Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

Nunez, Alfredo; Jamshidi, Ali; Wang, Hongrui

DOI
10.1109/TII.2018.2847736

Publication date
2018

Document Version
Accepted author manuscript

Published in
IEEE Transactions on Industrial Informatics

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

Alfredo Núñez, Senior Member, IEEE, Ali Jamshidi, Hongrui Wang, Student Member, IEEE

Abstract — This paper presents a Pareto-based maintenance decision system for rail welds in a regional railway network. Weld health condition data are collected using a train in operation. A Hilbert spectrum-based approach is used for data processing to detect and assess the weld quality based on multiple registered dynamic responses in the axle box acceleration measurements. The assessment of the welds is stochastic in nature and variant over time, so a set of robust and predictive key performance indicators is defined to capture the weld degradation dynamics during a given maintenance period. Using a scenario-based approach, two objective functions are defined, performance and the number of weld replacements. Evolutionary multi-objective optimization is employed to optimize the objective functions so that the trade-offs between performance and cost support decision-making for railway network maintenance. The results of the proposed methodology show that the infrastructure manager can localize field inspections and maintenance efforts on the area with the most critical welds. To showcase the capability of the proposed methodology, measurements from a regional railway network in Transylvania, Romania are employed.

Index Terms— Maintenance, Railway infrastructure, Acceleration measurements, Evolutionary multi-objective optimization

I. INTRODUCTION

Regional railway transport has always been challenging because traditional cost-benefit analyses normally suggest that railway transport is not economically feasible from an operations and maintenance perspective. This lack of economic feasibility can be explained by the relatively low demand and the dispersed nature of the populations in rural areas. However, from the social perspective, these regional railway lines are crucial for communities because they provide access to work and services [1]. Moreover, good access to transportation services can help to prevent the de-population of rural areas and reduce the concentration of capital, services and attractions in big cities, which will reduce urban traffic congestion and increase investment in areas such as agriculture and tourism. Hence, regional railways can be seen as a tool to decentralize opportunities over a territory when accounting for societal equality and to include regionalization in decision-making.

The main challenge faced by regional railways is that their operations and maintenance budget is constantly decreasing. Consequently, new technological solutions that can support the decision-making to expend scarce resources in a smart way are needed. To optimize maintenance costs, many intelligent monitoring systems have recently been used for railway infrastructures, which need big data analyses over the data collected [2], [3]. The range of sensing technologies has expanded rapidly, and sensor devices have simultaneously become cheaper. For example, networking technologies, WSNs, smart phones, accelerometers installed on trains, drones and video cameras have all developed and become less expensive [4], [5]. In this paper, extending the traditional concept of what a train does is proposed, which is transporting people or goods, to include performing monitoring tasks. In this way, the capabilities of trains can be fully used while providing information about the health conditions of the track. Therefore, axle box acceleration (ABA) measurements are used from a regular train in operation to provide information about the health conditions of rails. This information is useful for detecting shortsurface defects and corrosion and for determining the conditions of insulated joints and crossings.

The health conditions of rails can be estimated by tracing their degradation over time with the proper indicators [6], [7]. Degradation analyses allow infrastructure managers to be aware of critical locations by providing information about when degradation will reach a critical level, and this information can be used to mitigate the risk of a rail break. Rail degradation is mainly related to rail surface defects [8], and it is a stochastic process that changes over time and space. In this paper, analyzing weld health conditions based on multiple ABA measurements is focused. The detection and assessment of rail surface defects are normally realized by analyzing frequency-based features. Previously, methods such as the power spectrum density [9], short-time Fourier transform [10] and wavelet transform [11] methods have been adopted for detecting defects in wheels, axle bearings and rails. ABA
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

Signals from regional railways are particularly noisy and are affected not only by train speed and wheel conditions but also by a less accurate GPS location. To circumvent the usage of a fixed signal, this paper employs the Hilbert-Huang Transform (HHT) based on an adaptive signal that can be associated with physical excitations [12]. The Hilbert spectrum of ABA is employed as an indicator of the weld health conditions.

The use of ABA for assessing welds is inherently stochastic, so a set of robust and predictive key performance indicators (KPIs) is defined to capture the weld’s degradation dynamics during a given maintenance period. Using a scenario-based approach, two objective functions are used, the performance and the number of weld replacements. The latter function is related to maintenance costs. To facilitate decision-making, a methodology is proposed based on multi-objective optimization. Although the proposed methodology is applied to the rail welds, the structure can be generalized for other types of rail surface defects. Fig. 1 shows the four major steps of the methodology. First, an operational train is used to collect ABA signal data. Then, the health conditions of the rail are estimated by using the Hilbert spectrum of the ABA signal at the rail welds. Next, KPIs are established to characterize the severity of the detected welds. As the KPIs include the uncertainty induced by measurement-related variables, such as vehicle speed, condition of wheels, etc., a stochastic analysis is presented in the second step to provide robust KPIs. Moreover, different degradation scenarios are introduced to include the predictability of the KPIs. In the last step, a maintenance decisions optimization problem (MDOP) is solved to obtain Pareto-optimum maintenance solutions.

