of variations in the train speed are removed. As a result, recordings of roughly a thousand kilometre of rail amount to producing hundreds of Gigabytes of video data.

Deep convolutional neural networks (DCNNs) has been applied for different problems in the area of classification due to their algorithmic advantages (Krizhevsky et al., 2012; LeCun et al., 2015). We use a DCNN model in order to automatically estimate from the ABA signals the defect severity throughout the tracks based on a big data analysis. For training the DCNN, based on previous results (Jamshidi et al., 2017a,b; Faghih-Roohi et al., 2016), we obtain a set of labelled images with their severity. The labels used from the images samples are on a scale from 0 to above 4 according to the severity level of the defects visible in the squats found by analysis of rail images. Non-defect track images are assigned a value of zero and defects are assigned from 1 and above. The severity of the squat s can be used to represent the health conditions of the rail, $H_s(t)$, at the time instant t of the measurement as follows:

$$H_s(t) = \begin{cases} S_1 & \text{if } 0 < L_s(t) \leqslant 1 \\ S_2 & \text{if } 1 < L_s(t) \leqslant 2 \\ S_3 & \text{if } 2 < L_s(t) \leqslant 3 \\ S_4 & \text{if } 3 < L_s(t) \leqslant 4 \\ S_5 & \text{if } 4 < L_s(t) \leqslant 5 \end{cases} \tag{1}$$

where $L_s(t)$ is the measured level of severity, S_1 refers to a seed squat, S_2 is a light squat, S_3 is a moderate squat, S_4 is a severe squat,

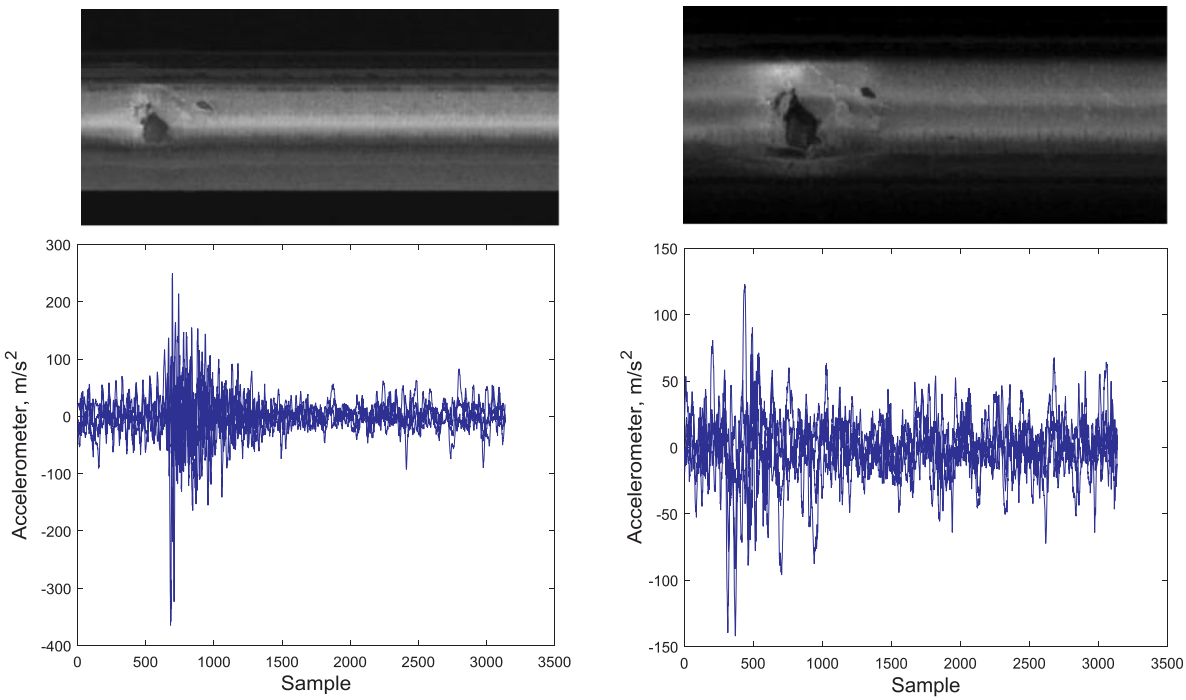

Fig. 3. ABA signals including acceleration matched with rail video frames.

and S_5 is a squat with risk of rail break. The images and their severity are matched with their corresponding ABA signal. To do so, a window of the ABA samples is defined with length of 3036 samples, covering full responses to local defects. This facilitates matching the signals with the video frames. Fig. 3 shows two samples of image data used for the severity analysis associated with the corresponding ABA signals. The labelled data is therefore split into two parts for training and testing. To keep consistency in the defect detection, the labelled samples are collected from different locations over the measured track and they cover all the types of squats. They are compiled into a training set for each of the classes. The dataset was obtained by manual labelling of the images by an expert. The labelled sample defects are then divided into a training set and a testing set. The sample size was 125 squats. The distribution of the squats classes in terms of severity set is 70 samples for S_1, 8 for S_2, 6 for S_3, 8 for S_4, and 33 for S_5. 75% of the data is assigned for training and 25% for validating of the network performance. The samples of the labelled images are composed of 125 different squats collected from different locations of the track. We train a convolutional neural network regression model using the samples. The average binary accuracy (defect vs. non-defect) of the network on all tested samples is taken into account. Although putting a high acceptance threshold on the network output response might increase the rate of false positive detection, we use the threshold to detect the correct classes of the defects, seed (trivial) defects, and the normal classes. Once the DCNN for the image data is trained, defects in the large pool of previously unseen samples can be found.

Using a set of convolutional layers, the defect features are included in the DCNN model as filters to recognize distinguishing features and to create a feature map. A Rectified Linear Unit (ReLU) is used as activation function after the convolution steps, as well as max-pooling layers in order to down-sample the outcome of each layer (Srivastava et al., 2014). The convolutional and pooling layer are finally attached to a sequence of three fully-connected layers to get class predictions (see in Fig. 4).

The separating rail observations (detecting squats using DCNNs) from track characteristics (determined by influential factors) is one of the major contributions of this paper. On the one hand, the DCNN is used to estimate the severity of the defects according to the ABA and image sources. This just gives the defect analysis (the rail observation in the Step 1) and not the rail health condition. On the other hand, track prior knowledge containing the influential factors can impact the rail health condition (Step 2) as those factors affect the quality of rail use over time (rail degradation). Thus, influential factors are collected to contribute the track characteristics for the estimation of rail health condition. For instance, a piece of rail positioned on a rail curve can get degraded faster than the same rail piece on a straight rail. To include track characteristics effects, the interdependency between the rail observations (the DCNNs) and the track prior knowledge is investigated in Step 3.

3. Step 2: Prior knowledge of the track

General characteristics of the railway track system can have a large influence in the initiation and growth of the rail defects. A list of some generic track characteristics that are potentially relevant to the appearance of rail defects are discussed next. The idea in this paper is to take seven factors into account as “general characteristics of track” as according to the literature survey, they are proved to be significantly influencing in the initiation and growth of the rail defects. In particular, we classified the seven influential factors

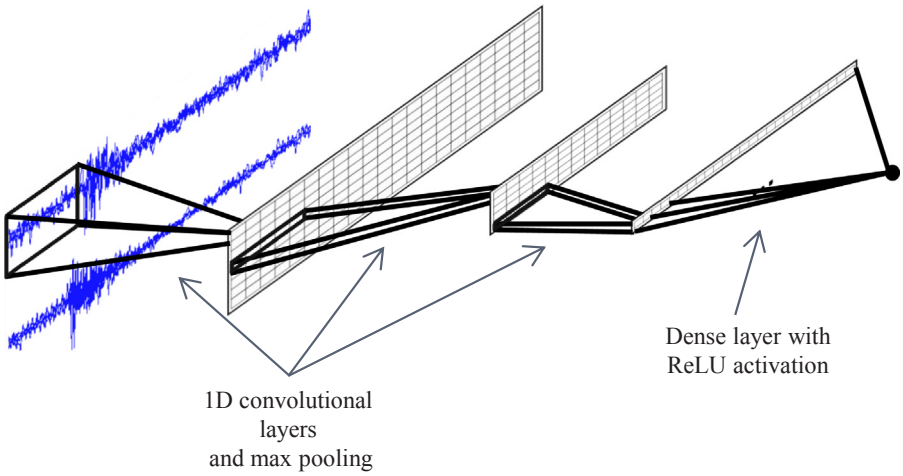

Fig. 4. The prediction model consists of 3 one-dimensional convolutional layers each followed by a max-pooling operation, and a ReLU activated dense layer on top, which results in the final scalar estimation of the severity.

based on Step 2 into three categories: (1) track profiles, (2) track irregularities, and (3) operational speed profile and tractive efforts. However, there are other factors that can affect the track. As an example, train traffic can be influential and has an important role in the actual rail health conditions. In this paper, we assume that the influence of the traffic tonnage, which increases the amount of contact force between wheel-rail, can be seen in the defect severities (the rail observation). Furthermore, tonnage will be an influential factor when predicting defect evolution over time. During the same time period, the rail defects in segments with a higher tonnage evolved faster than the defects in segments with a low tonnage. Additionally, observations indicate that a higher number of defects will be found in tracks with a higher tonnage. These are two possible ways to include the effect of the traffic tonnage in the proposed approach: (1) Indirectly via the effect of the tonnage in the rail observation. Condition monitoring measurements will automatically update both the appearance of new defects and the severity of the defects. (2) Directly via the inclusion of tonnage as influential factor. This case is most suitable when the infrastructure manager wants to predict the evolution of the defects; as tonnage will indicate how fast detected defects will evolve. In this paper, we do not include the prediction of the defect evolution, so in this case the indirect method via rail observation is conducted. Part of the future research is to consider the effect of tonnage within a predictive approach.

We employ various sources of information to obtain the prior knowledge of track using a big data analysis.

3.1. Track profiles

Track profiles are design features. Deviations of the track alignments (vertical, lateral, etc.) with respect to the nominal alignment can lead to track irregularities (Wang et al., 2012; Kawaguchi et al., 2005). Mutton et al. (1991) analyse the wheel-rail contact conditions in the curved and tangent track to investigate the influence of the lateral profile of the track on the initiation and growth of rail defects. Grassie (2012) reviews the research on squats and squat-type defects. The author concludes that squats are associated with straight tracks and gentle curves, but not with tight curves. Likewise, Li et al. (2008) report that squats in the Netherlands occur mainly on straight tracks and gentle curves. On the contrary, head checks occur mostly on the curved tracks of radii less than 3000 m (Li, 2009). In the current paper, the horizontal curvature of the track is taken into account. Furthermore, the rail segments are is defined based on the rail curvature. In this way, only one influential factor for the horizontal curvature is considered for one segment. The vertical profile is ignored as the corresponding changes in the Dutch railway network are small.

3.2. Track irregularities

The track geometry changes from the design geometry due to trains passing over the track. More passing trains could worsen the track geometry conditions. In the literature, the irregularity amplitude and its wavelengths are mostly used as the controlling factors of the track quality. The limits for those controlling factors are typically analysed using measurements and dynamic simulations. The presence of track irregularities was found to have an influential effect on RCF defect appearance (Nielsen et al., 2005). Track geometry problems are widely explained as one of the influential factors considering wheel-rail interactions, maintenance planning, and life of railway tracks. Irregularities have an impact on ride comfort and the traffic safety level. All those influences are therefore very critical in railway dynamics. Nonetheless, the critical level is directly related to track usage. In the literature, there are also different studies about the influence of track geometry on the track conditions and the track degradation. Thus, by considering the significant contribution of the track geometry in the track conditions and then subsequent maintenance plans, control of track irregularities plays an important role on facilitating condition-based maintenance planning (Andrade and Teixeira, 2011, 2012).

