Automatic Detection of Squats in Railway Infrastructure

Maria Molodova, Zili Li, Alfredo Núñez, *Member, IEEE*, and Rolf Dollevoet

***Abstract*—This paper presents an automatic method for detecting railway surface defects called "squats" using axle box acceleration (ABA) measurements on trains. The method is based on a series of research results from our group in the field of railway engineering that includes numerical simulations, the design of the ABA prototype, real-life implementation, and extensive field tests. We enhance the ABA signal by identifying the characteristic squat frequencies, using improved instrumentation for making measurements, and using advanced signal processing. The automatic detection algorithm for squats is based on wavelet spectrum analysis and determines the squat locations. The method was validated on the Groningen–Assen track in The Netherlands and accurately detected moderate and severe squats with a hit rate of 100%, with no false alarms. The methodology is also sensitive to small rail surface defects and enables the detection of squats at their earliest stage. The hit rate for small rail surface defects was 78%.**

***Index Terms*—Axle box acceleration (ABA), rail transportation maintenance, railway monitoring, surface defects on railway rails.**

I. Introduction

THE detection of anomalies in railway tracks in their early stage and their timely maintenance can prevent failures and traffic disruptions, and also minimize the long-term cost of the railway infrastructure. This paper focuses on the detection of squats, which are a class of short surface-initiated track defects [1]–[3]. Squats can initiate at small indentations, corrugations, and welds [4], [5]. When squats are detected at an early stage and the degradation is minor, the tracks can be easily treated by grinding a thin layer from the surface. Such early detection significantly reduces the maintenance cost of tracks because severe squats can lead to the replacement of the track section. In The Netherlands, squat-related costs are more than 5000 €/km per year. Because the railway infrastructure in The Netherlands is one of the most intensively used systems in Europe, the costs related to squats are relevant for the Dutch railway infrastructure manager (ProRail).

The Dutch railway network is approximately 2800 km long and includes 6500 km of tracks, 4700 km of electrified tracks, 8700 switches, 4500 bridges and tunnels, 3000 level crossings, and 380 stations. The system carries more than 1 200 000 passengers on 6000 trains per day. Thus, the prevention of operational disruptions in the system is of paramount importance, not only because of the high cost of replacing broken rails but also because delays and derailments are undesirable to the users of the system.

Systematic and periodical inspections must be performed to keep the tracks in good condition. Using computer-based tools, infrastructure managers can then systematically make maintenance decisions to minimize the total costs and guarantee the long-term quality of the infrastructure [6]. Several methods are available to monitor the health of railway tracks [7]–[11]. In The Netherlands, the current methods used to detect short-wave defects in railway tracks include visual inspections and ultrasonic and eddy current measurements [12], [13]. However, these types of inspections are most efficient at later stages of degradation and are not optimal. Further visual inspections are slow and laborious, and the results are dependent on the human operator [14]. Thus, the development of effective monitoring systems that assess the rail conditions more frequently, together with automatic detection methods to locate such relevant irregularities as squats, are key issues for efficient and robust infrastructure management.

Vertical axle box acceleration (ABA) has been employed by Dutch railways to detect such defects as corrugation and poor-quality welds since the mid-1980s [15]. The main advantage of ABA compared with other methods is its lower cost and ease in maintenance. Recently, enhanced methods using accelerometers have been proposed for monitoring track conditions [16]–[19]. Lee *et al.* [20] described a mixed filtering approach for estimating track irregularities using a set of accelerometers mounted on the axle box and bogie of high-speed trains. Well-tuned Kalman filters, compensators, and bandpass filters allow for good-quality estimations of the lateral and vertical irregularities. Shafiullah *et al.* [21] presented a communications protocol between the sensors (accelerometers) placed in a railway–wagon system to monitor the typical dynamic behavior of railway wagons. The protocol improved energy consumption and is sufficiently generic to be implemented in other wireless data communication processes. ABA measurements have been employed to detect track defects such as corrugation, welds, and poor-quality insulated joints whose effects can be easily observed in the ABA signal [22]–[24]. The combination of a good

Manuscript received May 31, 2013; revised October 18, 2013 and December 20, 2013; accepted February 10, 2014. Date of publication March 20, 2014; date of current version September 26, 2014. This research is supported by the Dutch railway infrastructure manager ProRail. This research is supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organization for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs. This paper was presented in part at the 16th International IEEE Conference on Intelligent Transportation Systems, The Hague, The Netherlands, October 6–9, 2013. The Associate Editor for this paper was J. V. Krogmeier.

The authors are with the Section of Road and Railway Engineering, Delft University of Technology, 2628 Delft, The Netherlands (e-mail: M.Molodova@tudelft.nl; Z.Li@tudelft.nl; A.A.NunezVicencio@tudelft.nl; R.P.B.J.Dollevoet@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2014.2307955

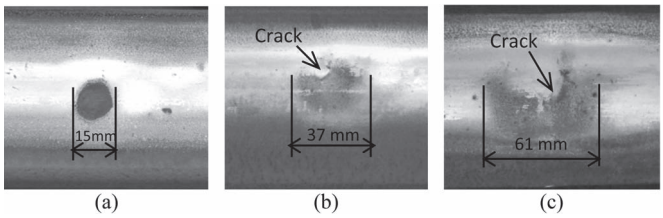

Fig. 1. (a) Light squat. (b) Moderate squat. (c) Severe squat.

measurement design, proper location of the accelerometers, and the use of intelligent methods to enhance the quality of the signals has been crucial in demonstrating the effectiveness of the prototypes, which is an initial step before the massive implementation of real-life railway-condition-monitoring solutions. All of the findings available in the literature can be integrated within a global monitoring system of railway conditions that can detect the different classes of defects.

However, to the best of our knowledge, the potential of this measuring method has not been fully exploited for the detection of all classes of squats. There are several reasons for this lack of widespread implementation, and they all present major challenges: 1) squats are typically found randomly and are isolated, i.e., only one squat is found in each location and 2) their response in the ABA signal (particularly light squats) cannot be easily observed without proper instrumentation and signal processing.

This paper presents the first results of a new methodology for the detection of squats that is based on a prototype designed in our group [25], and that was implemented and validated via extensive field tests on the Dutch tracks. The plan is to mount this prototype on in-service trains to provide an effective system for monitoring the entire railway infrastructure on a daily basis with continuous updates at a relatively low cost. This ability is particularly relevant for cases of rapid degradation, in which the typical six-month schedule of inspections is too long to apply proper corrective measures. The implementation of this system at a large scale will significantly enhance the safety of railway operations and reduce the life-cycle costs of the entire infrastructure.

