Robust and Predictive Fuzzy Key Performance Indicators for Condition-Based Treatment of Squats in Railway Infrastructures

Ali Jamshidi; Alfredo Núñez; Rolf Dollevoet; and Zili Li

Abstract: This paper presents a condition-based treatment methodology for a type of rail surface defect called squat. The proposed methodology is based on a set of robust and predictive fuzzy key performance indicators. A fuzzy Takagi–Sugeno interval model is used to predict squat evolution for different scenarios over a time horizon. Models including the effects of maintenance to treat squats, via either grinding or replacement of the rail, are also described. A railway track may contain a huge number of squats distributed in the rail surface with different levels of severity. This paper proposes to aggregate the local squat interval models into track-level performance indicators including the number and density of squats per track partition. To facilitate the analysis of the overall condition, a single fuzzy global performance indicator per track partition is proposed based on a fuzzy expert system that combines all the scenarios and predictions over time. The proposed methodology relies on the early detection of squats using axle box acceleration measurements. Real-life measurements from the Meppel-Leeuwarden track in the Dutch railway network are used to show the benefits of the proposed methodology. The use of robust and predictive fuzzy performance indicators facilitates the visualization of the track health condition and eases the maintenance decision process. DOI: 10.1061/(ASCE)IS.1943-555X.0000357. © 2017 American Society of Civil Engineers.

Author keywords: Key performance indicators design; Railway track condition monitoring and maintenance; Interval fuzzy models.

Introduction

During recent years, a modal shift from road to rail has been promoted in Europe. The idea is to increase the share of transport demand for mobility of people and freight, reduce road traffic congestion, make efficient use of the energy resources, and tackle the major challenges of climate change. Major contributions are needed in the optimal management of railway assets, evolving toward a more automated predictive operation where functional assets are monitored. This includes all the important indicators such as economic, safety, and societal impacts, considering the perspective of both railway infrastructure manager and users (Zoeteman 2001).

A typical set of railway assets is shown in Fig. 1, including the track, station, superstructure, substructure, communication, catenary, control room, signaling system, rolling stock, barrier, security, and surrounding. In order to monitor and properly maintain the railway assets, it is necessary to measure the evolution of important health condition indicators over time, also called key performance indicators (KPIs), for each of the critical assets. For example, in Fig. 1, \( J_{\text{label}}(t) \) relates to the KPI for the health condition of an asset called Asset, uniquely labelled as label at time \( t \). In the Netherlands, the assets in the railway network includes more than 3,000 km of track and 388 stations, being one of the densest networks in Europe. In this network, the design of an optimal maintenance plan for all its assets is a challenging problem. To optimally design the maintenance plans, the infrastructure manager requires crucial information of each asset (Stenström et al. 2015a) and maintenance decision making considering risk-averse situations (Rockafellar and Royset 2015). Thus, the optimal maintenance plan is a necessity because of the high demand from users and government for a better quality of service and the need of keeping costs as low as possible.

Maintenance performance indicators evaluate the system performance and can be used to guarantee that these assets operate at an acceptable level of functionality and safety. Parida and Chattopadhyay (2007) propose a general systems framework using a hierarchical structure of multicriteria maintenance performance measurements. Åhrén and Parida (2009) apply the same framework to the case of benchmarking railway infrastructure maintenance operations. Three different hierarchical levels are proposed: strategic level for top management decisions, tactical level for middle management, and functional level for supervisors and operators. The general framework requires effective measurements of the health condition of the assets considering that the different assets degrade at different rates due to the effect of different exogenous sources. Particularly, the focus of this paper is to design robust and predictive fuzzy performance indicators for health condition monitoring of railway tracks, considering a particular major type of rolling contact fatigue (RCF) called squat (Li et al. 2015).

In the Netherlands more than 40% of the railway maintenance budget is allocated yearly to track maintenance (Zoeteman and van Meer 2006; Zoeteman et al. 2014). The presence of RCFs accelerates track degradation, which negatively influences its health condition. It also increases the noise level that affects people living in the surroundings and in a worst case affects safety because severe RCFs can result in derailment. For track maintenance to be...
effective, the planning should consider not only costs but also the dynamics of RCFs. Complex interactions between environment, vehicles, wheels and track interface, and structure and also different behaviors under maintenance operation such as grinding and rail replacement can be considered. Patra et al. (2009) modeled rail degradation by a time-to-failure function using million gross tons (MGT) measurements and approximately 12 failure events; they proposed decision making in a Monte Carlo simulation setting. The maintenance operations are modeled as different cost functions, including rail grinding costs, track tamping costs, and rail lubrication costs, among other maintenance operations. Stenström et al. (2015b) assess the value of preventive maintenance in comparison with corrective maintenance. The idea is to analyze the cost-benefit of using preventive maintenance including four different maintenance costs: maintenance inspections, repair of potential failures, repair of functional failures, and service and production loss. In the case study for a Swedish railway line, the 10 costliest railway sections are found to have three times the tonnage compared with the sections with the lowest costs, and also the costliest sections experience 4.5 times more track failures. The conclusion is that the railway sections with the lowest total maintenance cost have implemented more preventive maintenance actions.