II. METHODOLOGY

A. ABA measurement system

There are different methods used to diagnose the condition of rail defects, including ultrasonic measurement [13], eddy current testing [14], and image recognition [15], among other technologies. Each method has different advantages and disadvantages. These methods help to check the crack development underneath rail surface defects (ultrasonic and eddy current) and to visually monitor rails (rail video images). However, these methods are not able to capture the dynamic response of welds under actual operation. By having a dynamic response, valuable information about welds can be acquired. A technology capable of (1) shortwave irregularity detection and (2) the capture of the dynamic response of welds is needed [16]. ABA measurement systems with these capabilities have been reported in different countries: the detection of lateral and vertical track irregularities in Korea [17]; the evaluation of wheel load fluctuation and rolling noise in Japan [18]; the detection of corrugations in Poland [19] and Italy [20]; the detection of rolling contact fatigue defects, damaged welds, and insulated joints in the Netherlands ([21], [22]); and the analysis of vertical track geometry in Spain [23]. ABA system implementation has different advantages: (1) ABA is a low-cost measurement system compared to other types of detection methods. (2) The ABA sensors are easy to maintain and (3) ABA can be implemented on in-service operational trains. Moreover, there is (4) the possibility to detect seed rail defects with no need for expensive and complex instrumentation and (5) the ability to estimate the severity level of the dynamic contact force. Additionally, this technology is suitable for regional networks because it is portable and cost-effective.

In this paper, an axle box acceleration (ABA) measurement system is used to detect welds and estimate conditions. In total, 16 accelerometers are mounted on the axle boxes (vertically and longitudinally). For mounting the sensors, small mounting studs have to be glued in dry conditions on each axle box.

Furthermore, a GPS antenna for positioning is used and installed on the roof of the train. The positioning system, beside the GPS antenna, uses tacho count which is able to detect pulses
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

at approximately 1 MHz. A schematic view of the ABA system is described in Fig. 2.

Welds highly excite the vertical acceleration of ABA systems [24]. The location of welds is well known by the infrastructure manager. If not, their locations can easily be obtained by field inspection or video camera systems. However, relating the welds to the ABA signals is not an easy task in the time domain, but it is possible in the frequency domain. To study the relation between the actual welds in the track with the welds detected by the ABA measurement system, a training and validation process is performed. First, a track field inspection is carried out to relate different welds detected by the ABA signals to the actual welds. Based on the examples for training, a validation to other welds is performed by examining similar frequency responses. Finally, predictions are performed for the rest of the track.

B. Rail welds

Welded rails have considerably improved the problems of rail wear and overloading, which are the main causes of rail breaks. Most field welding in the railway industry is carried out using the aluminothermic technique. Such welds are primarily associated with rail replacement, the installation of insulated rail joints and track construction activities. However, rail welds are typically subject to complex loading and high stresses at the rail head as train wheels pass and they are exposed to cyclic fatigue loading [25]. Thus, the rails with damaged welds produce vibrations that cause noise and ride discomfort. Over time, they are a safety concern because they contribute to less stable support for moving trains. Crack propagation due to brittle fractures is one of the main factors that lead to a severely defective weld [26]. In regional networks, brittle fractures accelerate the further degradation of defects due to poor maintenance operations.

C. Hilbert spectrum of the ABA signal

The Hilbert spectrum is obtained by combining empirical mode decomposition (EMD) with the Hilbert transform [27], namely, the HHT. EMD is a data-driven algorithm that adaptively decomposes a signal \( \{ \tau(t), t \in [0,T] \} \) into a number \( N_{IMF} \) of intrinsic mode functions (IMFs) \( v_r(t) \), \( r = 1, 2, \ldots, N_{IMF} \) and a residual \( r(t) \). The IMFs are selected using a sifting process that only terminates when a stopping criterion has been satisfied. In detail, the EMD algorithm can be described by Algorithm 1. The decomposition starts with the original signal \( \tau(t) \). The upper and lower envelopes, \( e_u(t) \) and \( e_l(t) \), of the signal are obtained by connecting all the signal maxima and minima, respectively, using spline interpolation. Then, the average envelope signal \( e_a(t) \) can be computed to obtain the estimated IMF \( v^{(i)}(t) \) on line 9. At this point, the estimated IMF will be checked by the IMF termination criterion.

\[
S(i) = \frac{1}{N_{IMF}} \sum_{i=0}^{N_{IMF}} \frac{|v_r^{(i-1)}(t) - v_r^{(i)}(t)|^2}{|v_r^{(i-1)}(t)|^2} < \varepsilon
\]

where \( \varepsilon \) is a positive number that typically ranges from 0.2 to 0.3. If the criterion is not fulfilled, the sifting process continues with the estimated IMF; otherwise, the \( r \)-th IMF is obtained as \( v_r(t) = v_r^{(i-1)}(t) \), and the global termination criterion will be checked to see if all the IMFs are obtained. The global termination criterion judges if the updated estimated IMF \( v_r(t) \) on line 13 has only one pair of extrema. If not, the sifting process goes on; otherwise, the signal residual \( r(t) \) can be finally obtained.