Using geometry measurements for the rail maintenance planning is of important considerations for the infrastructure manager

(Veit, 2007; Sharma et al., 2018). A maintained track geometry considerably contributes not only to train safety but also track health conditions. Furthermore, track geometry monitoring could help to prolong the effective track life time by managing the track degradation, the track health conditions, and subsequently the cost of the maintenance operations (Kawaguchi et al., 2005).

The measurement data has been used to develop statistical modelling of railway track irregularities in the last three decades. Track safety and ride comfort are among the first track irregularities analysis using field data. Hamid and Gross (1981) discuss the impact of track quality on track maintenance decisions and performance-based analysis of track geometry using a statistical model for a long track. The paper develops a degradation-based track conditions model to explain interaction between rail defects and performance indicators. A similar investigation has been carried out using linear models to capture the track response to a train load in terms of track irregularities and potential appearance of rail defects (Corbin and Fazio, 1981). Bing and Gross (1983) use a comprehensive model to predict the track quality for maintenance operations. They employed multiple data of traffic and train speed, track structure, and maintenance time slots to predict the track quality over time. In the current paper, based on the available data, we select three sets of irregularity-related influential factors including (1) the vehicle effect, which is a signal indicating the train ride quality based on several geometry measurements and operating trains characteristics, (2) track geometry, which is an indicator estimated based on a combination of different track geometry measurements such as horizontal alignment, the vertical alignment, and cant differences, and (3) track superelevation, which is the difference between the designed cant and the measured cant.

3.3. Operational speed profile and tractive efforts

Tractive effort and curving in the track are found to be potentially responsible for RCF-type rail damages (Grassie and Elkins, 2005). The review of the squat defects by Grassie (2012) reveals that these defects are associated with driving traction i.e. locomotives and power cars. Observations by Li et al. (2011) show the relationship between braking and squat occurrence in the Dutch railway network. The authors conclude that the traction performance of the rolling stock has a large influence on the initiation and growth of squats. They found many squats at pieces of a track where the gradient of the speed was the highest and the speed was low. Moreover, the low speed was also influential, as more frequent activation of the Anti-lock Brake System (ABS) system occurs at lower speeds. Tractive and braking efforts, which differ by the types of locomotives, can also influence the occurrence of RCF defects. A wide range of the Direct Current (DC) or the Alternating Current (AC) drive systems are used in different countries to provide the required power supply of the trains. A comparison is made between AC versus DC locomotives under diverse operational conditions in Australia to investigate the possible development of squats in the rail head (Vo et al., 2015). Scott et al. (2014) find that the most susceptible locations to the squat defects are those where low-speed running occurs with high wheel slip and low adhesion. They investigated the traction characteristics of the typical AC traction motors to find the potential link between the generation of defects and the rolling stock type. In the current paper, the speed profile of the typical rolling stock is investigated to determine its potential correlations with the occurrence of defect. The related effects considered in this paper includes: (1) train speed profile, which is the speed of the measurement train in km/h, (2) train acceleration profile, which is the acceleration of the measurement train in m/s^2, and (3) rail head wear, which estimates the difference between the measured height of the railhead and the nominal height of a new rail railhead in mm. The measurements are obtained with tacho signals, accelerometers, and scanning laser sensors mounted on the measurement train.

4. Step 3: Interdependency analysis

According to the track prior knowledge explained in Step 2, those track factors that are observed to be influential on rail conditions in the Dutch railway network are considered. We use the data available in the Dutch railway infrastructure monitoring system, BBMS (“Branche Breed Monitoring Systeem”), to acquire the signals of the influential factors. In this paper, we use both dynamic and static measurements to obtain the influential factors. After processing the measurements, the influential factors are calculated for a single measurement campaign. Part of the further research includes the use of historical measurements to study the evolution of the influential factors over time. Seven signals are chosen as influential factors that might significantly affect the rail conditions including (1) train speed profile, (2) train acceleration profile, (3) track horizontal curvature, (4) track geometry parameter, (5) rail head wear, (6) vehicle effect, and (7) track superelevation. In Fig. 5, a map is employed to show the track including all the seven influential factors. The data are captured over the whole track to analyse the dynamics of the track influential factors.

Hence, on the one hand we have a set of data over the track representing the track knowledge and on the other hand, squats are detected along the track with their severity and location using the ABA signals and the image data. The interdependency is defined by investigating how to match the location and the severity of a certain defect with the signals of the track influential factors. To do so, the track is partitioned into different segments and the interdependency is investigated per segment.

To numerically represent the severity of a segment, we consider the average of the severities of all the squats that are located in segment j:

$$\aleph_j(t) = \left(\frac{\sum_{s \text{ in segment } j} H_s(t)}{\sum_{s \text{ in segment } j} \delta_s(t)} \right) \tag{2}$$

where $H_s(t)$ is the severity of the squat s provided by the ABA detection algorithm for the measurement time t. The function $\delta_s(t)$ equals 1 when s is a squat, and equals 0 otherwise.

Regarding the processing of the datasets, once all the data sets (signals) over the track are acquired, the signals are processed

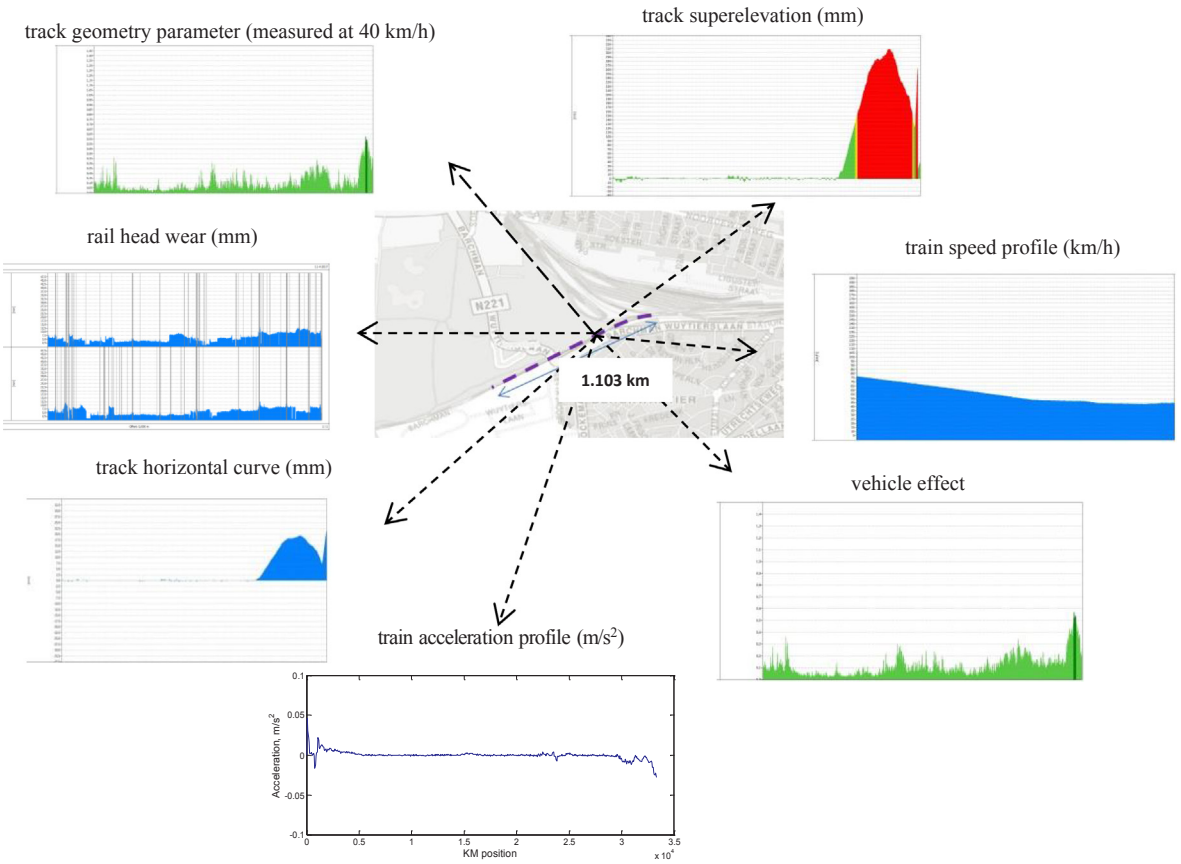

Fig. 5. A schematic representative of the map of the track influential factors in a piece of the track.

according to Eqs. (3) and (4). First, signals are normalized using (3), and then the influential factor is obtained by the average of the signal as in (4). The influential factor is then a "representative" value of the measured signal for that segment. So, the signals should all be normalized between L_{int} and L_{end} which are respectively the upper bound and lower bound of the interval selected for the normalization. The function can be expressed as:

$$\gamma^k_{j,\mathrm{Nor}}(x,\,t) = \frac{(\gamma^k_j(x,\,t) - \gamma^k_{j,\min}(t))(L_{\mathrm{end}} - L_{\mathrm{int}})}{\gamma^k_{j,\max}(t) - \gamma^k_{j,\min}(t)} + L_{\mathrm{int}} \tag{3}$$

where $\gamma^k_j(x,\,t)$ is the data for the *k-th* influential factor at the location x and time instant t, $\gamma^k_{j,\min}(t)$ and $\gamma^k_{j,\max}(t)$ are minimum and maximum values of the signal at the segment j. By considering $x^k_{j,\mathrm{avg}}$ as the location where average value of the data occurs (as a representative of each segment), the data value for the segment j is calculated according to:

$$\gamma^k_j(t) = \gamma^k_{j,Nor}(x^k_{j,avg},\,t) \tag{4}$$

where $\gamma^k_j(t)$ is the influential factor for the segment j and the time step t.

By considering a matrix containing $\gamma^k_j(t) = [(\gamma_1(t))^T,\,\ldots,(\gamma_{N_s}(t))^T]$ where N_s is the total number of segments and $\gamma_j(t) = [\gamma^1_j(t),\,\ldots,\gamma^n_j(t)]$, a clustering model is proposed as follows. We have selected the method called Fuzzy C-Means due to its simplicity. Based on the fuzzy clustering approach, a data point will belong to all the clusters but with a different membership degree. The closer to the centre of the cluster, the membership will be near to one. Points far away from a cluster will have a membership degree near to 0 (Ma et al., 2015). Just for illustration, three clusters are defined over the influential factors. The membership degree to the cluster determines how much a segment belongs to the cluster. The track is partitioned into 15 segments. Fig. 6 shows a schematic view of the clusters. As seen in the figure, segment 5 is highlighted by a rectangle indicating a high membership degree of the cluster 2 in the segment indicated by an arrow. Rail segment 4 has the higher membership to cluster 1; however, it does not belong to the cluster 1 as much as segments 1, 2, 3, 6, 14, and 15, which they all have membership values near to one. The results are used in order to obtain rail health conditions decision rules.

In this paper, five levels are defined including very low (L_1), low (L_2), medium (L_3), high (L_4), and very high (L_5) to represent the interdependency between the defect severity and the influential factors (for simplicity and interpretability of the data, linguistic terms such as very low, low, medium, high and very high are used).