The following section presents background information about squats and the experimental setup of the real-life specialized measurement system for trains.

II. System Description

A. Squats

Squats are rolling contact fatigue defects in rails that are generally considered to be surface initiated. The term "squat" originated from the defect's typical shape, which resembles the print that would be left behind by a very heavy gnome sitting on the rail. Fig. 1 presents reference photographs of the three classes of squats, i.e., light, moderate, and severe.

The bottom of the squats is typically rusty because the squats are too deep for contact with wheels or there is a network of cracks beneath the squat. Hence, the depth of a squat is typically greater than 0.05 mm, which is the typical compression of a loaded rail. Measurements of vertical–longitudinal profiles indicate that the maximum vertical deviations of the rail surface at squats are typically less than 0.2 mm.

The growth of squats is a result of the dynamic contact force between wheels and rails. The wavelength of the contact force and the consequential wavelength of the squats embodied in the length of the typical two lungs and the corrugation-like wave pattern of squats are a natural characteristic of the coupled wheel–track system [26]. For the Dutch tracks, the wavelength of the contact force and squats is typically between 20 and 40 mm. The wavelength increases with the severity of a squat, due to changes brought about in the local track system by the dynamic interaction at the squat; it can be up to 60 mm [1]. The wavelength and its change will be reflected in the vibrations of the wheels because of the dynamic wheel–rail interaction; the vibrations can be picked up by accelerometers at the axle box.

Observations of rail surface defects have indicated that not all defects can grow into squats. If the size of a rail surface defect exceeds a critical size, it may grow into a squat. For tracks in The Netherlands, the critical size for visual inspection is 6–8 mm for both the rolling and transverse directions when the traction and braking efforts are maximal. For this study, rail surface defects that exceed this critical size are considered light squats. Defects below this threshold are considered trivial defects and will likely disappear due to wear.

B. Monitoring of the Railway Infrastructure

ProRail provides maintenance and extensions to the railway network infrastructure in The Netherlands. To formalize the monitoring tasks of the infrastructure, ProRail uses condition scores to determine the status of the tracks based on several indexes. The presence of squats decreases the score and maintenance measures are suggested when the score of the track reaches certain thresholds. Let us represent the condition of the track at time t and location x with the function $H(t,x) \in \mathbb{D}$. This function can take discrete values $\mathbb{D} = \{S, A, B, C, N\}$ if the railway infrastructure at time t and coordinate x has a trivial defect S, light squat A, moderate squat B, or severe squat C. The value of this function is N when no squat is present at x.

The growth of squats depends on the track loading conditions but is generally a relatively slow process (it may take months for a squat A to evolve into a squat B). Thus, within a time interval (typically weeks), $H(t,x)$ is not expected to considerably change. Let us define function $H(t,x)$ during period $[t-T, t+T]$ as $H_T(x)$. To estimate function $H_T(x)$ over the tracks within the network of interest $x \in \mathrm{X}$, measurements taken on train i will be used $y_i(t), t \in [t-T, t+T]$. This set of measurements includes the position of the train $x_i(t)$, which is provided by a Global Positioning System (GPS); the velocity of the train $v_i(t)$; and a set of ABA measurements $a_i(t)$ in the vertical and longitudinal directions collected in different bogies. The railway track system is shown in Fig. 2.

The Road and Railway Engineering Section of Delft University of Technology has been analyzing squats by focusing on their causes and modeling their interactions with wheel–track systems [26], [27]. Field experiments and studies have demonstrated that the detection of squats requires high sensitivity, resolution, and accuracy; hence, a more accurate and reliable

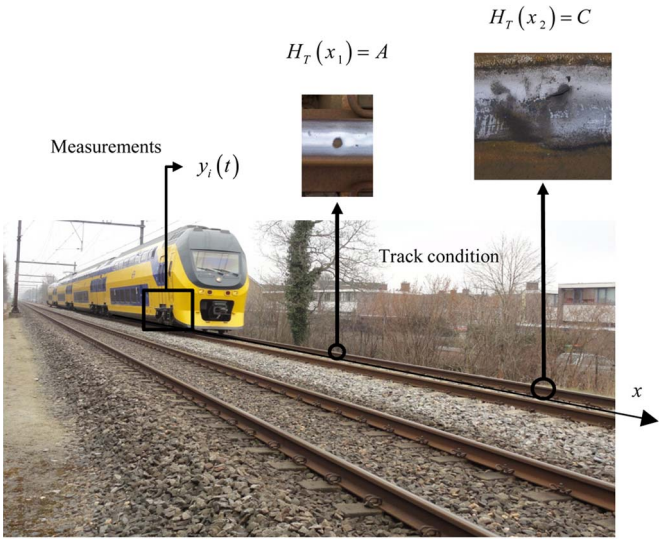

Fig. 2. Track conditions are assessed via signal processing methods using on-board measurements.

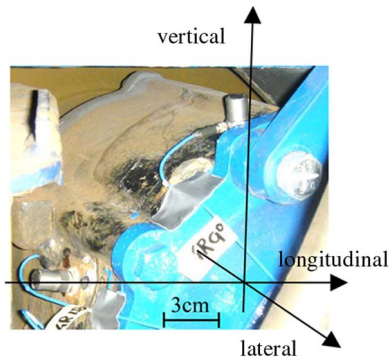

Fig. 3. Accelerometers mounted on the axle box in the vertical and longitudinal directions.

system with higher sensitivity and resolution than previous systems had to be developed. The implementation of the system is described in the following section.

C. Implementation of the Measuring System

A prototype of an ABA measuring system, which is described in a patent [25], has been installed on a specialized measurement train. The system includes improved instrumentation that is reliable for frequencies higher than 1 kHz to measure both vertical and longitudinal ABA signals. The longitudinal ABA signal increases the hit rate of detecting light squats due to the high signal-to-noise ratio. We used general-purpose uniaxial piezoelectric accelerometers that were calibrated by the manufacturer to comply with the accuracy standards. The sampling frequency of the ABA measurements was 25.6 kHz. Fig. 3 presents the accelerometers mounted on one of the axle boxes.