In the literature, different studies have been carried out to present how a degradation model for tracks can be embedded on asset management to facilitate maintenance plans. Track geometry measurements relying on statistical analysis are used to capture the track degradation effect (Sadeghi and Askarinejad 2010; Andrade and Teixeira 2011, 2012; Andrews 2012; Vale and Lurdus 2013; Nathanail 2014; Guler 2014; Weston et al. 2015). In those papers, different time-dependent degradation models are proposed; they can all be used to improve maintenance interventions. Estimation of the track safety and consideration of the probability of rail break has also been investigated (Schafer and Barkan 2008; Burstow et al. 2002; Sandström and Ekberg 2009). Detailed mechanical models can give many insights about the evolution of rail defects; however, the use of those models for maintenance planning operations requires sophisticated knowledge about the track and its operational conditions that are not always available or easy to obtain in practice. Fuzzy logic has increasingly been used in different fields; in particular, in the ones where uncertainties can influence the decision process. It is used to measure performance in different infrastructures by predicting failure of components (Senouci et al. 2014; Sadiq et al. 2004) and optimizing asset condition (Xu et al. 2014; Wang and Liu 1997) and decision making (Khatri et al. 2011). In this paper, the use of interval fuzzy model is proposed to capture the most important dynamics of squats in railway infrastructure from the maintenance operation point of view. The aim is to keep the prediction as simple as possible, but suitable enough to ease decision making in practice. The use of KPIs that are able to explicitly include the dynamics of the deterioration of the assets, together with an appropriate set of scenarios for the principal sources of stochasticities that might affect their performance, is recommended. A fuzzy Takagi–Sugeno (TS) interval model (Škrjanc 2011; Nuñež and De Schutter 2012; Sáez et al. 2015) is calibrated using real-life data collected over years of field tests and measurements, which helps to obtain numerical models capable of predicting squat growth over a time horizon under different possible scenarios and under different maintenance decisions.

Based on the interval fuzzy models for squats, a condition-based methodology for rails is proposed using different KPIs that are defined in a track-partition level that allows the grouping of defects located in a given track partition. In this methodology, number and density of squats are considered over a prediction horizon under three different scenarios, i.e., slow, average, and fast growth. Then, to facilitate visualization of the track health condition and to ease the maintenance decision process, a fuzzy global KPI based on fuzzy rules for each partition that merges the different KPIs over the prediction horizon and scenarios is proposed. The methodology is evaluated with data from a Dutch railway track, relying on the use of technology-based axle box acceleration (ABA) measurements capable of detecting the early-stage squats on the rail (Mołodova et al. 2014; Li et al. 2015). An introduction of the ABA measuring system is described in ABA-based health condition monitoring in railways, including the background of the ABA measurement system and its application in rail condition monitoring based on ABA.

Fig. 2 is the flowchart of the proposed methodology divided in three steps. In Step 1, relying on ABA measurements, the health condition of the track and severity are estimated. A list of defects is assumed to be provided by the detection algorithm. In Step 2,
using interval fuzzy TS model, the growth of each detected defect $i$ is evaluated over time and different possible evolution scenarios are considered. Three models are evaluated, with grinding, replacement, and without maintenance. The idea is to see the consequences of the maintenance operations on the detected squats for different scenarios over a prediction horizon. At the end, in Step 3, a global fuzzy KPI is used to describe the condition at a track partition level for a given travel direction, left and right rails. The global fuzzy KPI at a partition combines the effects of a vector of KPIs over a prediction horizon, considering three most representative defect evolution scenarios.

The paper is divided as follows. In next section, the main elements of the ABA-based detection methods are presented. Next, interval fuzzy models for squats are presented for three cases: without maintenance, after grinding, and after replacement. Then different KPIs are defined at a track partition level in order to aggregate the local dynamic behavior of squats. Because of the number of scenarios and prediction horizon, the fuzzy global KPI is proposed to facilitate decision making. Next the numerical results and discussion are presented. Finally, conclusions and further research are discussed.

ABA-Based Health Condition Monitoring in Railways

**Background of the ABA Measurement System**

There are different methods to diagnose the condition of rail defects, including ultrasonic measurements, eddy current testing, image recognition, and guided wave–based monitoring, among other technologies. Each has different advantages and disadvantages. In this paper, a technology capable of detecting defects in an early stage is needed, thus the use of ABA measurements is considered (Li et al. 2008; Molodova et al. 2014). Li et al. (2015) investigated the feasibility of detecting early-stage squats using an ABA prototype. It is reported that squats could be detected by analyzing the frequency content of the ABA signals in the wavelet power spectrum. In practice, the useful frequency band for early detection of squats ranges from 1,000 to 2,000 Hz and 200 to 400 Hz (Molodova et al. 2014).

In the literature, it has been reported that ABA systems can be employed to detect surface rail defects like corrugation, squats, and welds in poor condition. The ABA system offers the advantages of (1) having a lower cost than other types of detection methods, (2) being easy to maintain, and (3) ability to be implemented in in-service operational trains. Other significant advantages that ABA offers over similar measurement systems are (1) the ability to detect small defects with the absence of complicated instrumentation and (2) the ability to indicate the level of the dynamic contact force (Molodova et al. 2015).