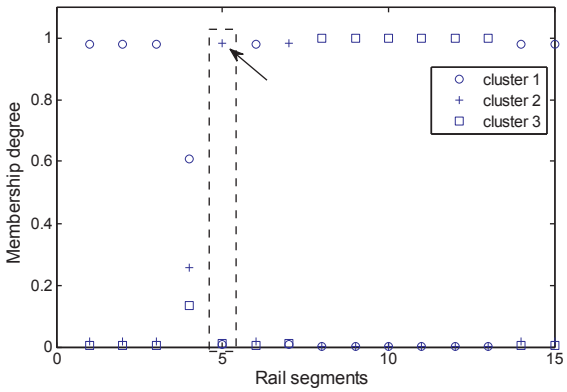

Fig. 6. Membership degree for all the segments based on their influential factors. The highlighted rectangle indicates that rail segment number 5 belongs highly to the cluster 2 (membership degree almost one, indicated by the arrow).

5. Step 4: Fuzzy inference model

In this paper, a fuzzy inference system is used to develop rules about the rail health condition-based on the influential factors $\gamma_j^k(t)$. The Mamdani fuzzy system approach is considered due to its interpretability and simplicity (Camastra et al., 2015; Tosun et al., 2011). To explicitly express the inference system, the Mamdani inference can be defined as follows:

$$Y_j^m(t) = f_{\text{Mamdani}}(\gamma_j^1(t), \gamma_j^2(t), \ldots, \gamma_j^k(t), \ldots, \gamma_j^n(t)) \tag{5}$$

where $Y_j^m(t)$ is the rail health conditions in section j and $\gamma_j^k(t)$ the k influential factor in section j. Fig. 7 shows the architecture of the inference model. In the first layer, the values of input variables, e.g. $\gamma_j^k(t)$ are used. The membership degrees of the inputs to the fuzzy values are obtained in layer 2 and employed to compute the rule truth values in layer 3. At the layer 4, according to the truth values of the various rules, the rail health conditions of each rule in the segment j is estimated.

The R fuzzy if-then rules are generated based on (5) to capture combinations of the influential factors. The purpose is to assign a membership degree to each $\gamma_j^k(t)$. Gaussian membership functions are used to fuzzify the inputs. The Gaussian type of membership function has been used because it is smooth and nonzero at all points (Markowski and Mannan, 2009; Xie, 2003). The Gaussian membership function is based on two parameters and can be represented as:

$$\text{Gaussian}(x;c,\sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \tag{6}$$

where for each membership function, c and σ are the parameters of the membership function. The parameters are tuned so that every membership function has around 30 percent overlapping with the neighboring functions. The rule r_i can be expressed as:

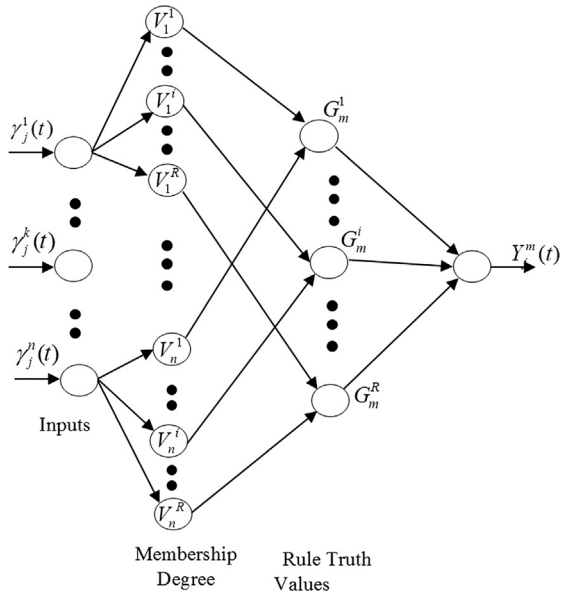

Fig. 7. Generic structure of the fuzzy decision model to compute the rail health conditions.

$$r_i\text{: If } \gamma_j^1(t) \text{ is } V_1^i \text{ and } \cdots \gamma_j^k(t) \text{ is } V_k^i \text{ and } \cdots \text{ and } \gamma_j^n(t) \text{ is } V_n^i \text{ then } Y_j^m(t) \text{ is } G_m^i \tag{7}$$

where V_k^i is the fuzzy set related to input variable $\gamma_j^k(t)$ and G_m^i is the fuzzy set of the rail health conditions selected based on the expert judgment for rule r_i. The minimum of the fuzzified input values is given as the rule truth value of each rule as seen in Fig. 7. The fuzzy set of the output is obtained by the Mandami union operator over all the rules. To defuzzify the output, the centre of gravity approach is applied so as to obtain a crisp value. The fuzzy inference system (Mamdani) is to map the inputs (the influential factors) to the output (the rail health condition) using a set of fuzzy rules. Thus, the fuzzy rules are components of the fuzzy inference system. To set the fuzzy rules, a questionnaire is provided to systematically analyse the combinations of possible inputs. As the judgment relies on the expert knowledge, it is prone to bias. Thus, the investigation is used to support the experts on the validation of the judgements. The inclusion of the investigation results in the questionnaire, helps the expert to visualize the effect of $\gamma_j^k(t)$ over the segment j on the actual rail health conditions. Furthermore, as the questionnaire will lead to a model of the rail condition using the knowledge of expert, the expert qualified to fill out the questionnaire is a rail maintenance engineer or a rail inspection expert. The expert should have experience with both rail monitoring and rail maintenance. By using the proposed methodology, the infrastructure management company will benefit from systematically keeping the knowledge of rail experts in the company. So, in case a rail expert is not available, the railway company can still use the previously developed rules or update them according to new infrastructure requirements. In the questionnaire, two options are given including “influential” “non-influential”. Then, the experts are asked to rank between 1 and 2 the effect of the combination of influential factors into the health conditions of the rail. A major contribution of the fuzzy system is to include non-crisp values (fuzzy values) in the output (the rail health condition). Although a binary approach is used for the questionnaire, (1) we can capture the fuzzy dynamics on the rail health condition and (2) we cover all the rule combinations. Otherwise and with using five-level ranking, number of the rules created would be too much time consuming for the experts whereas some of those rules would be useless in the decision making. Moreover, the five-level ranking is used to just improve the visualization quality of the interdependency analysis. The questionnaire is converted into a fuzzy inference system, where the rules are given by the options of the questionnaire (two possible fuzzy sets per influential factor) and the output fuzzy sets of each rule are given by the answers of the experts (three possible fuzzy sets).

6. Step 5: Rail maintenance decisions

After estimating the rail health conditions for each segment, the entire rail can be evaluated according to the estimated health conditions. The aim is to find the most critical pieces of the track for the condition-based planning of grinding operations. Squats can be treated by grinding completely when they are at an early stage of growth or they can effectively be kept at safe level (to avoid having disastrous consequences) when they are severe. In this paper, a clustering method is proposed to grind the most critical pieces of the track efficiently based on predefined maintenance time slots determined by the infrastructure manager. As different tracks have different maintenance time slots, it is important to consider the available time slots to carry out the grinding operation. In the Dutch railway network, the time slots vary from one railway station to the next railway station. This means that not all segments of a long track that include different railway stations have the same maintenance time available for doing grinding. The grinding planning is formulated as in Fig. 8.

As depicted in Fig. 8, if maintenance time is still available after the grinding, the clustering approach can be applied to the other critical track pieces to effectively utilize the whole available maintenance time slot. Hence, the infrastructure manager makes sure that the maintenance time is fully used to avoid inducing extra maintenance costs.

The clustering approach strives to cover as many severe defects using as few clusters as possible within the limited maintenance

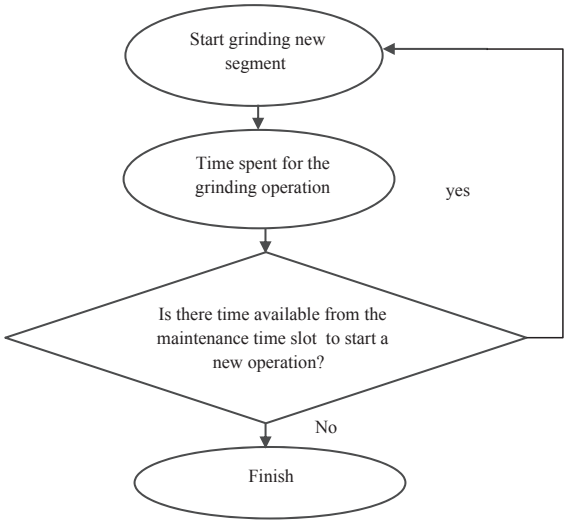

Fig. 8. The proposed simplified grinding planning scheme.

Table 1
A list of the notations of the clustering model.

Notations	Definitions
N_c	Number of clusters
N_d	Number of squats
$\underline{\xi}$,$\bar{\xi}$	Track starting and ending positions, respectively
ω	Defect severity
T_s	The setup time for a grinding maintenance operation.
T_t	Duration of maintenance time slot
d_g^{start} and d_g^{end}	Track starting and ending positions of the g-th cluster
Δd_{min} and Δd_{max}	Minimum and maximum size of each cluster
v_G^{on}	Grinding machine speed running over the track
v_G^{off}	Driving speed of the grinding machine when the machine is off
T_G^{on}	The time needed to switch from grinding to driving
T_G^{off}	The time needed to switch from driving to grinding

time slot, which usually takes 8–10 h at night in the Dutch railway network (this depends on the type of operations, and it could change per day, week, and year). The proposed clustering approach assigns a defect, e.g. a squat, to a cluster. The model includes not only the squat position, but also the squat severities acquired by the ABA system measurement and rail image data. The proposed grinding model is elaborated in the previous work of the authors (Su et al., 2017). Table 1 presents the notations used in the model.