Three accelerometers, including one in the vertical direction and two in the longitudinal directions, were mounted on each of the four axles of a bogie. This design reduces the chances that the squat will not be recorded because the wheel traveled along a different trajectory on the rail due to its inherent

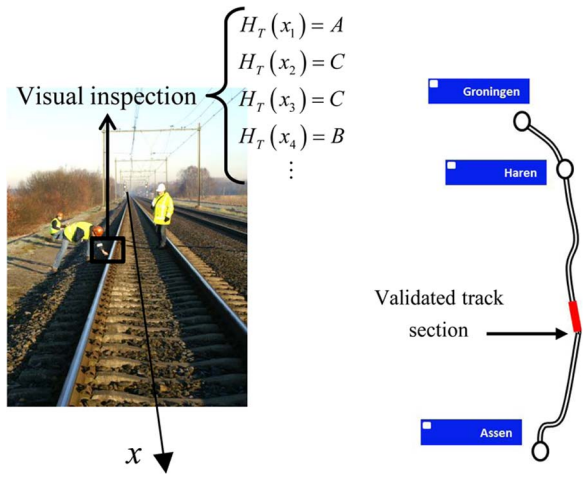

Fig. 4. Monitored track and visual inspection on the track between Groningen and Assen, The Netherlands. Extensive field measurements provided the list of defects and locations used to validate the methodology.

oscillatory movements (hunting) or because the squat is smaller than the width of the rolling band (so that not every wheel will necessarily run over it). The data recorded by the prototype system also included the GPS coordinates and train speed.

D. Validation Data

Every six months, a measurement train collects information about the infrastructure over the entire Dutch network in a database called IRISsys. The database contains photographs of the rail tops, the locations of insulated joints, switches, bridges, viaducts, and level crossings, and other information about the infrastructure. In this database, the locations of short track irregularities can be identified by the GeoCode of the track and the conventional kilometer location. The kilometer location is also used in this study to position ABA measurements and identify the locations of squats.

This study required the most complete and up-to-date information about the short track irregularities. Therefore, detailed visual inspections were performed (see Fig. 4). These inspections provided the following validation data: the GPS coordinates of the defects; photographs; vertical–longitudinal profiles of the rail, which were measured with the RAILPROF device; and MINIPROF measurements of the cross-sectional profile of the rail at the short track irregularities, welds, and insulated joints.

The validated track sections were located on the track between Groningen and Assen in The Netherlands. The following section describes how the measured ABA signals are analyzed according to their wavelet power spectrum (WPS) response to detect the squats.

III. Wavelet Analysis and Squats

A. Frequency Analysis Using the Wavelet Transform

The ABA measurements in the time domain are not sufficient to detect small defects. The proper frequency content of the ABA must be used to improve the detection of moderate and light squats. Several techniques are available to investigate the

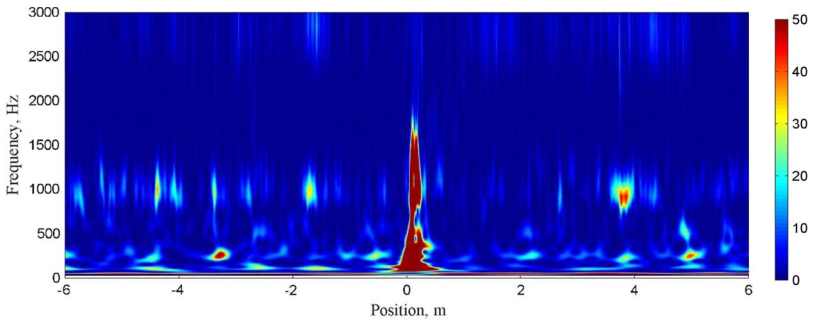

Fig. 5. Scalogram of the ABA signal around a moderate squat located in a track section between Groningen and Assen, The Netherlands.

frequency content of the signal, but not all of these techniques provide good information. The main drawback of the short-time (or windowed) Fourier transform (STFT) is the selection of the window size. For the ABA analysis, a tradeoff between the time and frequency resolution was observed when the window size was changed. For example, a shorter window size results in a lower frequency resolution and higher time resolution. A different window size was required for each type of defect to detect the different classes of squats and assess a broader range of track elements.

Wavelet analysis is used in this paper because, among other reasons, it has the advantage that the time–frequency representation is not dependent on the window size. The continuous wavelet transform (CWT) is a time–frequency analysis tool in which the observed function is multiplied by a group of shifted and scaled wavelet functions. The CWT can be defined as follows [28]:

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi^* \left(\frac{(n'-n)\delta_t}{s} \right) \tag{1}$$

where x_n is a time series with a time step of δ_t; n is the time index; $n' = 0, \ldots, N-1$ is the time shift operator; ψ is a mother wavelet, which is a locally limited function; $\psi^*((n'-n)\delta_t/s)$ is a family of wavelets deduced from the mother wavelet by different translations and scaling; $*$ indicates a complex conjugate; s is a wavelet scale, $s > 0$; and $W_n(s)$ are wavelet coefficients.

The wavelet scale is related to the Fourier period (or inverse frequency). The relationship for a particular wavelet function can be derived by finding the wavelet transform of a pure cosine wave with a known Fourier period and then computing the scale at which the WPS reaches its maximum. According to Parseval's theorem of energy preservation, the energy of the wavelet transform is equal to the energy of the original signal in the time domain. The physical meaning of the CWT can be described as the correlation between the original signal and the scaled wavelet at a delay n'. The wavelet transform can be also considered as a linear filtering operation that involves a set of parallel filters. For a more detailed discussion on wavelets, see [29]. The main advantage of the CWT compared with the STFT is its high time and frequency resolutions [18]. Therefore, wavelet analysis is appropriate for the investigation of nonstationary phenomena with local changes in the frequency components, such as structural damage detection and crack identification [30]–[32]. In this paper, the Morlet function is used as a mother wavelet. The Morlet function is defined as

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2} \tag{2}$$

where ω_0 is a nondimensional frequency. The power spectrum of a wavelet transform is defined as the square of the wavelet coefficients, i.e.,

$$\left| W_n^2(s) \right|. \tag{3}$$

The plot of the WPS is termed as a scalogram. A vertical slice of a wavelet plot is a measure of the local spectrum. An example of such a plot is shown in Fig. 5, which presents the scalogram of the ABA signals around a moderate squat. The red color around the zero position represents a high signal energy level that is caused by an impact at the squat. In this study, the scalograms are used to define the time–frequency relationship of squats with the ABA signal.