**Rail Condition Monitoring Based on ABA**

In this study, the authors are users of the ABA detection methodology presented in Li et al. (2015) and Molodova et al. (2014); thus, it is assumed that a list of squats and their location are available. Define the counter of squat defects as $i = 1, 2, \ldots, N_{\text{defects}}$, where $x_i$ represents the position of squat $i$. The variables $H(x, k)$ and...
L(x, k) are defined as the real rail condition and real squat length, respectively, defined at position x and time step k. The focus is only on positions xi where squats are detected. To simplify the notation, it is assumed H_i(k) = H(x_i, k) and L_i(k) = L(x_i, k) represent the severity and the length of squat i at time step k.

To systematically classify squats in terms of severity, the terminology used in Smulders (2003), UIC Code (2002), and Rail Damages (2001) is followed. The definitions of these three references are compatible with one another. Although the transition between one class to the other is not always abrupt, fixed values for those transitions have been defined according to the practical experience of the system. Depending on the squat length L_i(k), measured in millimeters, the severity of the squat can be used to represent the health condition of the rail at location x_i as follows:

\[ H_i(k) = \begin{cases} 
S & \text{if } 0 \leq L_i(k) < 8 \\
A & \text{if } 8 \leq L_i(k) < 30 \\
B & \text{if } 30 \leq L_i(k) < 50 \\
C & \text{if } 50 \leq L_i(k) < 60 \\
RC & \text{if } L_i(k) \geq 60 
\end{cases} \]

where S = seed squat; A = light squat (A squat); B = moderate squat (B squat); C = severe squat (C squat); and RC = squat with risk of derailment. The boundaries were defined based on general guidelines to classify squats.

Fig. 3 gives an example of defect growths collected from field measurements in the Meppel-Leeuwarden track. In the figure, the x-axis represents the kilometer position of the track where the squats are located and the y-axis indicates time in three different months, Month 0 (moment of the measurement), Month 6, and Month 12. In the diagram, A squats are drawn as circles and B squats are squares. Different squats grow with different rates. In the average case, the track measurements show that it takes approximately 9 months for an A squat of 20 mm to evolve into a B squat of 30 mm.

In this study, the ABA measurements are used to develop a model for defect evolution. For each squat, the related energy of the ABA is available using wavelet spectrum analysis and advanced signal processing methods (Moledova et al. 2014). Relying on the ABA measurement, the energy values of the ABA signals can be calculated at every position x at time step k as E(x, k). From the energy signal, the interest is only in those locations with squats, namely, E_i(k) = E(x_i, k). For using the energy of the ABA signal to predict the squat length evolution, a correlation between the squat length and energy of the ABA signal was performed. Photographs from track visits of several years are used to measure the lengths of the squats and to fit the piecewise linear correlation model. The estimated length \( \hat{L}_i(k) \) of squat i at step k as function of the energy value E_i(k) is given by

\[ \hat{L}_i(k) = \begin{cases} 
q_1 & \text{if } E_i(k) < 80 \\
q_2 & \text{if } 80 \leq E_i(k) < 170 \\
q_3 & \text{if } 170 \leq E_i(k) < 300 \\
q_4 & \text{if } E_i(k) \geq 300 
\end{cases} \]

where the slope of local linear functions is \( q_m, m = 1, \ldots, 4 \), and the bias \( q_m, m = 1, \ldots, 4 \), are adjusted to the specific track. For Eq. (2), previous works have been used (Li et al. 2011, 2015; Moledova et al. 2015). In general, it is possible to say that the correlation coefficient and residual standard are affected by the speed of the measurement train. In this paper, it is assumed that the measurement is performed at commercial speed as was done for the test measurement so far, and segments that were measured out of a reasonable range of speed were disregarded.

A global view of Step 1 of the methodology, estimation of track health condition based on ABA, is presented in Fig. 4. As shown in the figure, in order to estimate length \( L_i(k) \), the energy value \( E_i(k) \) is calculated using the ABA measurement. Hence, relying on the estimated squat lengths, the rail health condition \( H_i(k) \) can be approximated. In the figure, a squat is detected with an energy value \( E_i(k) = 145 \text{ mm}^2/\text{s}^2 \), the estimated squat length \( \hat{L}_i(k) = 43 \text{ mm} \), and the estimated health condition \( H_i(k) = B \).

### Interval Fuzzy Models for Squats

#### Maintenance-Oriented Models for Squats

Typically, maintenance slots in the Dutch railway network are decided based on long- and short-term planning for preventive and corrective maintenance, respectively. In the long term, the contractor should inform the asset manager at least 1 year before cyclic grinding for using the equipment needed. In the short term, normally the maintenance is performed when the squats are in the last stage of growth (C squat). Thus, a predictive approach by employing well-designed KPIs should aim to improve both short- and long-term planning by (1) keeping a good balance between costs and health condition of the track, (2) simplifying the design of maintenance plan over the whole time horizon, and (3) indirectly increasing the track safety.