We call $[\underline{\xi}, \bar{\xi}]$ the physical range, and clusters located within the physical range are called active clusters. Also, the setup time, T_s, typically includes the machine travelling time, preparation time and finishing time for a maintenance operation. The start and end locations of the *g*-th cluster are the decision variables of the clustering problem. Thus, the grinding model can be formulated as the following non-smooth optimization problem (Su et al., 2017):

$$\max_{\{d_g^{start}, d_g^{end}\}_{g=1}^{N_c}} \sum_{g=1}^{N_c} \sum_{l=1}^{N_d} \omega_l I_{d_g^{start} \leqslant X_l \leqslant d_g^{end}} + \sum_{g=1}^{N_c} I_{d_g^{end} > \bar{\xi}} \tag{8}$$

subject to

$$I_{d_g^{start} \leqslant X_l \leqslant d_g^{end}} = \begin{cases} 1 & \text{if } d_g^{start} \leqslant X_l \leqslant d_g^{end} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$I_{d_g^{end} > \bar{\xi}} = \begin{cases} 1 & \text{if } d_g^{end} > \bar{\xi} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

$$d_1^{start} \geqslant \underline{\xi} \tag{11}$$

$$d_{N_c}^{end} \leqslant \bar{\xi} + 2N_c(\Delta d_{min} + \varepsilon) \tag{12}$$

$$\Delta d_{min} \leqslant d_g^{end} - d_g^{start} \leqslant \Delta d_{max} \quad \forall\, g \in \{1, \ldots, N_c\} \tag{13}$$

$$d_{g+1}^{start} - d_g^{end} \geqslant \varepsilon \quad \forall\, g \in \{1, \ldots, N_c - 1\} \tag{14}$$

$$d_g^{start} \leqslant \bar{\xi} \;\Rightarrow\; d_g^{end} \leqslant \bar{\xi} \quad \forall\, g \in \{1, \ldots, N_c\} \tag{15}$$

$$\sum_{g=1}^{N_c} I_{d_g^{end} \leqslant \bar{\xi}} \cdot \left(\frac{d_g^{end} - d_g^{start}}{v_G^{on}} + T_G^{on} + T_G^{off} \right) + \sum_{g=1}^{N_c - 1} I_{d_g^{end} \leqslant \bar{\xi}} \cdot \frac{d_{g+1}^{start} - d_g^{end}}{v_G^{off}} \leqslant T_t - T_s \tag{16}$$

The indicator function I_σ takes value 1 if the statement σ is true, and 0 otherwise. The first term in the objective function (8) rewards the squats covered by a cluster depending on their severities, while the second term serves to maximize the number of non-active clusters, i.e. minimize the number of active clusters. The second term in (8) counts the number of clusters outside the physical range, i.e. non-active squats. As the total number of available clusters *Nc* is fixed, maximizing the number of non-active clusters is equivalent to minimizing the number of active clusters. The active cluster is defined via the kilometre positions of its start and end points within the physical range $[\underline{\xi}, \bar{\xi}]$. A non-active cluster is outside the physical range and has no physical meaning. We use the idea of non-active cluster to be able to have idle clusters. Also, X_l is the kilometre location of the *l*-th squat. Constraints (11), (12) set the distance range of the clusters. Note that the upper bound ξ_{max} is set as indicated to allow the situation of non-active cluster, i.e. all clusters are located outside the physical range $[\underline{\xi}, \bar{\xi}]$. The term ε in (12) is included to avoid the overlapping of clusters. To determine ε, we suggest just to take a tiny positive value (like $\varepsilon = 0.001$ m). When ε is high, the distance between clusters will be higher and interesting rail pieces might not get covered by a cluster. Constraint (13) restricts the size of each cluster. The minimal and maximal size of a cluster is indeed determined by operational considerations of the grinding machine. The minimal size of a cluster is usually

Fig. 9. Schematic track map between two stations, Amersfoort and Weert.

set to be the shortest length that the grinding machine can manipulate. The maximal size of a cluster should be less than the length of the rail considered. Constraint (13) restricts the size of each cluster. Constraint (14) ensures that the clusters are not overlapping, where the small positive parameter ε is the minimum distance between two clusters. So, there may be track sections between clusters that will be not included in the grinding planning. The constraint (15) forbids fractional clusters. The fractional cluster means that the start and end points of a cluster must both be inside or outside the physical range. We only allow to use active clusters (start and end points are both inside the physical range) and non-active clusters (start and end points both outside the physical range). The constraint (16) is the time limit constraint to ensure that the resulting clusters can be processed within the given maintenance time slot. The left-hand side of constraint (16) computes the total maintenance time, including the time to grind the active clusters (first term), the time for the machine to travel between the clusters (second term), and the setup time T_s. Constraint (16) guarantees that the total maintenance time to execute the clustering plan is less than the duration of the maintenance time slot T_s. The non-smooth optimization problem (8)–(16) can either be solved by gradient-free algorithms like pattern search and genetic algorithms, or transformed into an MILP problem following the standard procedure described in Bemporad and Morari (1999). In Su et al. (2017), the clustering method was employed as part of the low-level optimization, in a setup where the decisions are based on prediction including uncertainties via a scenario-based chance-constrained approach.

7. Numerical results

The track Amersfoort-Weert in the Netherlands is selected as a case study (nearly 125 km of track). The track passes through Utrecht, Geldermalsen, 's-Hertogenbosch, and Eindhoven to reach the destination (Weert) (Fig. 9). The whole track is partitioned into 15 segments to take all the signals of the influential factors per segment into account. Also, the definition of the segments is based on track curvature, which means that each curve is included into one segment regardless the segment sizes.

The squat problem is aimed in the case study due to the fact that: (1) squats are one of the most commonly observed defects on rails, (2) squat-related costs are more than 5000 euro/km per year in The Netherlands. Although the rail grinding helps to treat all type of rail defects, e.g. corrugation, head checks and wheel burn, the optimal maintenance decisions proposed in the current paper focus on the squat problem and for the other rail defect types, it is crucial to take the effect of those defects in the maintenance decisions into account. For the estimation of the actual rail conditions, as explained in Section 2, the images are analysed using image processing to detect the ones including squats.

The rail image analysis is defined based on the input images that are down-scaled to 375 × 275 pixels and converted into gray-scale images. The sequence of three fully-connected layers translates the extracted high-level features from the previous layers into 3 classes representing the normal rail, trivial defects (seed squats), and squats. The normal class includes all the components in a healthy state, including plain rails, switches, welds, possible non-defect contaminations, etc. Trivial defects appear in the form of indentations or small damages to the rail head, while squats are usually defects that are fully grown deformations cracking the rail surface. The overlap between different rail images can cause mismatch between rail images and the ABA signal and might affect the estimation of the rail health conditions. To avoid this, first the video frames are pre-processed and the overlaps are removed. Then, we align images with the ABA signal using GPS tags and different reference points of the rail infrastructure (such as switches, crossings, joints, etc.). Fig. 10 shows the mean absolute error of the detection algorithm as a function of the training epoch of the network for both training and validation data. The 75% of the data is assigned for the training and 25% is to validate the network performance. The samples of the labelled images are composed of 125 different squats collected from different locations of the track. Fig. 11 shows the comparative predictions and the ground truth values for all samples in the test set. Thus, although the number of samples is limited, as the samples were picked up from different locations and vary from light to severe squats, one can argue that the

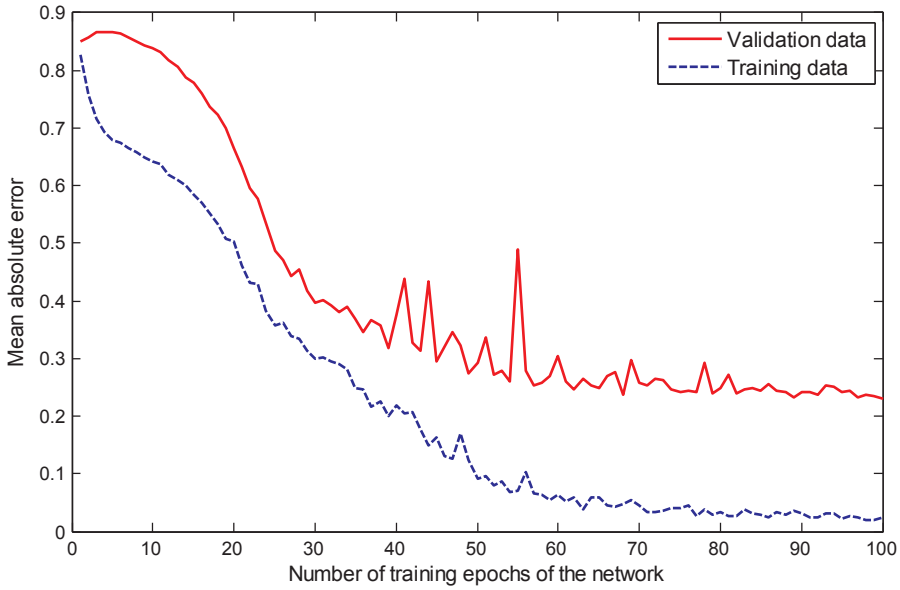

Fig. 10. Mean absolute error (MAE) between the ground truth severities and the predictions. The network is trained using 75% of the data and is validated on the remaining 25%.

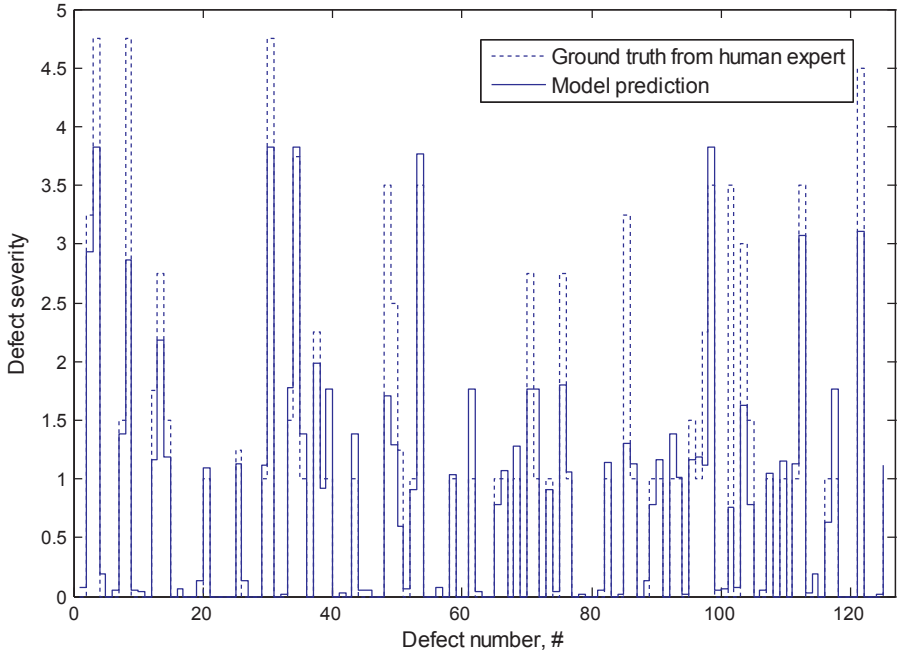

Fig. 11. Ground truth values provided by a human expert by estimating the defect severities from defect images vs prediction of the severity level from the ABA signal.

dataset covers all the interesting cases.

Finally, the trained model is used with the new samples provided from the target track and predictions based purely on ABA are calculated. Fig. 12 shows a sample plot of the results by the detection algorithm, which are used as the rail actual health conditions, and shows the position of the defects and their severity.

The time needed for training is 40 h per 1500 examples. Once the network is trained, it is used to find squats in the large pool of previously unseen samples (prediction). Unlike the training time, the prediction time is insignificant (30 s per 15,000 examples). The prediction result then has an average binary accuracy (squat vs. normal) of 96.9%. The detected squats are then analysed in terms of

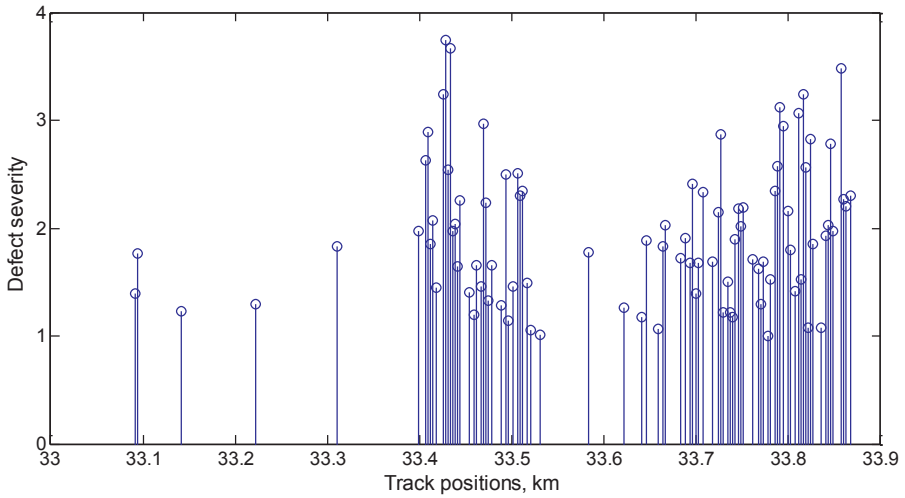

Fig. 12. A sample of defect locations versus defect severity between kilometre 33 and 33.9 in the track Amersfoort-Weert.