In this paper, we propose the use of the scale-averaged wavelet power (SAWP) for the automatic detection of squats. This function captures the variation of the spectrum in a signal, and thus, the system triggers the detection when the power spectrum of the frequencies related to the squats is higher than a given threshold. The SAWP is defined as the weighted sum of the WPS over the scales s_{j1} to s_{j2} [33], i.e.,

$$\overline{W}_n^2 = \frac{\delta_j \delta_t}{C_\delta} \sum_{j=j_1}^{j_2} \frac{|W_n(s_j)|^2}{s_j} \tag{4}$$

where n is the time index, δ_j is a scale step, δ_t is the time step, and C_δ is an empirically derived constant for each wavelet function. The SAWP can be used to examine fluctuations in power over a range of scales.

To detect squats, the SAWP is calculated in the frequency bands related to the squats, or the "signature tunes" of the squats, which were derived from finite-element simulations in previous research [27]. The main "signature tunes" of squats, which are defined as a frequency with a local maximum power spectrum, are approximately 300 Hz and 1060–1160 Hz for the Dutch tracks and a short time–frequency response up to 2000 Hz. In practice, the maximum power spectrum is not exactly at 300 or 1060 Hz because of the track and contact conditions and the frequency resolution of the power spectrum.

Because the SAWP is a time series, squats can be identified by finding the local maxima of the SAWP that exceed a certain

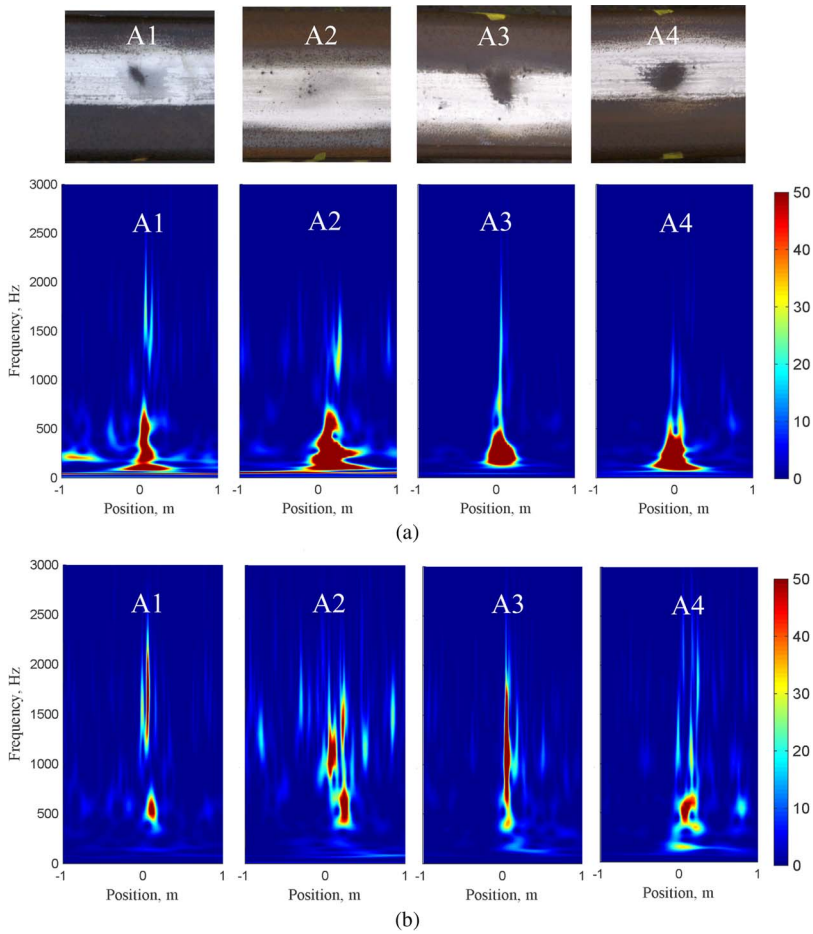

Fig. 6. WPS of vertical and longitudinal ABA at light squats. (a) WPS of the vertical ABA. (b) WPS of the longitudinal ABA.

threshold. In general, a higher SAWP indicates a more severe defect. A constant threshold of 0.5 m^2/s^4 was used for the analyzed tracks. The threshold determines the performance of the detection algorithm. Thresholds are determined to maximize the hit rate and reduce the number of false alarms. The thresholds can be adapted to satisfy the requirements of the infrastructure managers depending on the track properties.

In the following section, the relationship between different classes of squats and the frequency characteristics of the ABA are presented with wavelet scalograms. In the following sections, only real-life ABA measurements over real defects on the track between Groningen and Assen in The Netherlands are used as typical examples.

B. Light Squats and Trivial Defects

A frequency response of up to 2.5 kHz was observed in both the vertical [see Fig. 6(a)] and longitudinal [see Fig. 6(b)] ABAs at light squats. Photographs of the squats are provided at the top of Fig. 6. There are two major frequency responses: below approximately 800 Hz and from 800 to 2500 Hz. The frequency response above 800 Hz is stronger in the longitudinal ABA.

Trivial defects are defects that are smaller than the critical size of the squats (6–8 mm), and such defects will disappear due to wear. If these small defects can be detected, larger defects can be also detected. The scalogram of the longitudinal ABA measurements of trivial defects displays the presence of frequency components between 800 and 2000 Hz (see Fig. 7). This observation confirms the similarity of the frequency characteristics of trivial defects and light squats. The trivial defects that correspond to their WPS are shown in Fig. 7. These defects are less than 8 mm in size, and all of them subsequently disappeared due to wear.

C. Signature Tunes of Moderate and Severe Squats

Fig. 8(a) presents examples of the WPS at moderate and severe squats. The WPS at these squats have two areas with strong responses: one is below 600 Hz with the maximum around 250–350 Hz, and the other is between 600 and 2000 Hz with the maximum at approximately 1000–1300 Hz. The squats whose responses are presented in Fig. 8 have lengths of 30, 40, and 50 mm. With increasing squat length, the WPS response below 600 Hz becomes stronger, whereas the response between 600 and 2000 Hz becomes relatively weak, and the maximum frequency decreases.