The experimental results show that each squat can grow at different rates. The estimation of squat lengths can be affected by the subjectivity of human error. For instance, one source of uncertainty comes from the fact that visually only the rusty area of the defects is used to measure length, while the defect might be longer. Fuzzy systems can work under subjective environments. In the proposed methodology, the design of the global fuzzy KPI deals with the subjectivity. The definition of a low or a high number of defects will depend on the subjectivity of the infrastructure manager, and on how this information is incorporated for maintenance decision making. In order to generalize this characteristic, interval fuzzy models can be used to capture the stochasticities of different scenarios for the squat growth. The upper bound of the interval represents a worst-case scenario, while the lower bound represents...
a slow growth rate scenario. In the interval fuzzy models, the average behavior is given by a TS fuzzy model. This is used to approximate nonlinearities by smoothly interpolating affine local models. Each local model is involved in the global model based on the activation of a membership function. According to literature, the identification of interval fuzzy models is divided into three steps: clustering method to generate fuzzy rules, identification of the TS local linear parameters (average model), and identification of the fuzzy variance for each rule (Škrljanc et al. 2005; Škrljanc 2011). In this paper, the use of the interval fuzzy models in Nuñez and De Schutter (2012) and Sáez et al. (2015) is proposed, which includes Gustafson–Kessel clustering, local identification of the linear parameters, and optimization of a parameter $\alpha$ to adjust the width of the interval, minimizing both the area of the band and number of data points outside the band.

The general problem of interval defect evolution is as follows. Consider different defect growth scenarios $h = h_1, h_2, \ldots, h_H$, time steps $i = k, k+1, k+2, \ldots, k+N_p$, and the maintenance action at time step $k$, $u(k)$. The prediction model for the growth of a squat can be written as

$$\hat{L}_i^h(k+1) = f_j^h[L_i(k), u(k)], \quad x_i \in [x_j, x_{j+1})$$

where $\hat{L}_i^h(k+1)$ is the estimation of the length of the squat $i$ located in the track partition $j$ at the time step $k+1$ considering the scenario $h$. The model considers the effect of maintenance $u(k)$ and the initial condition of the squat $L_i(k)$. Depending on the location of the squat $i$, $x_i$, a local model is used corresponding to the track partition $j$ where the squat is located, $x_i \in [x_j, x_{j+1})$. It is assumed that the dynamics for different squats are similar if they are in the same track partition under the same scenario.

In this paper, three maintenance actions are considered, $u(k) = \{u_1, u_2, u_3\}$, where $u_1$ is without maintenance, $u_2$ is grinding, and $u_3$ is replacement. Three scenarios are also evaluated, $h = h_1, h_2, h_3$, where $h_1$ represents slow-growth, $h_2$ average-growth, and $h_3$ fast-growth scenarios.

### Dynamics of Squats without Maintenance

In the absence of maintenance, $u(k) = u_3$, the prediction model for the average growth scenario, $h_2$, is formulated based on the TS fuzzy model

$$\hat{L}_i^{h_2}(k+1) = \sum_{r=1}^{N_p} \beta_{jr} [Li(k)] L_{jr}(k)$$

$$L_{jr}(k) = a_{jr} L_i(k) + b_{jr} \beta_{jr}[L_i(k)] = \frac{A_{jr}[L_i(k)]}{\sum_{r=1}^{N_p} A_{jr}[L_i(k)]}$$

where $a_{jr}, b_{jr}$ are parameters of the fuzzy local model on rule $r$, $r = 1, 2, \ldots, N_p$, and $\beta_{jr}[L_i(k)]$ is normalized activation degree of the rule $r$. In this paper the use of Gaussian functions is proposed to model the membership degrees, $A_{jr}[L_i(k)] = \exp\{-0.5c_{jr,1}[L_i(k) - c_{jr,2}]^2\}$, defined by parameters $c_{jr,1}$ and $c_{jr,2}$ given by the Gustafson–Kessel clustering algorithm.

Once the TS model is obtained, the slow-growth and fast-growth scenarios are used as lower and upper bounds of the average growth scenario, $\hat{L}_i^{h_2}(k+1)$, respectively. The equations can be defined as

$$\hat{L}_i^{h_1}(k+1) = \sum_{r=1}^{N_p} \beta_{jr} [Li(k)] [L_{jr}(k) - \alpha_{h_1} \Delta_{jr}[L_i(k)]]$$

$$\hat{L}_i^{h_3}(k+1) = \sum_{r=1}^{N_p} \beta_{jr} [Li(k)] [L_{jr}(k) - \alpha_{h_3} \Delta_{jr}[L_i(k)]]$$

where $\Delta_{jr}[L_i(k)] = \sigma_{jr}[1 + \psi_{jr} \psi_{jr}^T]^{-1} \psi_{jr}^{0.5}$

In this paper, three maintenance actions are considered, $u(k) = \{u_1, u_2, u_3\}$, where $u_1$ is without maintenance, $u_2$ is grinding, and $u_3$ is replacement. Three scenarios are also evaluated, $h = h_1, h_2, h_3$, where $h_1$ represents slow-growth, $h_2$ average-growth, and $h_3$ fast-growth scenarios.

Fig. 4. Global scheme of the main components of Step 1: estimation of track health condition based on ABA (images by Alfredo Núñez)
scenarios, respectively; and $\alpha^h$ and $\alpha^s$ are tuning parameters in the fast-growth and slow-growth scenarios, respectively. Moreover, $\Psi^f, \Psi^s$, and $\Psi^r$ are the covariance matrix, regression matrix, and variance of the local model.