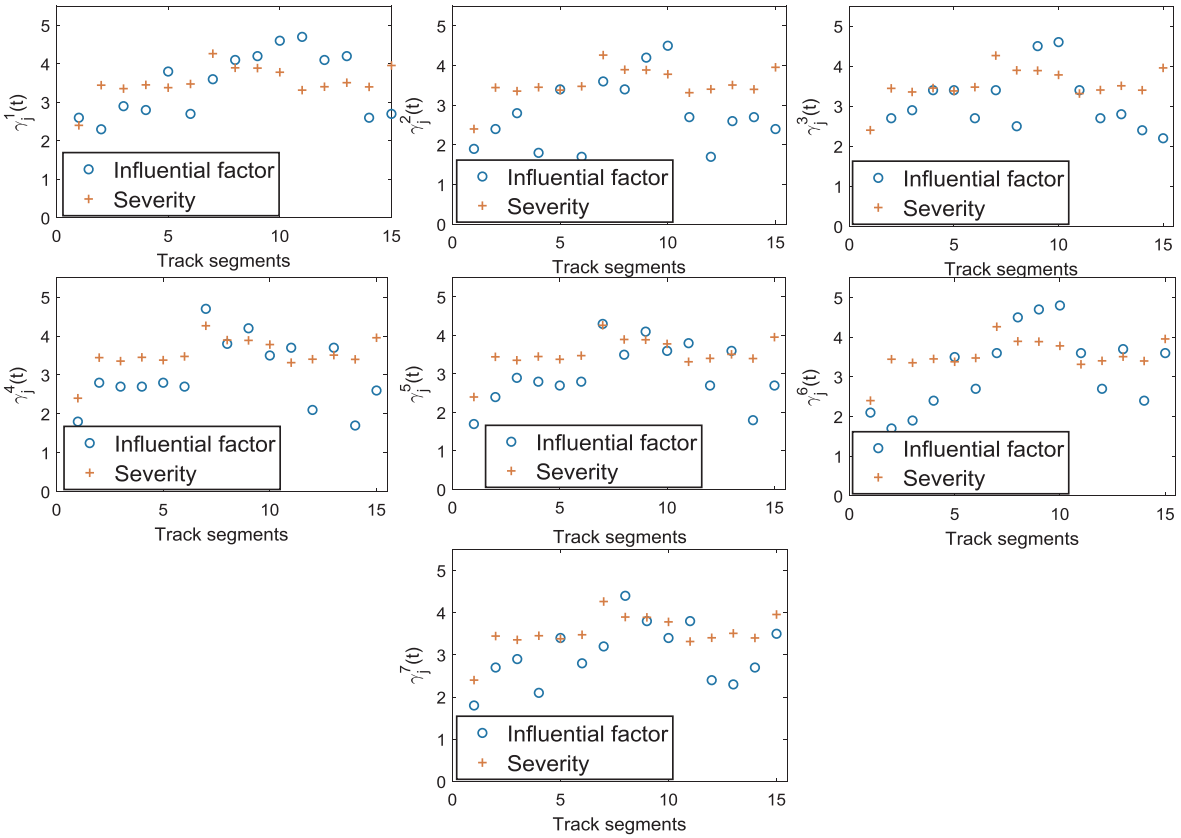

Fig. 13. interdependency analysis between defect severity and track influential factors over 15 track segments where $\gamma_j^1(t)$, $\gamma_j^2(t)$, $\gamma_j^3(t)$, $\gamma_j^4(t)$, $\gamma_j^5(t)$, $\gamma_j^6(t)$ and $\gamma_j^7(t)$ are the train speed profile (m/s), train acceleration profile (m/s^2), track horizontal curve (mm), track geometry parameter (measured at 40 km/h), rail head wear (mm), vehicle effect, and track superelevation (mm), respectively.

the severity according to Step 2.

In Fig. 13, to perform the interdependency analysis we have compared the defect severities within a segment with each influential factor (track characteristic). This information can be used to guide the design of fuzzy rules created from interviews with experts about the relation between health conditions and influential factors. Based on the interdependencies, a set of fuzzy rules is defined to estimate the health conditions based on the influential factors as obtained in (3). All the rules are given the same weight. Moreover,

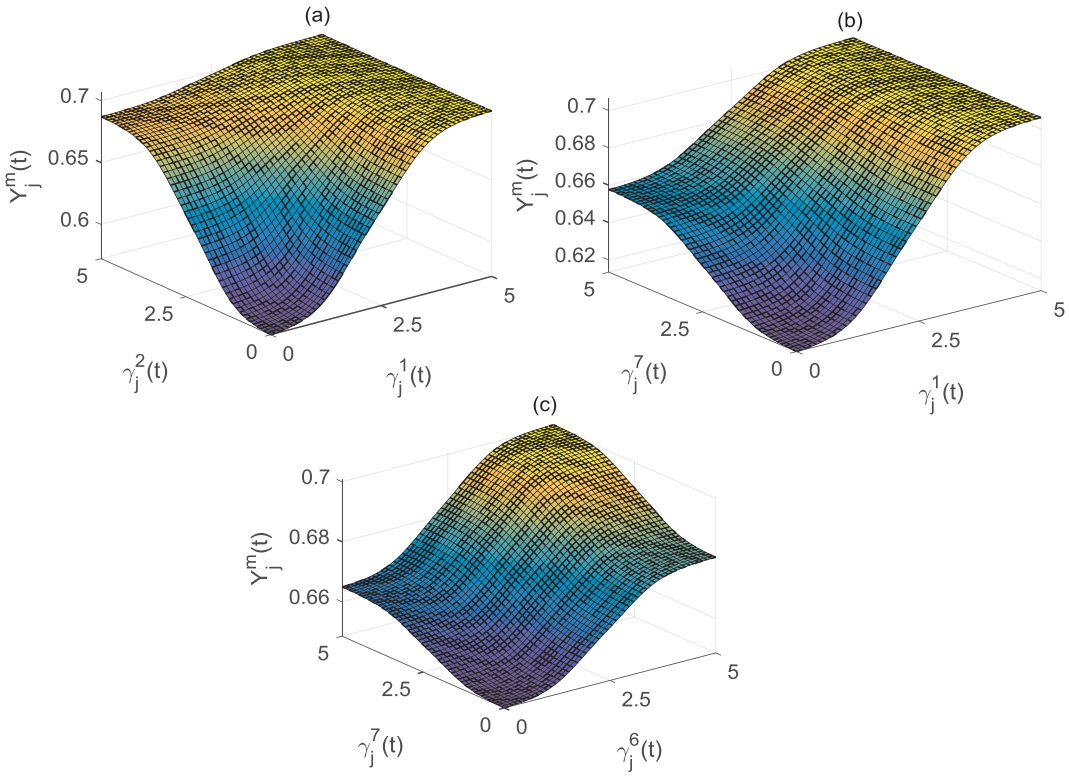

Fig. 14. Examples of how the fuzzy rail conditions rules are defined according to the interdependency analysis.

all the input variables are combined through the rules. In this paper, 127 fuzzy if-then rules are generated in order to meet the possible interdependencies. Furthermore, based on the fuzzy rules, the sensitivities of the health conditions to the influential factors are captured as shown in Fig. 14. This figure presents how the influential factors model the rail health conditions, varying from fully healthy (severity equal to zero) to completely unhealthy (severity equal to one), while all the other influential factors are assumed to be fully healthy (equal to zero). Three plots are used to show the sensitivity.

Variation of the inputs of an expert in the questionnaire can lead to different final maintenance decision results. Several experts are asked to fill out the questionnaire so that variations cause by single expert are reduced. Among all the influential factors, train speed has the highest effect on the grinding decision and superelevation has less influence. An increase of 20% in the train speed related influential factor gives an 8% increase on the rail health condition, whereas an increase of 20% in the superelevation related influential factor gives 5%. A misestimation of 20% in a single factor gives at most 8% difference in final results error in the case of changing train speed related influential factor with superelevation related influential factor.

As an example in Fig. 14(a), the effect of two input variables namely train speed, $\gamma_j^1(t)$, and train acceleration profile, $\gamma_j^2(t)$, respectively, is presented. As shown in the figure, the train speed changes over the track affect the rail health conditions more in comparison with the train acceleration profile. This is an indication of the importance of the train speed for maintenance decisions. Fig. 14(b) depicts the influence of the speed profile versus the superelevation, $\gamma_j^7(t)$. The plot shows that the rail health conditions cannot get excited by the influence of the superelevation as much as the effect of the train speed profile. In Fig. 14(c), the vehicle effect, $\gamma_j^6(t)$, is compared with the superelevation $\gamma_j^7(t)$. As can be seen in the plot, the both factors are not as influential as the train speed and the train acceleration on the rail health conditions. However, the vehicle effect can influence the health conditions more in comparison with the superelevation. Therefore, an increase in the most influencing factors, i.e. $\gamma_j^1(t)$ and $\gamma_j^2(t)$ can increase the criticality of the segment up to requiring maintenance. If this criticality goes beyond the given rail health conditions of other rail segments of the track, then the grinding decision changes directly. Therefore, the infrastructure manager should take the segments with higher train speed profile and train acceleration into account in the maintenance plan.

Relying on the fuzzy model, the rail health conditions is estimated. Each segment is evaluated based on the health conditions as shown in Fig. 15 and Table 2. Table 2 presents the results of the case study. Given the influential factors, the rail health conditions based on the fuzzy inference system is estimated. Although some rules might not be needed as they might never apply in practice, we aimed to develop a questionnaire that captures all the possibilities to have a full coverage of inputs. Using the proposed inference system, any rail segment can be evaluated with given influential factors. Table 2 gives an example on how the fuzzy inference system performs. The influential factors are obtained from rail field measurements and the last column is calculated using the fuzzy rules.