The vibrations caused by squats can be also transmitted through the axle and measured on the opposite end of the axle

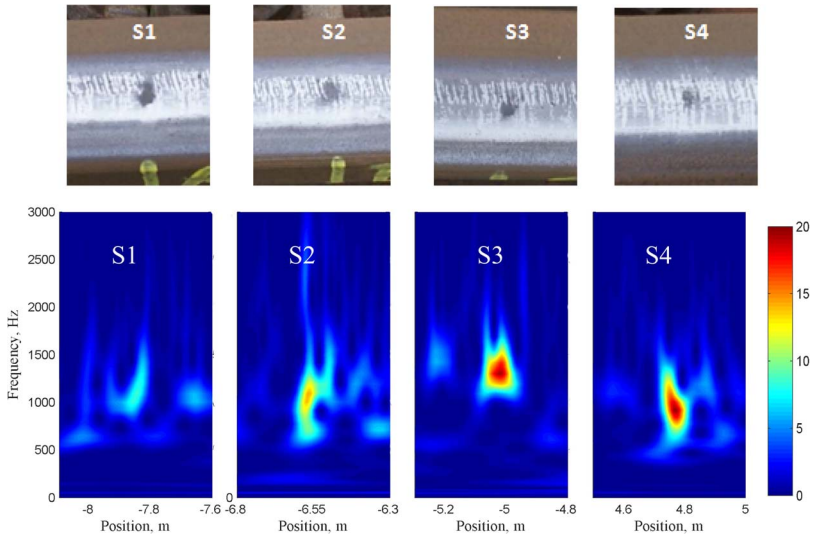

Fig. 7. WPS at trivial rail surface defects found in The Netherlands.

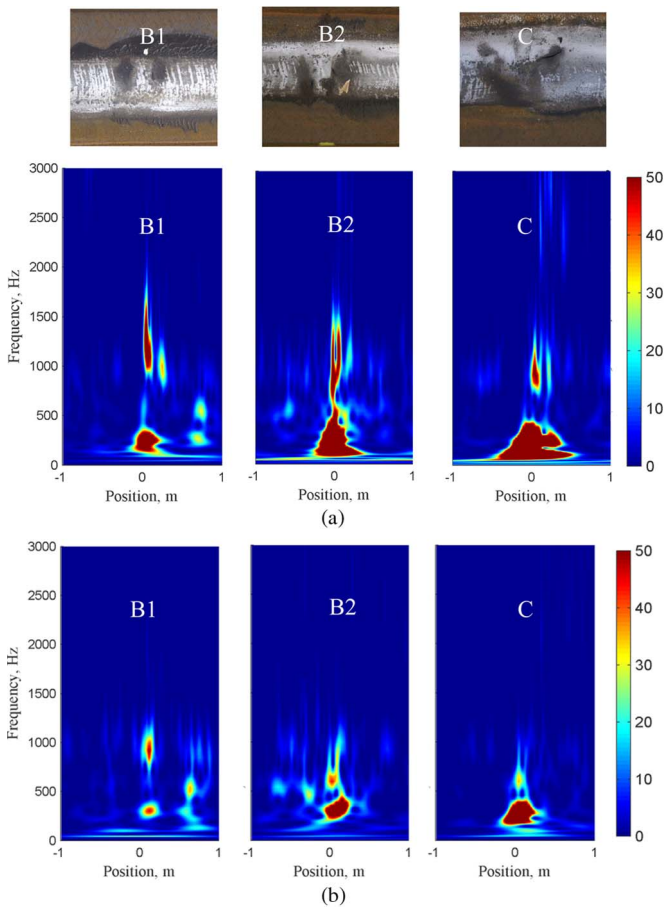

Fig. 8. WPS at typical moderate and severe squats found in The Netherlands. (a) WPS at squats. (b) WPS on the opposite axle.

[see Fig. 8(b)]. The response measured on the opposite end is weaker, and the frequencies are lower than the response measured at the original end. This fact should be considered when implementing the automatic detection process; the responses on the two axle boxes must be compared to determine the rail on which the squat is located.

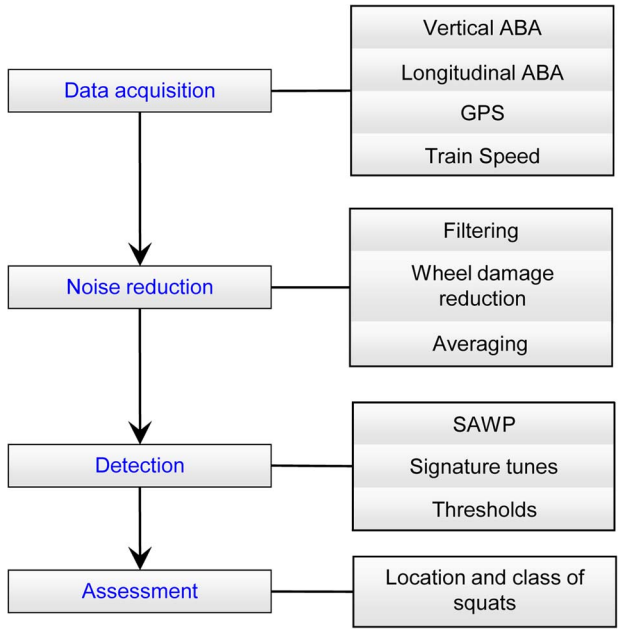

Fig. 9. Detection procedure.

The following section describes how the instrumentation presented in Section II and the wavelet analysis in Section III are used to design the automatic detection method for squats.

IV. Automatic Detection

The process for detecting squats includes data acquisition, preprocessing of the measured data to reduce the noise, detection of squats, and assessment. Fig. 9 presents the flowchart of the detection algorithm. The main steps of this algorithm are described below.

A. Data Acquisition

The data from the measurement train are recorded as described in Section II-C. The information includes multiple