Fig. 5 depicts the proposed interval fuzzy models including 177 data points used to capture the squat evolution in different stages of growth. A subset of the data used for analysis is included in Table 1. The A squats from 8 to 30 mm in length have no or shallow cracks. The B squats ranging from 30 to 50 mm grow quickly. The B squats evolve to C squats when the network of cracks beneath the squat gets further spread. All three stages are shown by reference photos of an A squat, B squat, and C squat in Fig. 5.

### Rail Grinding Effect

Squats can be effectively treated by grinding when they are in the early stage of growth. Cyclic rail grinding keeps control of not only maintaining the rail profiles but to plan track maintenance efficiently (Magel and Kalousek 2002). Fig. 6 depicts squat growth before and after grinding, with black circles indicating those squats that did not disappear after grinding. As seen in the figure, some A squats are located in the effective zone of grinding such that these squats have a zero length after grinding. Those A squats that are imminent to becoming B squats are located in the ineffective zone for grinding as well as B squats and C squats. Moreover, three growth scenarios in the effective zone are specified to capture the squat evolution rate. Even though grinding severe squats postpones rail replacement, it could accelerate squat evolution because the cracks have not totally disappeared.

The growth model for squat $i$ by considering grinding effect can be expressed as

$$ L^h_i(k+1) = \begin{cases} L_i(k) & L_i(k) \leq L_G^{eff} \\ z_h^i [L_i(k) - L_G^{eff}] + L_i(k) & L_i(k) > L_G^{eff} \end{cases} $$

where $L_G^{eff}$ = critical squat length that estimate effectivity of grinding, with $L_G^{eff}$ approximately 20 mm in Fig. 6 for a grinding depth of 1.0 mm; and $z_h^i$ = slope of the linear model in the ineffective zone for grinding for different scenarios $h$, slow-, average-, and fast-growth scenarios.

### Rail Replacement Effect

When the squat severity becomes worse and cracks grow considerably, grinding is not efficient anymore. Therefore, replacement is the only solution. Because replacing a piece of rail takes time and is costly, optimal decision making for when and where the rail should be replaced is important. Higher rail (larger radius) and low rail (smaller radius) have different degradation behaviors (Patra et al. 2009), thus usually only the most needed rail is replaced. Rail replacement is performed using welds to connect the new rail with the old one. After replacement, the rail surface defects will totally disappear by the installation of new rail, whereas development of new squats will depend on various factors, such as track conditions and MGT. In the case of the welds, because they are composed by materials with different properties than the rails, they are prone to squat defect appearance (Lewis and Olofsson 2009).

---

**Table 1. Subset of Data Used for Squat Analysis Including Defect Position and Visual Length at Times $k$ and $k+1$**

<table>
<thead>
<tr>
<th>Squat</th>
<th>Position (km)</th>
<th>$L_i(k)$ (mm)</th>
<th>$L_i(k+1)$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>104.8438</td>
<td>30.7260</td>
<td>34.7465</td>
</tr>
<tr>
<td>2</td>
<td>105.1051</td>
<td>37.7420</td>
<td>40.5086</td>
</tr>
<tr>
<td>3</td>
<td>105.1404</td>
<td>33.2264</td>
<td>37.0496</td>
</tr>
<tr>
<td>4</td>
<td>105.2116</td>
<td>34.2207</td>
<td>37.7799</td>
</tr>
<tr>
<td>5</td>
<td>105.3215</td>
<td>46.7870</td>
<td>49.1017</td>
</tr>
<tr>
<td>6</td>
<td>105.3901</td>
<td>33.0151</td>
<td>36.8862</td>
</tr>
<tr>
<td>7</td>
<td>105.4195</td>
<td>19.1797</td>
<td>21.6607</td>
</tr>
<tr>
<td>8</td>
<td>105.4269</td>
<td>20.2236</td>
<td>22.5435</td>
</tr>
<tr>
<td>9</td>
<td>105.4344</td>
<td>9.4918</td>
<td>12.4747</td>
</tr>
<tr>
<td>10</td>
<td>105.4561</td>
<td>37.2979</td>
<td>37.0903</td>
</tr>
<tr>
<td>11</td>
<td>105.4613</td>
<td>22.8311</td>
<td>24.6695</td>
</tr>
<tr>
<td>12</td>
<td>105.4953</td>
<td>19.5933</td>
<td>22.0216</td>
</tr>
<tr>
<td>13</td>
<td>105.5827</td>
<td>14.5360</td>
<td>16.7962</td>
</tr>
<tr>
<td>14</td>
<td>105.5852</td>
<td>19.5432</td>
<td>21.9787</td>
</tr>
<tr>
<td>15</td>
<td>105.6353</td>
<td>11.0032</td>
<td>13.9019</td>
</tr>
<tr>
<td>16</td>
<td>105.6591</td>
<td>25.1642</td>
<td>27.1955</td>
</tr>
<tr>
<td>17</td>
<td>105.7462</td>
<td>15.4564</td>
<td>17.7552</td>
</tr>
<tr>
<td>18</td>
<td>106.3105</td>
<td>28.7262</td>
<td>32.2161</td>
</tr>
<tr>
<td>19</td>
<td>106.8735</td>
<td>55.1141</td>
<td>57.1707</td>
</tr>
<tr>
<td>20</td>
<td>107.2845</td>
<td>17.8761</td>
<td>20.4044</td>
</tr>
</tbody>
</table>