Algorithm 1: General steps of EMD

Input: \( \tau(t) \) (original signal)
Output: \( v_r(t) \) (IMFs), \( r(t) \) (residual)
1. \( \ell \leftarrow 1; \)
2. \( i \leftarrow 1; \)
3. \( v_r^{(i)}(t) \leftarrow \tau(t); \)
4. while global termination criterion not fulfilled do
5. while IMF termination criterion not fulfilled do
6. \( \{ e_u(t), e_l(t) \} \leftarrow \text{FindEnvelopes}(v_r^{(i)}(t)); \)
7. \( e_a(t) \leftarrow \frac{[e_u(t) + e_l(t)]}{2}; \)
8. \( i \leftarrow i+1; \)
9. \( v_r^{(i)}(t) \leftarrow v_r^{(i-1)}(t) - e_a(t); \)
10. \( v_r(t) \leftarrow v_r^{(i)}(t); \)
11. \( \ell \leftarrow \ell+1; \)
12. \( i \leftarrow 1; \)
13. \( v_r^{(i)}(t) \leftarrow \tau(t) - \sum_{k=1}^{i} v_r(t); \)
14. \( r(t) \leftarrow v_r^{(i)}(t); \)
15. return \( v_r(t), r(t); \)

As a result, the sum of all IMFs and the residual can be used to reconstruct the original signal as follows:

\[
\tau(t) = \sum_{i=1}^{N_{IMF}} v_r(t) + r(t)
\]
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

The extracted IMFs are considered as inherent signal modes contained in the original signal that can be associated with a physical meaning or cause. The iterative sifting process automatically decomposes the original signal into IMFs that possess different frequency bands decreasing from 1 to \( \ell = N_{\text{IMF}} \). Thus, EMD can be applied to ABA signals to extract signal features by dividing the full measurement frequency band into several sub-bands of interest. The next step of HHT is to compute the instantaneous frequency of the extracted IMFs. Concretely, the analytic form of IMFs is first computed using the Hilbert transform:

\[
z_\ell(t) = v_\ell(t) + i : H[v_\ell(t)] = a_\ell(t)e^{i\theta_\ell(t)}
\]
where \( H[d_\ell(t)] \) denotes the Hilbert transform of IMF \( v_\ell(t) \).

The following relationship is established:

\[
a_\ell(t) = \sqrt{v_\ell^2(t) + H[v_\ell(t)]^2}
\]
\[
\theta_\ell(t) = \arctan \left( \frac{H[v_\ell(t)]}{v_\ell(t)} \right)
\]

Therefore, the instantaneous frequency \( \omega_\ell(t) \) of IMF \( v_\ell(t) \) can be obtained as

\[
\omega_\ell(t) = \frac{d\theta_\ell(t)}{dt}
\]

The instantaneous frequency represents the degree of frequency variation in the IMF \( v_\ell(t) \) at time \( t \). This feature is similar to the variations of the wavelet coefficient that reflect the signal variations within a certain frequency band. However, in theory, when using the IMFs that are extracted from the signal itself instead of the signals reconstructed from a mother wavelet, the frequency variations are more closely related to the physical nature of the signals, namely, the excitations of ABA at the welds. Finally, the Hilbert spectrum of the signal \( \tau(t) \) is defined as a function of instantaneous frequency and time as follows:

\[
S(\omega_\ell, t) = \text{Re} \sum_{\ell=1}^{N_{\text{IMF}}} c_\ell \{ a_\ell(t)e^{i\theta_\ell(t)} \}
\]

where \( \text{Re} \) denotes the real part of the operator for a complex signal. The Hilbert spectrum of a single IMF \( v_\ell(t) \) represents the energy variations of the instantaneous frequency with time, and it is given by

\[
S_\ell(\omega_\ell, t) = \text{Re} \{ a_\ell(t)e^{i\theta_\ell(t)} \}
\]

The Hilbert spectrum can provide indicators for many aspects of application in the sense of a power spectrum. Since the IMFs have different frequency bands that will be inherited by their Hilbert spectra, the employment of Hilbert spectra becomes crucial to determine on which frequency ranges the indicator is based. Thus, a linear combination of the Hilbert spectra of IMFs is proposed as the KPI for rail defect detection as follows:

\[
\gamma(t) = \sum_{\ell=1}^{N_{\text{IMF}}} c_\ell \text{Re} \{ a_\ell(t)e^{i\theta_\ell(t)} \}
\]

where \( c_\ell \) (\( c_\ell \geq 0 \)) is the weight coefficient for the spectrum of the \( \ell \)-th IMF with \( \sum c_\ell = 1 \). For ABA signals measured under different conditions or for the detection of different types of defects with varying frequency responses, the distribution of weight coefficients varies depending on the indicative frequency band.

D. Description of the KPIs

When performing multiple ABA measurements, the data related to a given weld are subject to a different set of stochasticity. The most influencing uncertainties are related to the way the wheel hits the defect. Sources of stochasticity come from sensors, train speed, wheel quality, the track and the welds. There are various strategies to include uncertainties in the optimization [28]. A scenario-based approach is used to cope with the uncertainties related to the multiple measurements, \( m = 1, \ldots, M \), and robust KPIs are defined. Three different severity scenarios are proposed: slow, average and fast severity scenarios. The KPI for the average severity scenario, \( h_2 \), can be expressed as

\[
\gamma^h_2(k) = \frac{1}{M} \sum_{m=1}^{M} \gamma^m_2(k)
\]

where \( \gamma^m_2(k) \) is the KPI using measurement \( m \) at the weld \( w \) (located at position \( x_w \)). The other two KPIs are defined as

\[
\gamma^f_2(k) = \max \{ \gamma^m_2(k) \}, m = 1, 2, \ldots, M
\]
\[
\gamma^s_2(k) = \min \{ \gamma^m_2(k) \}, m = 1, 2, \ldots, M
\]

where \( \gamma^f_2(k) \) and \( \gamma^s_2(k) \) are the fast and slow severity scenarios for the weld \( w \), respectively.