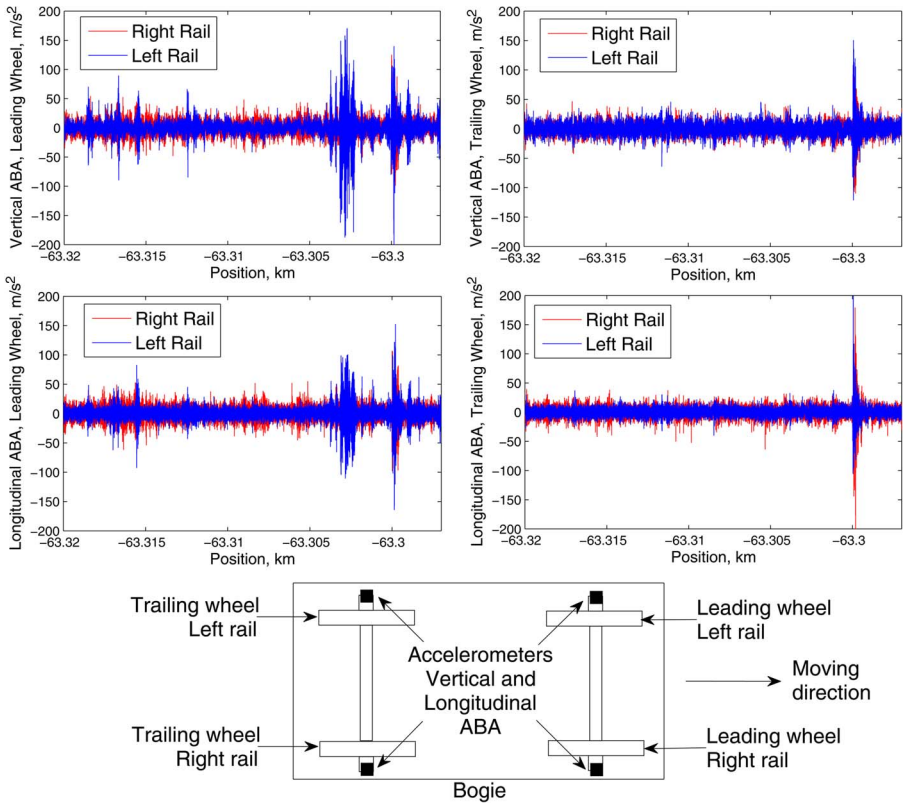

Fig. 10. Data obtained from the measurement system.

vertical and longitudinal ABA signals, GPS measurements, and the train speed. Fig. 10 provides an example of the raw data set for a short measurement period.

B. Noise Reduction

The measured data are then processed using the following steps.

1) Filtering of the ABA signals: The data were filtered using the signal processing toolbox of MATLAB. A lowpass Butterworth filter with a cutoff frequency of 2000 Hz was applied so that the responses at squats were fully captured.
2) Reduction of the influence of wheel damage on the ABA signals: If the wheel is not in good condition, the signals can still be used after additional signal processing is applied. Wheel damages are easy to detect than squats because they cause a periodic impact between the wheel and rail at a wavelength of the circumference of the wheel. The vertical ABA signal from the sensor located closest to the damaged wheel will exhibit repetitive peaks that appear approximately every 3 m. The excitation can be transferred to the other wheel by the axle but not to sensors located on the other axle. The problem of the repetitive peaks in the vertical ABA was solved by removing the repetitive pattern from the signal.
3) Noise reduction by coherent averaging of the repetitive ABA signals: The coherent averaging method [34] is based on the principle that a time signal that is measured immediately after applying a stimulus contains the invariant response to the stimulus and a noise component. When averaging several similar time signals, all of the invariant responses are systematically added, whereas the random noise components are summed and tend toward zero. Noise reduction is more effective when a large number of averaged samples are used. To apply an averaging technique, it is necessary to know the exact moment at which each stimulus occurs. In the case of ABA measurements, repeated signals were recorded on the same track section and overlapped by cross correlation of the signals in such a way that the responses at certain short track defects began at the same location.

The effect of the signal processing procedure can be clearly observed when analyzing the SAWP signal (see Fig. 11).

In Fig. 11, the peaks of the SAWP signals after processing are more intense and evident, which makes the detection procedure easier. The defects are denoted by red asterisks. The defect at 3.8 m is a moderate squat with $H_T(3.8) = B$. The other defects at $x \in t\{1.2, 2.8, 4.0, 6.9, 9.7, 13.0, 15.3, 18.2\}$ are all trivial defects; hence, $H_T(x) = S$.

C. Detection

The squats are predicted by calculating the SAWP of the ABA signals defined in (4) in the frequency bands related to the squats. As shown in Section III, the bands of interest vary for different stages of squats. Therefore, to include the frequencies with the maximum power spectrum for each class of squats, we define the different frequency bands to calculate the SAWP:

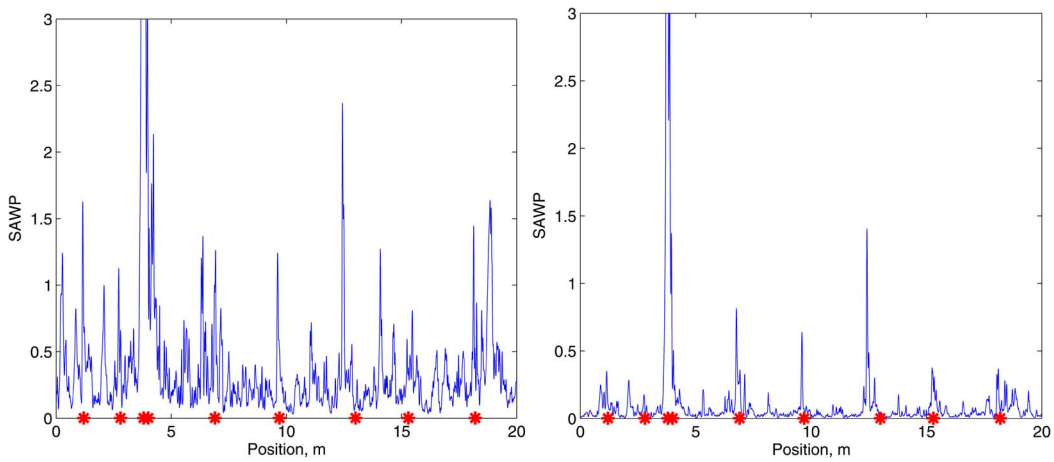

Fig. 11. SAWP before and after signal processing. Real defects are shown with red asterisks on the horizontal axis. The detection of small defects is much easier with the new instrumentation and method.

from 200 to 400 Hz and from 1000 to 2000 Hz for light squats and from 200 to 400 Hz for moderate and severe squats.

The locations of the squats are predicted by the values of the SAWP that exceed a certain threshold. A constant threshold of 0.5 m^2/s^4 was used for the analyzed tracks, which provided a good tradeoff between maximizing the hit rates and minimizing the false-alarm rate. The threshold is also exceeded by other track components, such as welds and insulated joints. If the locations of those components are known, they can be excluded from the squat analysis.

The ABA signal is influenced by the train speed. The measurement train used in this study had a nearly constant speed in the range between 100 and 110 km/h. In practice, under low speed levels, we have experienced a reduction in the power spectrum intensity in the high-frequency range, making the detection of small squats more difficult. However, when the train speed varies within the range of 60–100 km/h, using 3-D finite-element model simulation results, it has been proposed the use of quantitative relationships with the signature tunes and maximum ABA using a regression model [35].

D. Assessment

To validate the proposed method, the track was visually inspected (as explained in Section II-D) to properly quantify the false alarms and hit rate. The numbers of hits (H), misses (M), false alarms (F), squats, and predictions were calculated (see Fig. 12 and Table I).