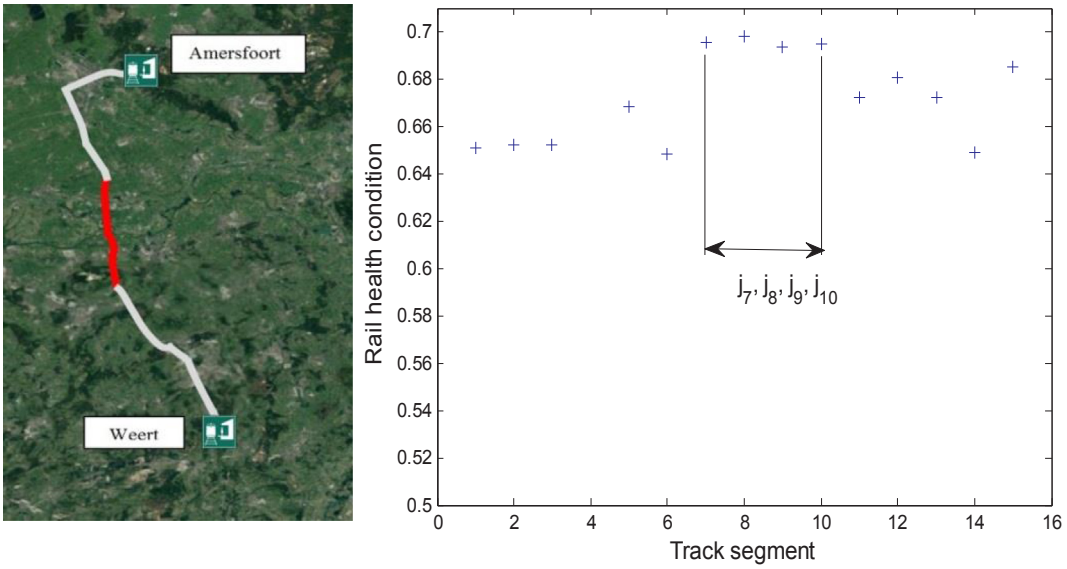

Fig. 15. Rail health conditions over the track segment showing the most critical pieces of the track highlighted with a red line on the map. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Calculated influential factors per segment and estimated rail health conditions using the proposed fuzzy inference system.

Segments	$\gamma_j^1(t)$	$\gamma_j^2(t)$	$\gamma_j^3(t)$	$\gamma_j^4(t)$	$\gamma_j^5(t)$	$\gamma_j^6(t)$	$\gamma_j^7(t)$	$Y_j^m(t)$
1	2.6	1.9	1.1	1.8	1.7	2.1	1.8	0.6507
2	2.3	2.4	2.7	2.8	2.4	1.7	2.7	0.6522
3	2.9	2.8	2.9	2.7	2.9	1.9	2.9	0.6521
4	2.8	1.8	3.4	2.7	2.8	2.4	2.1	0.4929
5	3.8	3.4	3.4	2.8	2.7	3.5	3.4	0.6683
6	2.7	1.7	2.7	2.7	2.8	2.7	2.8	0.6481
7	3.6	3.6	3.4	4.7	4.3	3.6	3.2	0.6957
8	4.1	3.4	2.5	3.8	3.5	4.5	4.4	0.6982
9	4.2	4.2	4.5	4.2	4.1	4.7	3.8	0.6938
10	4.6	4.5	4.6	3.5	3.6	4.8	3.4	0.6949
11	4.7	2.7	3.4	3.7	3.8	3.6	3.8	0.6721
12	4.1	1.7	2.7	2.1	2.7	2.7	2.4	0.6803
13	4.2	2.6	2.8	3.7	3.6	3.7	2.3	0.6721
14	2.6	2.7	2.4	1.7	1.8	2.4	2.7	0.6489
15	2.7	2.4	2.2	2.6	2.7	3.6	3.5	0.6849

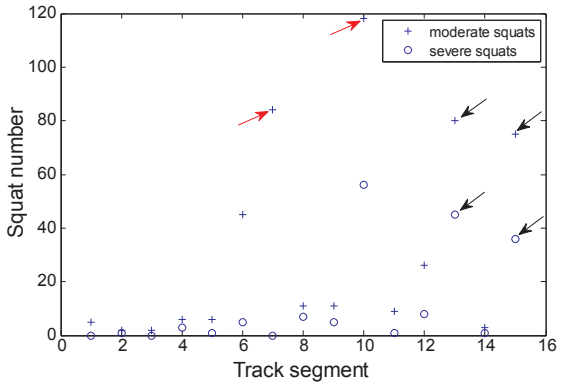

Fig. 16. Number of moderate and severe squats within different segments. Red arrows show the target segments including the most severe squats and black arrows indicate the alternative options for grinding in case the enough time remains within the maintenance time a lot after maintaining the target segments. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

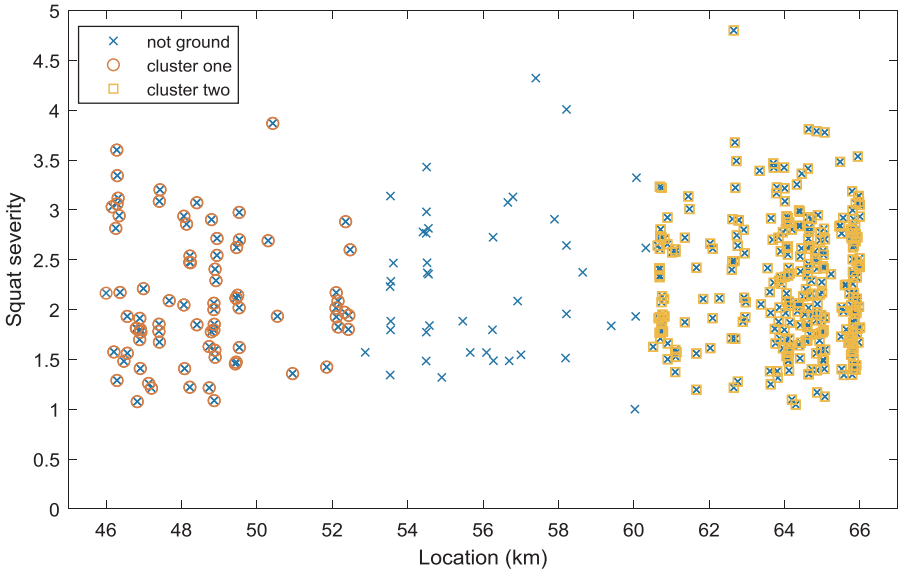

Fig. 17. Result of the grinding model that determines the optimal clustering of the target track. Squats are marked by either a colored square or a colored circle if they are covered by a cluster.

Fig. 15 also indicates that j_7, j_8, j_9, j_{10} obtain the highest values of the rail health conditions. It means that those segments have a critical health conditions compared to other rail segments. These segments highlighted by the red line in the figure belong to the track between railway stations Geldermalsen and 's Hertogenbosch. Furthermore, the rail actual conditions (rail observation) is depicted in Fig. 15. The figure shows the number of squats over the full track from Amersfoort to Weert. The defects are detected based on the proposed detection model described in Section 2.

The segments with the most severe squats are distinguished by two different arrows in Fig. 16. As seen in the figure, the segments 7 and 10 include the highest number of squats. Thus, the segments 7, 8, 9 and 10 are selected as the critical segments to be maintained. Depending on the available maintenance time slots, the track can be ground. If after grinding the above-mentioned segments, there is time to perform maintenance in the rest of the rail network, the segments 13 and 15 are candidates to be maintained (marked by black arrows). From the figure, j_{13} has more squats than j_{15}, but according to Fig. 15, j_{13} is more critical in terms of the health conditions and also it is shorter in length, which increases the squat density distributed on the track. Therefore, segment 15 is chosen as the alternative option.

Once the critical segments are determined, the optimal clustering model is used to cover squats subject to the time limit imposed by the maintenance slot. The proposed clustering model is able to treat the most important squats. In this paper, the length of the maintenance time slot is set to 8 h. This 8 h time period includes the setup time which covers the time required for transportation, machinery, personnel, etc. The most relevant squats are covered by a cluster, as the clustering model penalizes a squat outside any cluster by its severity. Hence, the most important squats are treated by grinding, even when the maintenance slot is not long enough.

This is not normally the case for cyclic grinding, which is currently the most used method in the Dutch network in which the full track is ground. For cyclic grinding, the grinding machine starts grinding from the start point going towards to the end of the track without any guarantee to capture the most important squats. Fig. 17 shows the clustering result between the stations Geldermalsen and 's Hertogenbosch covering the critical segments, i.e. j_7, j_8, j_9, j_{10}. The target track is around 20 km as shown in the x-axis of the figure. According to the proposed detection model, the squat severity is estimated as indicated in the y-axis. The grinding model proposes two clusters within the maintenance time slot capturing the most severe squats by considering the density of the squats. In this way, the grinding machine starts grinding from the beginning of the track to reach the 52.42 km, then the machine stops working to drive to cluster two (the transfer time is supposed negligible), which starts at the track position 60.51 km until the end of the track. Moreover, the number of the defects between (remaining track piece) 52.46 km and 60.51 km is much less (43 defects and average severity 2.10) than in the first cluster between 46 km until 52.4 km (77 defects and average severity 2.15) and the second cluster between 60.51 km until 66 km (187 defects and average severity 2.25). Thus, although we have defects between 52.46 and 60.51, by considering (1) the maintenance time slot limitation and (2) maintenance priority of the segments 7 and 10 (Fig. 17) in terms of higher value of health condition, the defects between 52.42 and 60.51 remain with no maintenance intervention until the next maintenance time slot. Without the proposed clustering model, the grinding machine will not be able to capture the most important squats, either at the beginning of the track or at the end of the track. Some severe squats would therefore remain untreated, which would increase maintenance costs and the probability of rail failure.

The cost to employ the grinding machine is 35 k euro for one night considering 10 h. Note that 10 h is fixed meaning that for

shorter maintenance time slots (shorter than 10 h) the cost is the same. Thus, the infrastructure manager will be charged the same amount of money, although the machine is used for less than 10 h. Thus, in case there would be 2 h extra time available after finishing the grinding of the critical rail pieces, the infrastructure manager has the chance to fill all the available time to keep the grinding machine running. In that case, according to the proposed methodology, the grinding machine can be transferred to the segment 10, j_{10}, which has a more critical health conditions compared to the rest of the target track. Then, the infrastructure manager can ensure that traffic-free hours have been used to treat all the most important squats over a long track.

8. Conclusions

In this paper, we have proposed an integrated approach for maintenance decision system of the railway infrastructures. The methodology includes infrastructure conditions monitoring and maintenance decision making. The proposed approach is applied to the condition-based treatment of squats, with big data information coming from a track in the Dutch railway network. The algorithm makes use of both axle box acceleration signals and rail video images, which contribute a huge amount of data. The use of both rail data sources reduces the detection error of the surface defects. Moreover, we have used the track characteristics of the Dutch railway network, enabling the infrastructure manager to interconnect the track influential factors with the actual rail health conditions. We therefore investigated how to define a list of decision actions to support the decisions regarding the maintenance plan by analysing the above-mentioned interdependency. The results propose a maintenance decision approach based on the actual conditions of the rails but together with the insights resulting from the influential factors. We proposed a partitioning in 15 different segments for a track that can be considered quite long (105 km). The maintenance decision system is proposed using a clustering model to perform grinding over the critical pieces of the rail. The results include the most severe squats covered by the maintenance clusters. Thus, although not all the squats are treated, the infrastructure manager can make sure that there is considerably less safety risk or high maintenance cost until the next rail measurement campaign. To include possible practical limitations. Then, we include the maintenance time slot as a constraint problem. Different pass numbers of the grinding machine, resulting in different grinding depths, have an impact on the rail defect risk after grinding. Different pass numbers also lead to different grinding speeds. The current clustering model considers only one grinding depth, meaning one fixed pass number and grinding speed.