---

**Fig. 5.** Interval fuzzy model for squat growth in the case study track (images by Alfredo Núñez)

**Fig. 6.** Squat growth before grinding and after grinding classified in two effective and ineffective zones for grinding operations; in this case, the depth of the grinding was approximately 1.0 mm.
Figs. 7(a and b) show squat growth before and after rail replacement. Fig. 7(a) shows the squat growth between welds where all the squats will disappear after replacement. The model assumes that no squats will appear during a long horizon by considering that newly developed squats can be detected in the next measurement campaign. Fig. 7(b) shows squat growth on the welds in a period after replacement. The exact time instant when the growth starts is related to quality of the weld. This means that for those welds that have good quality, the starting point would be much later.

In the case of between welds, the squat length after replacement is equivalent to zero during a time horizon $N_2$. The growth model on the weld can be expressed according to the time $N_2$, when squat can appear. Before time $k + N_2$, no squat is present in the weld, while at $k + N_2 + 1$ the squat will start to appear and evolve based on the proposed growth scenarios

\[
\hat{L}_i^h(x_{w_2}, k + t) = 0 \quad t = 1, 2, \ldots, N_1, h = h_1, h_2, h_3
\]

\[
\tilde{L}_i^h(x_{w_2}, k + t) = 0 \quad t = 1, 2, \ldots, N_2, h = h_1, h_2, h_3
\]

\[
\hat{L}_i^h(x_{w_2}, k + N_2 + 1) = \begin{cases} 
\frac{T_i^{TS} (\Delta L_i)}{1 - \Delta L_i} & \text{if } h = h_1 \\
\frac{T_i^{TS} (\Delta L_i)}{1 - \Delta L_i} & \text{if } h = h_2 \\
\frac{T_i^{TS} (\Delta L_i)}{1 - \Delta L_i} & \text{if } h = h_3 
\end{cases}
\]

where $x_{w_2}$ = some position between the welds; $x_{w_2}$ = location of the weld; and $\Delta L_i$ = small value that triggers the growth when squat $i$ starts evolving at the thermite weld at time instant $k + N_2 + 1$. After the squat appears, the interval fuzzy model will capture its evolution over time.

**Key Performance Indicators for Rail Health Condition**

**Description of KPIs**

The monitoring of the evolution of a single squat might not be practical from the maintenance perspective. Aggregated information over bigger track partitions can facilitate the infrastructure manager decision over the maintenance plans. In the case of squats, KPIs are proposed considering the number of A, B, and C squats and the number of squats with potential risk of rail break called RC squats, at different times $t$ and different growth scenarios $h$. Moreover, because a significant number of B and C squats near each other indicate a high potential risk to track safety, a KPI is proposed relying on a measure of density of squats B and C. Assume the function $\delta_{h,j}^d(x, k)$ is provided by the ABA detection algorithm for the current instant of measurement $k$. The function equals to 1 if a squat type $d \in \{A, B, C, RC\}$ is located at position $x$, instant $k$, partition $j$, and growth scenario $h$, and equals to 0 otherwise. Used as the initial condition, and relying on the interval fuzzy model, it is possible to predict $\delta_{h,j}^d(x, t)$ for any time horizon, $t = 1, \ldots, N_p$. The growth of new squats during the prediction horizon is not considered in this paper because it is assumed that new squats will be detected in the next measurement campaign at instant $k + 1$, where the models can be updated according to the new conditions.

The KPIs of squat numbers at partition $j$, instant $t$, and scenario $h$ can be expressed as

\[
y_{h,j}^A(t) = \sum_{x \in [x_j, x_{j+1}]} \delta_{h,j}^A(x, t)
\]

\[
y_{h,j}^B(t) = \sum_{x \in [x_j, x_{j+1}]} \delta_{h,j}^B(x, t)
\]

\[
y_{h,j}^C(t) = \sum_{x \in [x_j, x_{j+1}]} \delta_{h,j}^C(x, t)
\]

\[
y_{h,j}^{RC}(t) = \sum_{x \in [x_j, x_{j+1}]} \delta_{h,j}^{RC}(x, t)
\]

Also, to estimate the density of B and C squats $\delta_{h,j}^{BC}(x, t)$, a window is defined around the coordinate $x$ (in this paper, the window is 50 m in track length). The function $\delta_{h,j}^{BC}(x, t)$ equals the number of squats B or C in the moving window $[x − 0.025, x + 0.025]$. The KPI density for partition $j$, instant $t$, and scenario $h$ can be defined as the area of the density function as follows:

\[
y_{h,j}^{BC}(t) = \sum_{x \in [x_j, x_{j+1}]} \frac{\delta_{h,j}^{BC}(x, t)}{x_{j+1} - x_j}
\]

Define a vector containing all the KPIs called $y_{h,j}(t)$ for partition $j$, instant $t$, and scenario $h$.
where \( y_{h,j}(t) = [y_{h,j}^A(t), y_{h,j}^B(t), y_{h,j}^C(t), y_{h,j}^{RC}(t), y_{h,j}^{ABC}(t)]^T \)

Mamdani Fuzzy KPI

For Step 1, a Mamdani fuzzy expert system is used to calculate a single KPI (Mamdani and Assilian 1975). Even though the Mamdani fuzzy system approach was proposed more than 40 years ago, it is still popular because of its simplicity and interpretability (Camastra et al. 2015; Rezaei et al. 2015; Ozgur 2013). In this case, 32 fuzzy if-then rules are generated. The aim is to assign a membership degree to each KPI to represent the contribution of each KPI in the rail health condition.