The hit rate (HR) is defined as the ratio of the number of correct predictions to the total number of observed defects, which is the sum of the hits and misses. The false-alarm rate (FA) is defined as the ratio of the number of falsely predicted defects to the total number of predicted defects, which is the sum of the hits and the falsely predicted. They are, namely

$$HR = \frac{H}{H+M} \tag{5}$$

$$FA = \frac{F}{F+H}. \tag{6}$$

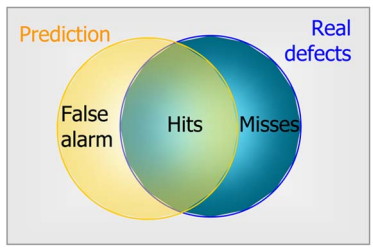

Fig. 12. Automatic detection method is designed to maximize the hits and minimize the misses and false alarms.

TABLE I
ASSESSMENT OF PREDICTION

	Observed	Not observed
Predicted	Hit (H)	Falsely predicted (F)
Not predicted	Miss (M)	Correct rejection

For accurate predictions, the hit rate should be maximized, and the false-alarm rate should be minimized.

V. RESULTS

A. Detection of Light Squats and Trivial Defects

This section focuses on predicting squats using the detection algorithm proposed above. The SAWP was calculated using (4). To include only the frequencies that are related to squats, the WPS was multiplied by a weight function at each moment in time. The weight function for light squats was equal to 1 at frequencies between 200 and 400 Hz and between 1000 and 2000 Hz and 0 at other frequencies. The SAWP was then calculated over all scales.

Six 15- to 25-m-long track sections were investigated. The sections contained 51 defects. The validation track sections were located between Groningen and Assen in The Netherlands. After the measurement train collected the data, a visual inspection was performed as explained in Section II. The SAWP of the track sections and the detected locations of the defects are shown in Fig. 13.

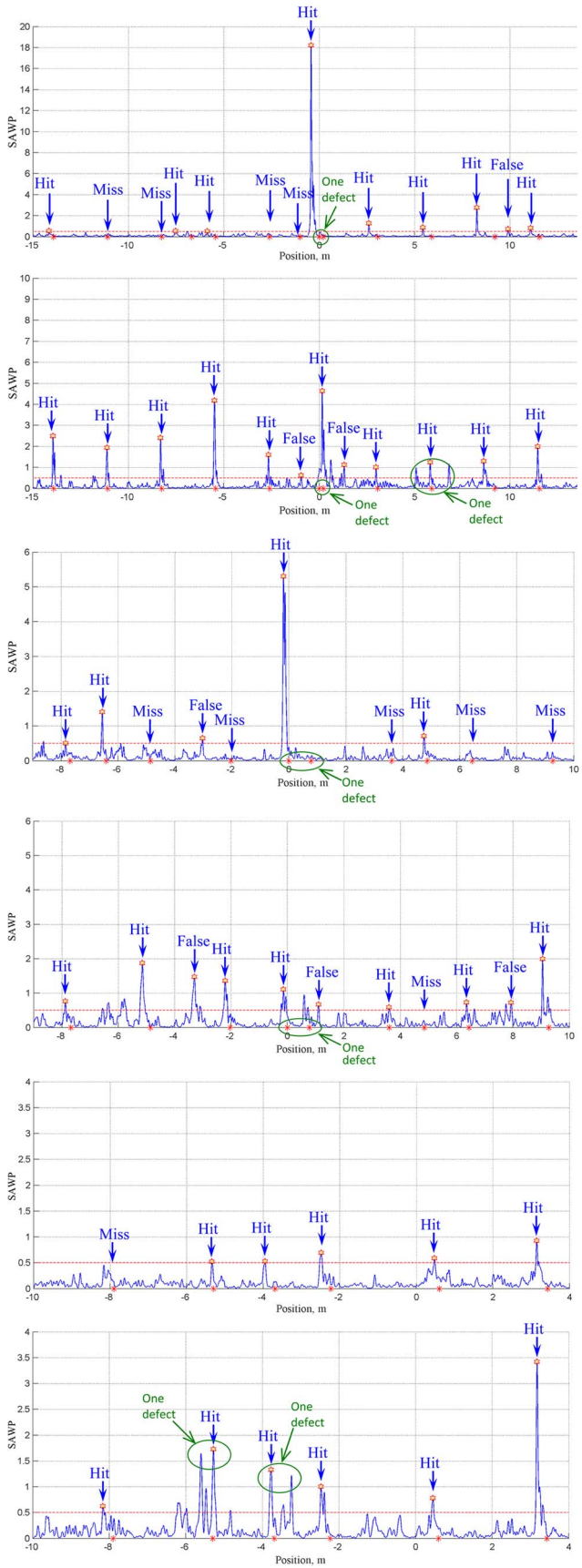

Fig. 13. SAWPs at six different sections. Real defects are denoted by red asterisks on the horizontal axis and the predicted defects are denoted by red stars at the SAWP peaks. If the distance between two defects is less than 1 m, they are considered as one defect, as is indicated with the green ovals.

TABLE II
HIT RATE OF TRIVIAL DEFECTS

Location	No. of defects	No. of predictions	No. of detected defect	No. of false alarms	HR	FA
1	12	9	8	1	67%	11%
2	10	12	10	2	100%	17%
3	9	5	4	1	44%	20%
4	8	10	7	3	88%	30%
5	6	5	5	0	83%	0%
6	6	6	6	0	100%	0%
Total	**51**	**47**	**40**	**7**	**78%**	**15%**

The red asterisks on the horizontal axis denote the real defects in the track, and red stars at the peaks denote their predicted locations. If the distance between two defects was less than 1 m, they were considered as one defect. The threshold for the detection was 0.5 $\mathrm{m}^2/\mathrm{s}^4$; this threshold was chosen to maximize the number of hits and minimize the number of false alarms using the training data. Further developments of the methods to obtain good thresholds (track structure dependent) and the possibility to determine which rail (right or left) has the defect with higher accuracy, among other topics, are the subject of future research.

Table II presents the hit rate and the numbers of correct and false predictions. Based on the 51 defects, the hit rate is 78%, and the false-alarm rate is 15%. Some of the defects located at x were found to be trivial defects by visual inspection: $H_T(x) = S$. Light squats were easier to detect because they are larger than these defects. Hence, the hit rate for light squats is expected to be higher than 78%.