E. Degradation model

The generic degradation process of a rail infrastructure component is stochastic. The degradation can be stopped by applying maintenance actions. In this paper, only replacement is considered. The maintenance actions per weld are binary: the weld is replaced or not replaced. For a set of welds, the optimization variable at time step \( k \) is \( u(k) = [u_1(k), \ldots, u_w(k)] \in \{0, 1\}^W \), where \( W \) is the number of welds. The following model is proposed to describe the effects of the maintenance actions at a weld located at \( x \)

\[
\gamma^s(k+1) = F_s^e \{ \gamma^s(k), u_s(k) \}
\]
\[
\gamma^s(k+1) = F_s^i \{ \gamma^s(k) \}
\]

if \( u_s(k) = 0 \) no maintenance

if \( u_s(k) = 1 \) replacement
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

For other types of defects, such as squats or corrugation, (12) should be extended to include grinding or other maintenance actions. To quantify the effects of maintenance on the degradation over time, predictions over the complete maintenance horizon should be conducted [29] for the three degradation scenarios, namely, $h = h_1, h_2, h_3$, by solving (12) recursively

$$
\gamma^w_h(k+t) = F_h\{\gamma^w_h(k+t-1), u_w(k+t-1)\}, t = 1, ..., N_p
$$

(13)

The predicted health condition of the weld $w$ could be considered as an interval prediction, $[\gamma^w_h(k+t), \gamma^w_h(k+t)]$, and it is obtained from the KPIs measured from the ABA signal.

F. Optimization of rail maintenance decisions

Because the objective functions of maintenance operations are usually conflicting, i.e., a solution that optimizes one objective may not optimize others, the use of multi-objective optimization is considered. When there are multiple objectives that vary over time, optimization by tuning the weights [30] is formalized so that the trade-off between Pareto optimal solutions is known as the Pareto front. In this section, a real-life case study is provided to show the capability of the proposed methodology. The regional track Bartolomeu-Zârmești in Romania is used as the case study (Fig. 3). A compactRIO measurement computer is used to collect 16 ABA channels (vertical and longitudinal) and store in the FPGA (field programmable gate array) at 25600 Hz. The tacho and GPS data are also stored in the FPGA. Next, a data block composed of 200 samples is transferred to a computer. The ABA data processing is later performed offline after collecting all the data from the measurement campaign.

The measurement data size per kilometer depends on the

$$
J^w_h(k) = \min_{\{u(k)\}} \{J_1(u(k)), ..., J_5(u(k))\} \text{ subject to } u(k) \in U
$$

(14)

Let’s assume minimization of the $G \geq 2$ objective functions simultaneously. The solution to (14) is known as the Pareto optimal set. A solution $u^*_h(k)$ is said to be Pareto optimal if and only if another solution $u(k)$ does not exist, such that

$$
J^w_h(u(k)) \leq J^w_h(u^*_h(k)) \forall g = 1, ..., G
$$

(15)

$$
J^w_h(u(k)) < J^w_h(u^*_h(k)) \text{ for at least one } g \in \{1, ..., G\}
$$

(16)

The set of all objective function values corresponding to the Pareto optimal solutions is known as the Pareto front. In this paper, two objective functions are considered, the performance and cost functions. For the first objective function, the contribution of the weld $w$ to the degradation scenario $h$ can be expressed as

$$
J^w_h(k) = \sum_{i=1}^{N_p} \alpha_i \gamma^w_h(k+i)
$$

(17)

where $\gamma^w_h(k)$ is the KPI at the time step $k$ of weld $w$ under severity scenario $h$ and $\alpha_i$ is a weight applied across the time predictions of the KPIs. According to the predictive and robust KPIs provided with the ABA signals, the objective function $J^w_h$ is defined to cover all the welds and all the scenarios over the maintenance period $N_p$.

$$
J^w_h(k) = \sum_{i=1}^{N_p} \alpha_i \gamma^w_h(k+i)
$$

(18)

where $J^w_h$ is the KPI at the time step $k$ and $\alpha_i$ is the weight per scenario. The second function $(J^w_h)$ is the number of replacements and is directly related to the budget availability and time constraint determined based on the decisions of the infrastructure manager.

$$
J^w_h(k) = \sum_{i=1}^{N_p} (1 - u_i(k+i))
$$

(19)

Furthermore, a decision window $(D_w)$ is used based on the given rail replacement length. If more than one weld can be covered by one replacement operation, it will be grouped in advance into a cluster $W_i$. The replacement operation will then improve the conditions of all the welds in the cluster with only one replacement operation. In this case, the following rule is applied:

$$
\text{if } u_i(k)=1, w \in W_i \Rightarrow u_i(k)=0, \forall \omega \in W_i \setminus \{w\} \land \gamma^w_{i+1} = F_h(\gamma^w_i, 1), \forall \omega \in W_i
$$

(20)

where $W_i$ is the $i$-th cluster of welds.

Algorithm 2 shows the general architecture of an MDOP. To solve the MDOP problem, the six optimization algorithms are selected and described in Section III through two case studies.