While this paper is focused on the analysis of squats, the results are applicable to the analysis of other types of rail defects like corrugations, damaged insulated joints, welds and other types of RCF defects. To apply the proposed methodology to all those defects, the infrastructure manager will need to analyse the rail observations in terms of that specific type of defect versus the track characteristics to define the list of decision rules. The methodology for the design of the rules is flexible, so they can be adapted to different railway networks. In the further research the interdependency analysis can be conducted at a more detailed level, for instance at every kilometre or even at meter of track. In future research, based on the influential factors it will be possible to anticipate the rail conditions much better, so predictive maintenance could be achieved. The maintenance operations could be different from one type of defect to another, but the general methodology can be adapted, as far as the defects can be grouped into different clusters. In addition, the proposed methodology can be linked to a rail maintenance cost analysis to reduce life cycle cost (LCC). Also, by having different measurement sets of rail data, a prediction model of how the defects can grow over time could be added to the methodology, correlated to the influential factors. This will help the infrastructure manager to predict the rail health conditions in advance and also to prolong the maintenance decision time horizon. In the future work, we will consider a flexible number of passes of the grinding machine to obtain more efficient clustering plans. Another topic for further research is to evaluate the methodology for different regions to investigate the influence of exogenous factors like environmental factors to the decision rules and consequently the maintenance decision rules.

Acknowledgements

This research is part of the NWO/ProRail project “Multi-party risk management and key performance indicator design at the whole system level (PYRAMIDS)”, project code 438-12-300, and the STW/ProRail project “Advanced monitoring of intelligent rail infrastructure (ADMIRE)”, project 12235, which are partly funded by the Ministry of Economic Affairs. The authors also would like to thank INSPECTATION for providing us with image data and technical support.

Appendix A

To generate the fuzzy rules, a questionnaire was filled out by an expert. In the questionnaire we asked to use linguistic terms e.g. non-influential = 0 and influential = 1 for the factors (column 1 to column 7). Then a score between 0 and 2 is given to the rail health conditions (the last column) considering the combination of situations from the influential factors. Healthy = 0, Average = 1, Unhealthy = 2 (see Table 3).

Table 3
The questionnaire.

$\gamma_j^1(t)$	$\gamma_j^2(t)$	$\gamma_j^3(t)$	$\gamma_j^4(t)$	$\gamma_j^5(t)$	$\gamma_j^6(t)$	$\gamma_j^7(t)$	$Y_j^m(t)$	$\gamma_j^1(t)$	$\gamma_j^2(t)$	$\gamma_j^3(t)$	$\gamma_j^4(t)$	$\gamma_j^5(t)$	$\gamma_j^6(t)$	$\gamma_j^7(t)$	$Y_j^m(t)$
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2
0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	2
0	0	0	0	0	1	0	1	1	0	0	0	0	0	1	2
0	0	0	0	0	1	1	1	1	0	0	0	0	1	0	2
0	0	0	0	1	0	0	1	1	0	0	0	0	1	1	2
0	0	0	0	1	0	1	1	1	0	0	0	1	0	0	2
0	0	0	0	1	1	0	1	1	0	0	0	1	0	1	2
0	0	0	0	1	1	1	2	1	0	0	0	1	1	0	2
0	0	0	1	0	0	0	1	1	0	0	0	1	1	1	2
0	0	0	1	0	0	1	1	1	0	0	1	0	0	0	2
0	0	0	1	0	1	0	2	1	0	0	1	0	0	1	2
0	0	0	1	0	1	1	2	1	0	0	1	0	1	0	2
0	0	0	1	1	0	0	1	1	0	0	1	0	1	1	2
0	0	0	1	1	0	1	1	1	0	0	1	1	0	0	2
0	0	0	1	1	1	0	2	1	0	0	1	1	0	1	2
0	0	0	1	1	1	1	2	1	0	0	1	1	1	0	2
0	0	1	0	0	0	0	1	1	0	0	1	1	1	1	2
0	0	1	0	0	0	1	1	1	0	1	0	0	0	0	2
0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	2
0	0	1	0	0	1	1	2	1	0	1	0	0	1	0	2
0	0	1	0	1	0	0	1	1	0	1	0	0	1	1	2
0	0	1	0	1	0	1	2	1	0	1	0	1	0	0	2
0	0	1	0	1	1	0	2	1	0	1	0	1	0	1	2
0	0	1	0	1	1	1	2	1	0	1	0	1	1	0	2
0	0	1	1	0	0	0	1	1	0	1	0	1	1	1	2
0	0	1	1	0	0	1	2	1	0	1	1	0	0	0	2
0	0	1	1	0	1	0	2	1	0	1	1	0	0	1	2
0	0	1	1	0	1	1	2	1	0	1	1	0	1	0	2
0	0	1	1	1	0	0	2	1	0	1	1	0	1	1	2
0	0	1	1	1	0	1	2	1	0	1	1	1	0	0	2
0	0	1	1	1	1	0	2	1	0	1	1	1	0	1	2
0	0	1	1	1	1	1	2	1	0	1	1	1	1	0	2
0	1	0	0	0	0	0	1	1	0	1	1	1	1	1	2
0	1	0	0	0	0	1	1	1	1	0	0	0	0	0	2
0	1	0	0	0	1	0	1	1	1	0	0	0	0	1	2
0	1	0	0	0	1	1	2	1	1	0	0	0	1	0	2
0	1	0	0	1	0	0	2	1	1	0	0	0	1	1	2
0	1	0	0	1	0	1	2	1	1	0	0	1	0	0	2
0	1	0	0	1	1	0	2	1	1	0	0	1	0	1	2
0	1	0	0	1	1	1	2	1	1	0	0	1	1	0	2
0	1	0	1	0	0	0	2	1	1	0	0	1	1	1	2
0	1	0	1	0	0	1	2	1	1	0	1	0	0	0	2
0	1	0	1	0	1	0	2	1	1	0	1	0	0	1	2
0	1	0	1	0	1	1	2	1	1	0	1	0	1	0	2
0	1	0	1	1	0	0	2	1	1	0	1	0	1	1	2
0	1	0	1	1	0	1	2	1	1	0	1	1	0	0	2
0	1	0	1	1	1	0	2	1	1	0	1	1	0	1	2
0	1	0	1	1	1	1	2	1	1	0	1	1	1	0	2
0	1	1	0	0	0	0	2	1	1	0	1	1	1	1	2
0	1	1	0	0	0	1	2	1	1	1	0	0	0	0	2
0	1	1	0	0	1	0	2	1	1	1	0	0	0	1	2
0	1	1	0	0	1	1	2	1	1	1	0	0	1	0	2
0	1	1	0	1	0	0	2	1	1	1	0	0	1	1	2
0	1	1	0	1	0	1	2	1	1	1	0	1	0	0	2
0	1	1	0	1	1	0	2	1	1	1	0	1	0	1	2
0	1	1	0	1	1	1	2	1	1	1	0	1	1	0	2
0	1	1	1	0	0	0	2	1	1	1	0	1	1	1	2
0	1	1	1	0	0	1	2	1	1	1	1	0	0	0	2
0	1	1	1	0	1	0	2	1	1	1	1	0	0	1	2
0	1	1	1	0	1	1	2	1	1	1	1	0	1	0	2
0	1	1	1	1	0	0	2	1	1	1	1	0	1	1	2
0	1	1	1	1	0	1	2	1	1	1	1	1	0	0	2
0	1	1	1	1	1	0	2	1	1	1	1	1	0	1	2

Appendix B

Transformation of the non-smooth optimization problem into an MILP problem, (8)–(16), according to the standard procedure described in Bemporad and Morari (1999).

First, we introduce the following binary variables:

$$\delta^{\text{end}}_{g,l} = 1 \;\Leftrightarrow\; d^{\text{end}}_{g} - x_l \leqslant 0 \tag{17}$$

$$\delta^{\text{start}}_{g,l} = 1 \;\Leftrightarrow\; d^{\text{start}}_{g} - x_l \leqslant 0 \tag{18}$$

$$\overline{\delta}_g = 1 \;\Leftrightarrow\; d^{end}_{g} - \overline{\xi} \leqslant 0 \tag{19}$$

$$\underline{\delta}_g = 1 \;\Leftrightarrow\; d^{start}_{g} - \leqslant 0 \tag{20}$$

$\forall\, g \in \{1,\ ...,N_{\text{c}}\},\ \forall\, l \in \{1,\ ...,N_{\text{d}}\}.$

Then we introduce the following variables:

$$z_{1,g} = \underline{\delta}_g d^{\text{start}}_{g} \tag{21}$$

$$z_{2,g} = \overline{\delta}_g d^{\text{end}}_{g} \tag{22}$$

$$z_{3,g} = \underline{\delta}_g d^{\text{start}}_{g+1} \tag{23}$$

$\forall\, i \in \{1,\ ...,N_{\text{c}}\}.$

Define

$$\delta = [\underbrace{\delta^{\text{start}}_{1,1}\cdots\delta^{\text{start}}_{N_{\text{c}},N_{\text{d}}}\ \delta^{\text{end}}_{1,1}\cdots\delta^{\text{end}}_{N_{\text{c}},N_{\text{d}}}}_{(\delta^{start})^T \text{ and } (\delta^{end})^T}\ \underbrace{\underline{\delta}_1\cdots\underline{\delta}_{N_{\text{c}}}}_{\underline{\delta}^T}\ \underbrace{\overline{\delta}_1\cdots\overline{\delta}_{N_{\text{c}}}}_{\overline{\delta}^T}]^T$$

$$z = [\underbrace{z_{1,1}\cdots z_{1,N_{\text{c}}}}_{z_1^T}\ \underbrace{z_{2,1}\cdots z_{2,N_{\text{c}}}}_{z_2^T}\ \underbrace{z_{1,1}\cdots z_{1,N_{\text{c}}}}_{z_3^T}]^T$$

These equations can be transformed into equivalent mixed integer linear model expressed as:

$$\min_{d,\delta,z} \sum_{g=1}^{N_{\text{c}}} \left(\sum_{l=1}^{N_{\text{c}}} \omega_l (\delta^{end}_{g,l} - \delta^{start}_{g,l}) + \overline{\delta}_g \right) \tag{24}$$

subject to constraint (11)–(14) and

$$\underline{\delta}_g - \overline{\delta}_g = 0 \quad \forall\ g \in \{1,\ ...,N_c\} \tag{25}$$

$$\sum_{i=1}^{N_{\text{c}}-1} \left(\left(\frac{1}{v^{\text{on}}_{\text{G}}} - \frac{1}{v^{\text{off}}_{\text{G}}} \right) z_{2,g} - \frac{1}{v^{\text{on}}_{\text{G}}} z_{1,g} + \frac{1}{v^{\text{off}}_{\text{G}}} z_{3,g} + (T^{\text{on}}_{\text{G}} + T^{\text{off}}_{\text{G}}) \overline{\delta}_g \right)$$
$$+ \frac{1}{v^{\text{on}}_{\text{G}}} (z_{2,N_{\text{c}}} - z_{1,N_{\text{c}}}) + (T^{\text{on}}_{\text{G}} + T^{\text{off}}_{\text{G}}) \overline{\delta}_{N_{\text{c}}} \leqslant T_{\text{t}} - T_{\text{s}} \tag{26}$$

References

Åhrén, T., Parida, A., 2009. Maintenance performance indicators (MPIs) for benchmarking the railway infrastructure: a case study. Benchmark.: Int. J. 16 (2), 247–258.

Andrade, A.R., Teixeira, P.F., 2011. Uncertainty in rail-track geometry degradation: Lisbon-Oporto line case study. J. Transp. Eng. 137 (3), 193–200.

Andrade, A.R., Teixeira, P.F., 2012. A Bayesian model to assess rail track geometry degradation through its life-cycle. Res. Transp. Econ. 36 (1), 1–8.

Attoh-Okine, N.O., 2017. Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering. John Wiley & Sons.