If \( y_{h,j}^A(t) \) is \( A_1^r \) and \( y_{h,j}^B(t) \) is \( A_2^r \) and \( y_{h,j}^C(t) \) is \( A_3^r \) and \( y_{h,j}^{RC}(t) \) is \( A_4^r \) and \( y_{h,j}^{ABC}(t) \) is \( G^r \)

where \( A_1^r, A_2^r, A_3^r, A_4^r, \) and \( G^r \) = membership functions for rule \( r \), and \( y_{h,j}^M(t) \) = output Mamdani KPI. The KPIs are first normalized, then Gaussian membership functions are used to fuzzify the KPIs. Also, to defuzzify, the center of gravity method is applied in order to obtain a crisp value at the end. Furthermore, relying on the fuzzy rules, interdependency of KPIs and the Mamdani KPI are captured as shown in Fig. 8. This figure presents how Mamdani KPI models the influence in the health of the track of two KPIs, varying from fully healthy (equals to 0) to completely unhealthy (equals to 1), while all the other KPIs are assumed to be fully healthy (equals to 0). In Fig. 8(a), a higher value for the BC density is much more relevant than the contribution of the number of B squats. In Fig. 8(b), a high number of C squats makes the most significant impact on the rail health condition. The rail condition will get highly unhealthy with high values of either density of the BC squats or number of C squats. In Fig. 8(c), a high number of RC squats will more strongly influence the health condition than the number of A squats. In Fig. 8(d), a high number of A squats or B squats will not have strong influence in the short term (the

![Fig. 8. Interdependency of KPIs over Mamdani KPI, \( y_{h,j}^M(t) \)](image_url)
condition moves between the values 0.28 to 0.37). However, the number of B squats more negatively affects the rail health condition than the number of A squats. Fig. 8 shows the intuitive fact that rail condition gets worse with an increasing number of squats from A to B to C to RC.

In general, the number of A squats will not have a significant impact on the current rail health condition. However, in the long term, if not ground, A squats will evolve into severe defects. In order to capture this and other dynamic effects, the prediction model is used, and the global KPI is calculated over time and under different scenarios.

**Fuzzy Global KPI**

Relying on defined Mandani KPIs $y^M_{kj}(t)$, a fuzzy global indicator is calculated to give a KPI over growth scenarios in partition $j$

$$J_{kj}^{rail}(k) = \frac{\sum_{h \in \{h_1, h_2, h_3\}} \sum_{t=k}^{t+k+7} w_{k} w_{t} J^M_{kj}(t)}{\sum_{h \in \{h_1, h_2, h_3\}} \sum_{t=k}^{t+k+7} w_{k} w_{t}}$$

**Numerical Results**

**Interval Fuzzy Models**

This section summarizes the simulation results to predict the squats length. A data set of squat lengths collected from different track visits are used to evaluate performance of the squat growth model. Identification data and validation data for the interval fuzzy TS model are selected randomly, using 60% of the data for identification and 40% for validation (Fig. 9).

To optimize the number of clusters, models from 2 to 10 clusters are tested. For each number of cluster, the RMS of the prediction error is used to determine the best model. During training, tuning parameters of the confidence interval fuzzy model are considered the same for the lower and upper fuzzy bounds. The idea is to obtain the optimum parameter $\alpha$ that results in a minimum number of data points outside the band where the band is as narrow as possible. Fig. 10(a) depicts the Pareto front of the normalized area of the band versus the normalized number of data points outside the band of $\alpha$ ranging from 0 to 40. Fig. 10(b) shows how $\alpha$ behaves in terms of area of the band. As shown in the Fig. 10(b), the area will reach a maximum value if $\alpha$ equals to 32.

In reality, the variance of the worst-case scenario is much larger than the best-case scenario; thus the assumption of a fixed $\alpha$ must be relaxed. Using full trajectories of different squats, ad hoc $\alpha^h$ and $\alpha^b$ were obtained to better fit the dynamics. The use of an interval fuzzy model for prediction is presented in Fig. 11, with a selected $\alpha = 1.5$ from the Pareto front and modified parameters $\alpha^h = 0.32 \times \alpha$ and $\alpha^b = 1.7 \times \alpha$. The squat length starts from a small defect in 8 mm to a severe squat in 60 mm. An important characteristic is when the predictive model reaches the highest bound in 60 mm. This happens for squats of 48 mm for the one-step-ahead prediction (within 6 months), and it will happen for squats of 18 mm in the case of four-steps-ahead prediction (within 24 months). For testing purposes, this model was evaluated

where $J_{kj}^{rail}(k) =$ fuzzy global indicator; $w_k =$ growth weight per scenario; and $w_t =$ weight exponentially showing time effect on the KPIs. In this way, different KPIs are aggregated into a single one that captures both stochasticities and evolution over time.

Fig. 9. Validation and identification data for the squat lengths; the trajectories from Li et al. (2010) were used as test data.