<table>
<thead>
<tr>
<th>Algorithm 2: general architecture of an MDOP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong>: $N$ (population size), $NR$ (number of reference points and archive size)</td>
</tr>
<tr>
<td><strong>Output</strong>: $P$ (final population)</td>
</tr>
<tr>
<td>1 $P \leftarrow$ Random Initialize ($N$);</td>
</tr>
<tr>
<td>2 $R \leftarrow$ Uniform Reference Point($NR$);</td>
</tr>
<tr>
<td>3 $A \leftarrow$ $P$;</td>
</tr>
<tr>
<td>4 $R' \leftarrow$ $R$;</td>
</tr>
<tr>
<td>5 while termination criterion not fulfilled do</td>
</tr>
<tr>
<td>6 $P' \leftarrow$ Mating Selection ($P$, $R'$);</td>
</tr>
<tr>
<td>7 $O \leftarrow$ Variation ($P'$, $N$);</td>
</tr>
<tr>
<td>8 $[A,R'] \leftarrow$ Ref Point Adapation;</td>
</tr>
<tr>
<td>9 $P \leftarrow$ Environmental Selection;</td>
</tr>
<tr>
<td>10 return $P$.</td>
</tr>
</tbody>
</table>

III. NUMERICAL RESULTS AND DISCUSSION

In this section, a real-life case study is provided to show the capability of the proposed methodology. The regional track Bartolomeu-Zârmești in Romania is used as the case study (Fig. 3). A compactRIO measurement computer is used to collect 16 ABA channels (vertical and longitudinal) and store in the FPGA (field programmable gate array) at 25600 Hz. The tacho and GPS data are also stored in the FPGA. Next, a data block composed of 200 samples is transferred to a computer. The ABA data processing is later performed offline after collecting all the data from the measurement campaign.

The measurement data size per kilometer depends on the
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

Train speed. The sampling rate is a fixed 25600 Hz. The average data size for the whole track (which covers 26.9525 km) is 273.44 MB/km for one measurement run. For a measurement campaign, ten different measurement runs are normally carried out over the track to assure the repeatability and reduce signal noise. The size of the full data is approximately 73698.916 MB.

The weather condition can differently influence rail conditions. High temperatures can increase the risk of incidents related to rail buckles. An excessive amount of snow and ice can cause huge problems for network performance: (1) a frozen catenary might damage the overhead wire and block the train from receiving electrical energy, or (2) a track switch could become dysfunctional when the temperature drop is extreme. Moreover, the welds can be ruptured in a rapid variation of temperature due to thermal expansion and contraction. Although it is an important topic, to the best knowledge of the authors, there is no study on how extreme weather conditions influence the quality of ABA measurement.

In the case study, the ABA system was installed on a passenger train. The speed was varied from 0 km/h at stations, up to 80 km/h. The signals collected at nearly 0 km/h do not contain the necessary excitation for analysis of welds. Signals approximately 70 km/h (higher than 60 km/h) have been selected for the processing. The coverage in the track of the case study is approximately 80% of the infrastructure. For the rail pieces at the train speed below 70 km/h (most of them at stations or near them), quantitative relationships with the signature tunes and maximum ABA should be incorporated using a regression model to make full use of the data collected.

Fig. 4 shows the IMFs decomposed from a piece of the track measured when the train is passing over a weld. The weld located at around the midpoint of the signal segment results in two peaks in ABA. The IMFs numbered from IMF1 to IMF9 show the modes of ABA in different frequency bands in a descending order. The mode that reflects the excitation of the weld can be identified as IMF4, where the two peaks attributed to the weld are almost zero. Thus, this mode can be considered as the result of EMD separating the vibration response of the weld from other sources of excitation. In this paper, the frequency band of the IMF4 of the ABA signal is employed as the best representation of the frequency response for welds. As given in (8), the weight coefficients $c_{\lambda}$ are thus determined to be $c_4 = 1$ and $c_{\lambda\neq4} = 0$. With this method, all the weld-like impacts will be detected, which are interesting locations for monitoring and rail replacement. Next, the amplitude of the Hilbert spectrum of the decomposed ABA is used to describe the severity of the welds.
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

Four welds were selected from the full track of the case study first to give a detailed description of the steps. In Fig. 5, the ABA signals of four welds are analyzed. As seen in Fig. 5, four measurements are used to show the impact of the welds on the ABA signals associated with their corresponding IMF4, and the amplitudes of the Hilbert spectrum are presented. All the measurements are obtained from the vertical acceleration signal of the left rail. Moreover, the accuracy of the welds’ positioning is weak in the regional network due to the geographic locations. In the case study of this paper, by defining a set of reference points in the ABA signals, together with tacho and GPS, the positioning of the welds is obtained with errors between 1 to 10 meters. In the welds of Fig. 5, the data positioning accuracy is approximately 1 meter when compared to ABA peaks due to welds in different measurements. The obtained amplitudes are used to indicate the severity of the weld damage.

In Fig. 6 (a), the KPIs are estimated, and it is possible to see the degradation behavior of the welds. Relying on the physical understanding of how a surface rail defect grows, a polynomial regression model is fitted using the least-absolute residual method to represent the stochasticity of the growth [31]. Fig. 6 (b) validates the maintenance decision results. As seen in the ranking, Weld 2 has the highest values in all the scenarios out of the four welds, and Weld 1 has the second highest rank. However, although Weld 3 has higher values than does Weld 4 in both fast and slow scenarios ($h_1$ and $h_2$), the average scenario of Weld 4 shows more criticality in comparison with Weld 3 (higher value). This plot can be used later to determine the extent that the optimization results are in agreement with the expected rankings of the welds.

In Case study one, four replacements are possible. In general, the number of replacements should be determined according to the decision of the infrastructure manager. The decision can be explained in terms of (1) budget limit and (2) maintenance time slots. In Case study one, the decision window $D_w$ is shorter than the distance between the welds, so only one weld is replaced by one replacement operation (no clustering needed).