Bemporad, A., Morari, M., 1999. Control of systems integrating logic, dynamics, and constraints. Automatica 35 (3), 407–427.

Bing, A.J., Gross, A., 1983. Development of railroad track degradation models. Transp. Res. Rec. 939, 27–31.

Budai, G., Huisman, D., Dekker, R., 2006. Scheduling preventive railway maintenance activities. J. Oper. Res. Soc. 57 (9), 1035–1044.

Caetano, L.F., Teixeira, P.F., 2016. Strategic model to optimize railway-track renewal operations at a network level. J. Infrastruct. Syst. 22 (2), 04016002.

Camastra, F., Ciaramella, A., Giovannelli, V., Lener, M., Rastelli, V., Staiano, A., Staiano, G., Starace, A., 2015. A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference. Expert Syst. Appl. 42 (3), 1710–1716.

Corbin, J.C., Fazio, A.E., 1981. Performance-based track-quality measures and their application to maintenance-of-way planning. Transp. Res. Rec. 802, 19–27.

Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., De Schutter, B., 2016, July. Deep convolutional neural networks for detection of rail surface defects. In: International Joint Conference on Neural Networks (IJCNN), pp. 2584–2589.

Fan, Y., Dixon, S., Edwards, R.S., Jian, X., 2007. Ultrasonic surface wave propagation and interaction with surface defects on rail track head. Ndt & E Int. 40 (6), 471–477.

Fumeo, E., Oneto, L., Anguita, D., 2015. Condition based maintenance in railway transportation systems based on big data streaming analysis. Proc. Comp. Sci. 53, 437–446.

Gandomi, A., Haider, M., 2015. Beyond the hype: big data concepts, methods, and analytics. Int. J. Inf. Manage. 35 (2), 137–144.
Ghofrani, F., He, Q., Goverde, R.M., Liu, X., 2018. Recent applications of big data analytics in railway transportation systems: a survey. Transp. Res. Part C: Emerg. Technol. 90, 226–246.
Grassie, S.L., Elkins, J.A., 2005. Tractive effort, curving and surface damage of rails: Part 1. Forces exerted on the rails. Wear 258 (7–8), 1235–1244.
Grassie, S.L., 2012. Squats and squat-type defects in rails: the understanding to date. Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit 226 (3), 235–242.
Hajizadeh, S., Li, Z., Dollevoet, R.P., Tax, D.M., 2014, August. Evaluating classification performance with only positive and unlabeled samples. In: Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR). Springer, Berlin, Heidelberg, pp. 233–242.
Hamid, A., Gross, A., 1981. Track-quality indices and track degradation models for maintenance-of-way planning. Transp. Res. Board 802, 2–8.
He, Q., Kamarianakis, Y., Jintanakul, K., Wynter, L., 2013. Incident duration prediction with hybrid tree-based quantile regression. In: Advances in Dynamic Network Modeling in Complex Transportation Systems. Springer, New York, NY, pp. 287–305.
He, Q., Li, H., Bhattacharjya, D., Parikh, D.P., Hampapur, A., 2015. Track geometry defect rectification based on track deterioration modelling and derailment risk assessment. J. Oper. Res. Soc. 66 (3), 392–404.
Heinicke, F., Simroth, A., Scheithauer, G., Fischer, A., 2015. A railway maintenance scheduling problem with customer costs. EURO J. Transp. Logist. 4 (1), 113–137.
Islam, D.M.Z., Laparidou, K., Burgess, A., 2016. Cost effective future derailment mitigation techniques for rail freight traffic management in Europe. Transp. Res. Part C: Emerg. Technol. 70, 185–196.
Jamshidi, A., Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., Dollevoet, R., Li, Z., Schutter, B., 2017a. A big data analysis approach for rail failure risk assessment. Risk Anal. 37 (8), 1495–1507.
Jamshidi, A., Núñez, A., Dollevoet, R., Li, Z., 2017b. Robust and predictive fuzzy key performance indicators for condition-based treatment of squats in railway infrastructures. J. Infrastruct. Syst. 23 (3), 04017006.
Kawaguchi, A., Miwa, M., Terada, K., 2005. Actual data analysis of alignment irregularity growth and its prediction model. Quart. Rep. RTRI 46 (4), 262–268.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Adv. Neural Inform. Process. Syst. 1097–1105.
Lasisi, A., Attoh-Okine, N., 2018. Principal components analysis and track quality index: a machine learning approach. Transp. Res. Part C: Emerg. Technol. 91, 230–248.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436.
Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., Hampapur, A., 2014. Improving rail network velocity: a machine learning approach to predictive maintenance. Transp. Res. Part C: Emerg. Technol. 45, 17–26.
Li, Z., 2009. Squats on railway rails. In: Wheel-Rail Interface Handbook. Woodhead Publishing Books, pp. 409–436.
Li, Z., Dollevoet, R., Molodova, M., Zhao, X., 2011. Squat growth—some observations and the validation of numerical predictions. Wear 271 (1–2), 148–157.
Li, Z., Molodova, M., Núñez, A., Dollevoet, R., 2015. Improvements in axle box acceleration measurements for the detection of light squats in railway infrastructure. IEEE Trans. Ind. Electron. 62 (7), 4385–4397.
Li, Z., Zhao, X., Esveld, C., Dollevoet, R., Molodova, M., 2008. An investigation into the causes of squats—correlation analysis and numerical modeling. Wear 265 (9–10), 1349–1355.
Liu, X., Dick, C.T., 2016. Risk-based optimization of rail defect inspection frequency for petroleum crude oil transportation. Transp. Res. Rec.: J. Transp. Res. Board 2545, 27–35.
Liu, X., Barkan, C., Saat, M., 2011. Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transp. Res. Rec.: J. Transp. Res. Board 2261, 178–185.
Liu, X., Saat, M., Barkan, C., 2012. Analysis of causes of major train derailment and their effect on accident rates. Transp. Res. Rec.: J. Transp. Res. Board 2289, 154–163.
Ma, F., Wu, Q., Yan, X., Chu, X., Zhang, D., 2015. Classification of automatic radar plotting aid targets based on improved fuzzy C-means. Transp. Res. Part C: Emerg. Technol. 51, 180–195.
Makino, T., Kato, T., Hirakawa, K., 2012. The effect of slip ratio on the rolling contact fatigue property of railway wheel steel. Int. J. Fatigue 36 (1), 68–79.
Makino, T., Neishi, Y., Shiozawa, D., Kikuchi, S., Okada, S., Kajiwara, K., Nakai, Y., 2016. Effect of defect shape on rolling contact fatigue crack initiation and propagation in high strength steel. Int. J. Fatigue 92, 507–516.
Mariani, S., Nguyen, T., Phillips, R.R., Kijanka, P., Lanza di Scalea, F., Staszewski, W.J., Fateh, M., Carr, G., 2013. Noncontact ultrasonic guided wave inspection of rails. Struct. Health Monit. 12 (5–6), 539–548.
Markowski, A.S., Mannan, M.S., 2009. Fuzzy logic for piping risk assessment (pfLOPA). J. Loss Prevent. Process Ind. 22 (6), 921–927.
Molodova, M., Li, Z., Núñez, A., Dollevoet, R., 2014. Automatic detection of squats in railway infrastructure. IEEE Trans. Intell. Transp. Syst. 15 (5), 1980–1990.
Mutton, P.J., Epp, C.J., Dudek, J., 1991. Rolling contact fatigue in railway wheels under high axle loads. In: Mechanics and Fatigue in Wheel/Rail Contact, pp. 139–152.
Nielsen, J.C., Ekberg, A., Lundén, R., 2005. Influence of short-pitch wheel/rail corrugation on rolling contact fatigue of railway wheels. Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit 219 (3), 177–187.
Parida, A., Chattopadhyay, G., 2007. Development of a multi-criteria hierarchical framework for maintenance performance measurement (MPM). J. Qual. Maintenance Eng. 13 (3), 241–258.
Peng, F., Ouyang, Y., 2012. Track maintenance production team scheduling in railroad networks. Transp. Res. Part B: Methodol. 46 (10), 1474–1488.
Peng, F., Ouyang, Y., 2014. Optimal clustering of railroad track maintenance jobs. Comput.-Aid. Civ. Infrastruct. Eng. 29 (4), 235–247.
Popović, Z., Radović, V., Lazarević, L., Vukadinović, V., Tepić, G., 2013. Rail inspection of RCF defects. Metalurgija 52 (4), 537–540.
Santos, R., Teixeira, P.F., 2011. Heuristic analysis of the effective range of a track tamping machine. J. Infrastruct. Syst. 18 (4), 314–322.
Sciammarella, C.A., Chen, R.J.S., Gallo, P., Berto, F., Lamberti, L., 2016. Experimental evaluation of rolling contact fatigue in railroad wheels. Int. J. Fatigue 91, 158–170.
Scott, D., Fletcher, D.I., Cardwell, B.J., 2014. Simulation study of thermally initiated rail defects. Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit 228 (2), 113–127.
Sharma, S., Cui, Y., He, Q., Mohammadi, R., Li, Z., 2018. Data-driven optimization of Railway maintenance for track geometry. Transp. Res. Part C: Emerg. Technol. 90, 34–58.
Song, Z., Yamada, T., Shitara, H., Takemura, Y., 2011. Detection of damage and crack in railhead by using eddy current testing. J. Electromagn. Anal. Appl. 3 (12), 546–550.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15 (1), 1929–1958.
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. Transp. Res. Part C: Emerg. Technol. 84, 92–123.
Tosun, M., Dincer, K., Baskaya, S., 2011. Rule-based Mamdani-type fuzzy modelling of thermal performance of multi-layer precast concrete panels used in residential buildings in Turkey. Expert Syst. Appl. 38 (5), 5553–5560.
Veit, P., 2007. Track quality–luxury or necessity. In: Railway Technical Review Special: Maintenance and Renewal, pp. 8–12.
Vo, K.D., Zhu, H.T., Tieu, A.K., Kosasih, P.B., 2015. Comparisons of stress, heat and wear generated by AC versus DC locomotives under diverse operational conditions. Wear 328, 186–196.
Wang, L., Wang, P., Quan, S., Chen, R., 2012. The effect of track alignment irregularity on wheel-rail contact geometry relationship in a turnout zone. J. Mod. Transp. 20 (3), 148–152.
Wen, M., Li, R., Salling, K.B., 2016. Optimization of preventive condition-based tamping for railway tracks. Eur. J. Oper. Res. 252 (2), 455–465.

Xie, M., 2003. Fundamentals of Robotics: Linking Perception to Action. World Scientific Publishing Company, pp. 54.
Xu, L., Zhai, W., 2017. A novel model for determining the amplitude-wavelength limits of track irregularities accompanied by a reliability assessment in railway vehicle-track dynamics. Mech. Syst. Sig. Process. 86, 260–277.
Zhuang, L., Wang, L., Zhang, Z., Tsui, K.L., 2018. Automated vision inspection of rail surface cracks: a double-layer data-driven framework. Transp. Res. Part C: Emerg. Technol. 92, 258–277.
Zoeteman, A., Dollevoet, R., Li, Z., 2014. Dutch research results on wheel/rail interface management: 2001–2013 and beyond. Proc. Inst. Mech. Eng., Part F: J. Rail Rapid Transit 228 (6), 642–651.
Zywiel, J., Oberlechner, G., 2001. Innovative measuring system unveiled. Int. Railway J. 41 (9).