Fig. 10. (a) Pareto front of number of data point outside versus area of the band; (b) area of band as function of alpha.
with another data set of the trajectories presented in Li et al. (2010). All of them are contained within the interval model.

**Fuzzy Global KPI for Track Health Condition**

The full track of the Meppel-Leeuwarden is used to show the proposed methodology. Fig. 12 shows a simple map of the track and the four partitions \( j_1, j_2, j_3, \) and \( j_4 \). The partitions can be adapted according to the maintenance plans or other design considerations. The partitions in this paper are all approximately 10 km long, except the last one which is 15 km long. Meppel is at kilometer 105, Leeuwarden is at kilometer 150, and the partitions are defined between the kilometers as \( x_{j_1} = 105, \ x_{j_2} = 115, \ x_{j_3} = 125, \ x_{j_4} = 135, \) and \( x_{j_5} = 150 \).

Fig. 13 shows the different KPI squat numbers over the four-steps-ahead prediction when no maintenance is performed. All the cases are calculated for the slow, average, and fast scenarios. In Fig. 13(a), the number of A squats tends to reduce over time because they are becoming B squats. In Fig. 13(b), the number of B squats increases because of the A squats becoming B squats, but after \( t = 12 \) the number of B squats decreases because most of them are becoming C squats. When no corrective maintenance is performed, it can be seen from Fig. 13(c) that after \( t = 12 \) a large number of C squats are in the track (worst-case scenario), which is a very expensive situation because the only solution will be to replace the rails. In Fig. 13(d), it is possible to see the moment when operational risk locations start to appear, indicating that maintenance should be done before the worst-case scenario indicates their appearance.

Fig. 14(a) shows how potential risk squats will start to appear over time. Fig. 14(b) shows the KPI related to density of B and C squats. As seen in Fig. 14(a), the first squats with high potential risk of derailment, RC squats, appear for the worst-case scenario at \( t = 12 \), in four kilometer positions \{130.9, 132.0, 132.5, 133.0\}. Three of those four locations were already detected at \( t = 0 \) in Fig. 14(b), while all of them are already present in the BC squat density signal at \( t = 6 \) for all scenarios. This means that within the first 12 months, the infrastructure manager is expected to take actions to prevent risk of derailment.
Fig. 15 collects all the scenarios and signals over the whole prediction horizon to indicate a single global fuzzy KPI for each track partition. Three cases are considered, no maintenance, grinding at $t = 0$, and local rail replacement at $t = 0$ for each severe squat. Maintenance can considerably improve the rail health condition, but to be fully efficient a combination of both grinding and replacement is necessary. After the maintenance operations, the condition is in the average condition range, where the potential risk of derailment is considerably lower during the prediction horizon. The following result allows the infrastructure manager to decide how to manage the rail in the future at each track partition. As in the case of the absence of maintenance operation, a cost of 0 Euro with the clear consequence of the bad rail health condition. In the case of the grinding effect and the replacement effect, the results can be

© ASCE 04017006-11 J. Infrastruct. Syst.

applied as an effective factor for cost analysis of the track maintenance plan.

Conclusion and Future Research

In this paper a condition-based monitoring methodology was developed for a type of surface defect in the rail called squats. This methodology is employed to construct an interval-based TS fuzzy prediction modeling in order to monitor the track condition over the maintenance time horizon per track partition.

The idea of using an interval fuzzy models is to capture all the possible growth scenarios. Based on the interval fuzzy models for squats, a condition-based methodology for railway tracks is proposed using different KPIs defined at a track partition level, allowing the grouping of defects located in a given track partition. In the methodology, the number and density of squats are considered over a prediction horizon under three different scenarios, slow, average, and fast growth. Then, to facilitate visualization of the rail health condition and to ease the maintenance decision process, a fuzzy global KPI is proposed based on fuzzy rules for each partition that combine the different KPIs over the prediction horizon and scenarios. Hence, the proposed methodology adds value by defining fuzzy global KPIs that are predictable over time to facilitate maintenance decision making of the rail. As an example, the KPIs obtained are presented for the Meppel-Leeuwarden track.

In further research, the study will be oriented into an optimization-based methodology to effectively reduce lifecycle costs and to fit the methodology closely to the real-life maintenance operations. The use of new predictive and robust KPIs defined for different parties will be based methodology closely to the real-life maintenance operations. The use based methodology to effectively reduce lifecycle costs and to fit the lifecycle model to control rolling contact fatigue.

Acknowledgments

Research sponsored by the Netherlands Organisation for Scientific Research (NWO)/ProRail project “Multiparty risk management and key performance indicator design at the whole system level (PYRAMIDS),” Project 438-12-300, which is partly financed by NWO and by the Collaborative Project H2020-MG-2015-2015 GA-398 636237b Needs Tailored Interoperable Railway—Needs Tailored Interoperable Railway (NeTIRail-INFRA). The authors also thank Meysam Naemi for the discussions about the effect and modeling of grinding.

References


Li, Z., Molodova, M., Zhao, X., and Dollevoet, R. (2010). “Squat treatment by way of minimum action based on early detection to reduce life cycle costs.” Proc., Joint Rail Conf., Rail Transportation Division, Urbana, IL.