Fig. 7 shows the Pareto front for four different MDOPs: 1) $h_1$ is a fast severity scenario, 2) $h_2$ is an average severity scenario, 3) $h_3$ is a slow severity scenario. Additionally, $\alpha_k$ is an exponentially growing parameter for the global KPI, which is the time indication in the prediction model. In addition, $\alpha_k$ is 0.8, 0.6 and 0.4 for the fast, average and slow severity scenarios, respectively. The black arrow in Fig. 7 when $J_2=2$ (two replacements allowed) shows that once the severity scenario changes from $h_1$ (fast degradation scenario) to $h_3$ (slow degradation scenario), the first objective function ($J_1$) gets a lower value because the degradation predictions are more “optimistic”. The selection of the degradation scenarios will influence the decision-making. A simple calculation shows an average improvement of 49.33% in the rail health condition when two replacements are made rather than zero.

<table>
<thead>
<tr>
<th>$J_2$</th>
<th>$H$</th>
<th>$J_1$</th>
<th>Maintenance decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$h_1\bigodot$</td>
<td>6.981</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_2\oplus$</td>
<td>4.665</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_3\otimes$</td>
<td>4.061</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td>1</td>
<td>$h_1\bigodot$</td>
<td>6.467</td>
<td>0 1 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_2\oplus$</td>
<td>2.851</td>
<td>0 1 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_3\otimes$</td>
<td>2.369</td>
<td>0 1 0 0</td>
</tr>
<tr>
<td>2</td>
<td>$h_1\bigodot$</td>
<td>2.465</td>
<td>1 1 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_2\oplus$</td>
<td>1.360</td>
<td>1 1 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_3\otimes$</td>
<td>1.099</td>
<td>1 1 0 0</td>
</tr>
<tr>
<td>3</td>
<td>$h_1\bigodot$</td>
<td>1.208</td>
<td>1 1 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_2\oplus$</td>
<td>0.6764</td>
<td>1 1 0 0</td>
</tr>
<tr>
<td></td>
<td>$h_3\otimes$</td>
<td>0.5339</td>
<td>1 1 0 0</td>
</tr>
<tr>
<td>4</td>
<td>$h_1\bigodot$</td>
<td>0</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td></td>
<td>$h_2\oplus$</td>
<td>0</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td></td>
<td>$h_3\otimes$</td>
<td>0</td>
<td>1 1 1 1</td>
</tr>
</tbody>
</table>

Fig. 5. Photos of welds for four different measurements with (a) the corresponding ABA signals in the vertical direction, (b) the IMF4 decomposed from the ABA signal, and (c) the Hilbert spectrum amplitude of IMF4.

TABLE I

MAINTENANCE DECISION RESULTS FOR EACH DEGRADATION SCENARIO
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

Table 1 presents the optimal decisions per scenario. The maintenance decision was equal to 1 when a replacement was made \((u=1)\) and 0 when no replacement was made \((u=0)\).

From the Table 1, Weld 1 and Weld 2 are the most important welds and are the most common candidates for replacement. This decision is in agreement with the results shown in Fig. 5, where the peaks of ABA and the amplitude of the Hilbert spectrum are much higher for these two welds. The maintenance decisions according to the global KPI are also analyzed. Weld 2 is selected when one replacement is allowed \((J_1=3.542, J_2=1)\). For two replacements \((J_1=1.793, J_2=2)\), Weld 1 and Weld 2 are selected. For three replacements \((J_1=0.8832, J_2=3)\), Weld 3 is included in addition to Weld 1 and Weld 2 in slow and fast scenarios, while in an average scenario, Weld 4 is added to the maintenance decision. This finding is in agreement with Fig. 6 (b). The analysis of the global KPI highlights the influences of Weld 1 and Weld 2 on the replacement decisions.

B. The whole track study

To use the proposed methodology to study the full track from Bartolomeu to Zâmești, the defect detection algorithm is run for the full track. Over the track, 2128 welds are detected in total. Considering only the welds detected at speeds above 60 km/h, the number of welds needed for optimization is 1849. The detection of welds is presented in Fig. 8 (a), which depicts their KPIs in terms of the track position in kilometers for a single measurement. Empty severity signals imply that the welds in those locations are not analyzed. In Fig. 8 (b), a close view of the track between 1.8 km and 2.4 km is shown. In Fig. 8 (b), the effects of the decision window \(D_w\) can be observed. All the damaged welds within the decision window are grouped to indicate the cluster severity. The cluster severity is the summation of the weld severities. In this paper, a decision window of 15 m is selected, but this window can be modified by the infrastructure manager if a reasonable alternative is found.

Then, the MDOP problem is solved and used by the infrastructure manager to make optimal maintenance decisions. A performance analysis is done to evaluate the reliability of the solutions. Six multi-objective optimization algorithms are compared (1) ARMOEA [32]; (2) NSGA-II [33]; (3) SPEA2 [34]; (4) GrEA [35]; (5) RSEA [36] and VaEA [37].

Although different MOEAs have been verified on different types of benchmark MOPs, some recent studies have noted that the performance of an MOEA can strongly depend on Pareto front shape of the problems. [38]. Hence, some MOEAs are more capable of dealing with regular Pareto fronts, whereas others are specifically tailored for problems with irregular Pareto fronts. For the maintenance optimization problem in this paper, a binary-based codification is used, i.e., on/off maintenance plan (replacing/not replacing). Considering the nature of a stochastic integer optimization problem, the results of the convergence and diversity of the Pareto fronts are obtained. The algorithm RSEA functions better than do the others algorithms according to Table II.

All six algorithms are run using the default setting and 30 independent runs. Populations of 500 solutions and 50000 generations were considered. The maximum number of rail replacements is set to 25 for demonstrations (a reasonable number that should fit the capacities of the replacement operations and budget). The results shown in Fig. 9 are achieved using the platform PlatEMO [39] and MATLAB 2017b at a desktop computer (2.60 GHz Intel Core i12, 32 GB)
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration of RAM). The Pareto fronts are shown in Fig. 9, and the true Pareto front was obtained from multiple runs over all the algorithms.

As shown in Fig. 9, the RSEA and SPEA2 algorithms approximate the true Pareto front better than the other algorithms. The RSEA converged to the true Pareto front more quickly and with less iteration. To evaluate these algorithms, the metrics that capture the convergence and diversity of the Pareto front approximation delivered by various algorithms can be analyzed. In this paper, the Normalized Hyper-volume (NHV) is included for performance comparisons. This metric is the only single set quality measure that is known to be strictly monotonic with regard to Pareto dominance; whenever a Pareto set approximation entirely dominates another one, the indicator value of the former approximation will be better. In addition, another metric is considered to include the effects of diversity analysis on the performances of the peer algorithms, namely, the spacing (Sp), which is widely accepted in the literature. The mean and standard deviation (Std.) listed in Table II are statistical results of the same algorithm that was run independently 30 times for the same test problem. The computation time is also calculated for each algorithm. According to the results, although AEROMA gives the fastest runtime, RSEA performs better in terms of convergence and diversity analysis.

Fig. 8. (a) Detection results for the track between Bartolomeu and Zănești. (b) A piece of the track is highlighted in red to show how a decision window ($D_w$) is defined over the track.

Fig. 10 shows the Pareto optimal solution for the maximum number of replacements (25 in total). Fig. 10 (e) depicts the optimal decision derived by the RSEA. Replacing pieces of the track appears to be critical based on the RSEA (similar to in 10.63 km), but this replacement decision is missing when the results of other algorithms are considered. Moreover, some common rail pieces were found that needed replaced, such as the track pieces between 16.92 km to 18.01 km.

The evolution of track deterioration compared to urban railway networks is lower because the railway traffic in regional railways is considerably lower. In the case study line, maintenance plans are organized regularly. A total of 25 locations for the rail replacements are considered to rank as the most important locations.

Fig. 9. Pareto front results of the three multi-objective optimization methods used to optimize rail maintenance between Bartolomeu and Zănești. The black circles show the true Pareto front, while the black cross shows the approximate Pareto front.

<table>
<thead>
<tr>
<th>ALGORITHMS</th>
<th>NHV</th>
<th>Spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMOEA</td>
<td>0.87151</td>
<td>0.52538</td>
</tr>
<tr>
<td>Std.</td>
<td>0.00018</td>
<td>0.04732</td>
</tr>
<tr>
<td>SPEA2</td>
<td>0.87197</td>
<td>1.05032</td>
</tr>
<tr>
<td>Std.</td>
<td>0.00016</td>
<td>0.15213</td>
</tr>
<tr>
<td>NSGAII</td>
<td>0.87187</td>
<td>0.85797</td>
</tr>
<tr>
<td>Std.</td>
<td>0.00016</td>
<td>0.18143</td>
</tr>
<tr>
<td>GrEA</td>
<td>0.87203</td>
<td>0.99232</td>
</tr>
<tr>
<td>Std.</td>
<td>0.00014</td>
<td>0.18264</td>
</tr>
<tr>
<td>RSEA</td>
<td>0.87229</td>
<td>1.47572</td>
</tr>
<tr>
<td>Std.</td>
<td>0.00012</td>
<td>0.13861</td>
</tr>
<tr>
<td>VaEA</td>
<td>0.87172</td>
<td>0.15339</td>
</tr>
<tr>
<td>Std.</td>
<td>0.00023</td>
<td>0.10437</td>
</tr>
</tbody>
</table>

Fig. 10. Optimal Pareto solutions for the rail maintenance decisions. The x-axis depicts the track distance (km) versus the maintenance decision, where 1 represents replacement and 0 represents no replacement. The following algorithms were used: (a) ARMOEA, (b) NSGAII, (c) GrEA, (d) SPEA2, (e) RSEA, and (f) VaEA.
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

To visualize the maintenance solution obtained by using the RSEA, a map in Fig. 11 is provided, indicating the 25 locations that are candidates for replacement. Moreover, a zoom-in plot is included to show a piece of the track. The track piece shows seven maintenance solutions labeled with the replacement ranking. For example, the location labeled ‘9’ has priority for replacement compared to the other locations with higher numbers.

![Image](image_url)

Fig. 11. Map of the maintenance solutions depicted with “+”. A zoom-in plot is attached to show the maintenance priorities with numbers.

### TABLE III

<table>
<thead>
<tr>
<th>The 6 peer algorithms</th>
<th>(30 \text{ s/data block})</th>
<th>(106.17 \text{ s/km})</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMOEA</td>
<td>mean 24.028 s Std. 0.083 s</td>
<td></td>
</tr>
<tr>
<td>SPEA2</td>
<td>mean 30.071 s Std. 0.157 s</td>
<td></td>
</tr>
<tr>
<td>NSGAII</td>
<td>mean 25.722 s Std. 0.0988 s</td>
<td></td>
</tr>
<tr>
<td>GrEA</td>
<td>mean 24.930 s Std. 0.114 s</td>
<td></td>
</tr>
<tr>
<td>RSEA</td>
<td>mean 26.187 s Std. 0.090 s</td>
<td></td>
</tr>
<tr>
<td>VaEA</td>
<td>mean 25.446 s Std. 0.091 s</td>
<td></td>
</tr>
</tbody>
</table>

Table III is presented to show the computational time of (1) data acquisition, (2) the detection algorithm, and (3) the optimization algorithms. For the data acquisition, the data block obtained from the accelerometers is transferred into raw files within 30 seconds. Finally, when a new raw file is created, the old file is transformed into a text file in 10 seconds, and both the raw file and the text file can be stored on a hard drive. Regarding the detection algorithm, the HHT-based detection algorithm requires an average of 2861.630215 seconds to analyze the data measured from the full track, 26.9525 km. Thus, the time needed is approximately 106.17 s/km. Moreover, six algorithms used for the optimization are listed in the table as well. As the processing times for these algorithms are not considerably high, the computational time difference between algorithms is not of major concern to the infrastructure manager.

Furthermore, simplicity and quickness are major indicators of the optimization algorithms to be used in practice. Relying on Table III, the ARMOEA algorithm is fastest compared to the others. However, when considering the simplicity plus the performance of the algorithms, RSEA and SPEA2 are interesting to use in practice as well.

### IV. CONCLUSIONS

In this paper, a Pareto-based maintenance decision system is proposed for rail welds. The proposed approach is applied for condition-based maintenance of rail welds based on information from a track in a regional railway network. The Hilbert spectrum approach is used to detect welds based on the ABA. The case study was conducted for track between Bartolomeu and Zârnești in Romania. To obtain the rail performance, a global KPI was proposed to indicate the conditions of the rail by including degradation scenarios and predictions to provide a clear picture of the network conditions during the maintenance period. The rail replacement number is simply considered as the second objective function, which is related to the budget of the infrastructure manager and the capacity to perform replacement operations during the maintenance period. Six algorithms were run using the same setup to optimize the rail maintenance decisions and evaluate the optimization results. Thus, the infrastructure manager not only receives Pareto optimal solutions but can also compare the results between the different algorithms. Among all six algorithms, the RSEA is the most reliable and has the most similarity to the true Pareto front compared to the other algorithms. Moreover, the coverage and diversity of the algorithms are tested using two performance metrics. The numerical results prove that the RSEA algorithm has good performance.

This analysis will drastically reduce the amount of time needed for field inspections. Instead of visiting each weld in the infrastructure, the managers can focus on the welds that provide Pareto optimal solutions based on ABA signals. In the future, a distributed system could be applied to the proposed framework to include larger sections of tracks. Moreover, by considering the track characteristics such as track geometry, a more elaborate framework for KPIs can be acquired. Another topic for future research is to consider the life-cycle costs as an objective function so that the rail replacement criterion can be reformulated according to different operation costs. Other objectives such as social value and environmental impact can also be included when adequate...
Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

tools are able to dynamically assess their effects. While this framework is focused on analyzing rail welds, the methodology can be used to analyze other types of rail defects, such as corrugations, squats and damaged insulated joints. The durability of the ABA measurement system has not been investigated yet, and it will be a topic of future research. In this study, the ABA system was already used for a week during the day in mixed dry and wet conditions, and no technical problems that might affect the results were reported. The effects of high temperature variations, such as snow conditions or high temperatures (at 30°C the rail might have a temperature approximately 70°C), are future research topics.

REFERENCES


Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration

Alfredo Núñez (M’02–SM’14) received the Ph.D. degree in electrical engineering from the Universidad de Chile, Santiago, Chile, in 2010. He was a post-doctoral researcher with the Delft Center for Systems and Control, Delft, The Netherlands. He is currently permanent staff of the Section of Railway Engineering, Delft University of Technology. His expertise is in the fields of maintenance of railway infrastructures, intelligent conditioning monitoring in railway systems, big data, risk analysis, and optimization. He has authored a book titled Hybrid Predictive Control for Dynamic Transport Problems in the Series of Advances in Industrial Control (Springer-Verlag, 2013). He has (co)authored about a hundred international journal papers and international conference papers. Currently, he is work package leader in the development of new sensor technologies (static, moving and crowd based sensors) for railway networks in Romania, Turkey and Slovenia, in the European project H2020 NeTIRail-INFRA project.

Ali Jamshidi received the MSc degree in the area of disaster risk analysis from Tehran University, Iran, in 2011. He is currently working on his Ph.D. project with the Section of Railway Engineering, Delft University of Technology, Delft, The Netherlands. His project is related to multi-party risk management and key performance indicator design at different system levels for railway infrastructures. He has been actively collaborating the Section of Railway Engineering in rail data measurements and rail data processing. His research interests include: risk analysis, fuzzy systems, predictive maintenance, failure analysis and infrastructure condition monitoring.

Hongrui Wang (S’15) received the B.S. degree in Electrical Engineering from Mao Yisheng Class, Southwest Jiaotong University, Chengdu, China, in 2012. He is currently pursuing the Ph.D. degree with the Section of Railway Engineering, Delft University of Technology, Delft, The Netherlands. His research interests include signal processing, machine learning and their applications in the assessment, detection, diagnosis and maintenance of railway infrastructure system.