A visual inspection at the same location after one year (period T') indicated that some of the defects were removed by natural wear of wheel–rail contact, which changed the condition function to $H_{T'}(x) = N$. The reasons that a defect evolves or does not evolve into a light squat $H_{T'}(x) = A$ will be addressed in future research. The most efficient threshold can be statistically determined by examining more defects. Such a threshold should distinguish light squats that will grow from trivial defects that will disappear due to wear.

B. Detection of Moderate and Severe Squats

The numerical simulations and ABA measurements indicated that the most pronounced frequencies at moderate and severe squats were between 250 and 350 Hz. To include the frequencies related to moderate and severe squats, the SAWP was calculated within the frequency bounds of 200–400 Hz.

The analyzed section of track between Groningen and Assen includes five moderate and severe squats on the right rail and one severe squat on the left rail; these squats are denoted by red circles in Fig. 14. The threshold for the detection of moderate and severe squats was chosen to be 0.5 $\mathrm{m}^2/\mathrm{s}^4$. Although the threshold for moderate and severe squats is coincidentally the same as that for trivial defects, the SAWP for detecting light squats and the SAWP for detecting moderate and severe squats included different frequencies. Using these thresholds, the hit rate for squats was 100%. As expected, some of the squats were detected in the signals from both the right and left rails.

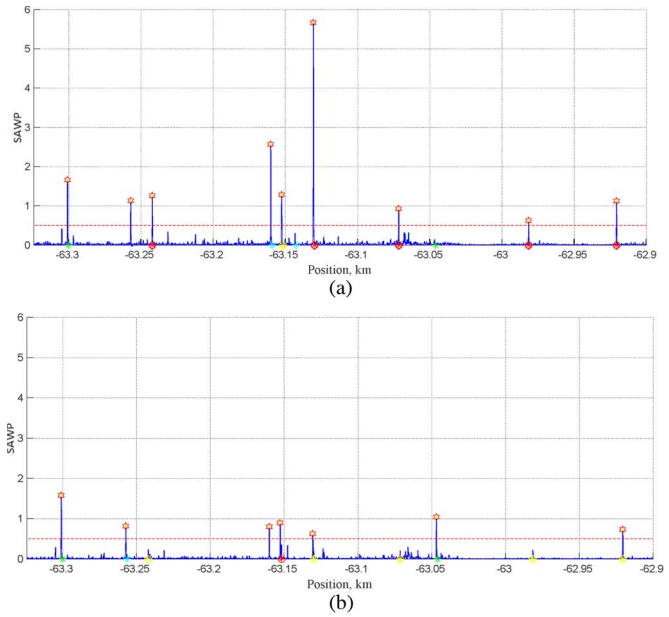

Fig. 14. Detection of moderate and severe squats. Red circles denote the moderate and severe squats, yellow circles denote the moderate and severe squats on the opposite rail, cyan asterisks denote artificial defects, green asterisks denote insulated joints, and red stars denote predictions. (a) Right rail. (b) Left rail.

VI. Conclusion

This paper has presented a new measuring system for the automatic detection of squats. The detection algorithm relies on the signature tunes of the squats, which were identified from numerical simulations in a previous work and were validated by field measurements. The detection algorithm was based on the SAWP. The averaging of the wavelet spectrum was performed at the frequency bands related to the squats. These frequency bands are different for light and for moderate and severe squats: 200–400 Hz and 1000–2000 Hz for light squats and 200–400 Hz for moderate and severe squats.

The thresholds for detecting squats on the Dutch tracks were empirically obtained. The hit rate using automatic detection for a mixture of trivial defects and light squats was 78%, and the false-alarm rate was 15%. These results can be improved by separating and ignoring the trivial defects that will disappear due to wear from light squats that will continue to grow. Such a distinction can be made by tuning the threshold. The hit rate for moderate and severe squats was 100%, with zero false alarms.

Using the presented measuring system with advanced data analysis will enable the detection of squats at their earliest stage, when early corrective or preventive actions can be taken. Such an effective maintenance policy can significantly reduce the life-cycle costs of a track that is affected by squats. Future research will focus on improvements to the algorithms and on implementing the system on revenue trains to monitor the network on a continuous basis. Management of the large amount of data collected is important for continuous monitoring because the data can be stored and postprocessed at a later date. Moreover, the accuracy of automatic detection under various track structures and other stochasticities are topics for further research.

Acknowledgment

The authors would like to thank the Associate Editor and anonymous reviewers for their valuable comments and suggestions.

References

[1] Z. Li, R. Dollevoet, M. Molodova, and X. Zhao, "Squat growth—Some observations and the validation of numerical predictions," *Wear*, vol. 271, no. 1, pp. 148–157, May 2011.
[2] *UIC Code, Rail Defects, 4th ed.*, International Union of Railways, Paris, France, 2002.
[3] J. Smulders, "Management and research tackle rolling contact fatigue," *Railway Gazette Int.158*, no. 7, pp. 439–442, Jul. 2003.
[4] P. Clayton and M. Allery, "Metallurgical aspects of surface damage problems in rails," *Can. Metall. Q.*, vol. 21, no. 1, pp. 31–46, Jan. 1982.
[5] Rail Damages, The Blue Book of RailTrack, U.K., 2001.
[6] A. Zoeteman, "Life cycle cost analysis for managing rail infrastructure: Concept of a decision support system for railway design and maintenance," *Eur. J. Transport Infrastruct. Res.*, vol. 1, no. 4, pp. 391–413, 2001.
[7] E. Magel, A. Tajaddini, M. Trosino, and J. Kalousek, "Traction, forces, wheel climb, and damage in high-speed railway operations," *Wear*, vol. 265, no. 9/10, pp. 1446–1451, Oct. 2008.
[8] P. Gullers, L. Andersson, and R. Lundén, "High-frequency vertical wheel–rail contact forces—Field measurements and influence of track irregularities," *Wear*, vol. 265, no. 9/10, pp. 1472–1478, Oct. 2008.
[9] J. C. Nielsen, "High-frequency vertical wheel–rail contact forces—Validation of a prediction model by field testing," *Wear*, vol. 265, no. 9/10, pp. 1465–1471, Oct. 2008.
[10] C. Delprete and C. Rosso, "An easy instrument and a methodology for the monitoring and the diagnosis of a rail," *Mech. Syst. Signal Process.*, vol. 23, no. 3, pp. 940–956, Apr. 2009.
[11] D. Peng and R. Jones, "Modelling of the lock-in thermography process through finite element method for estimating the rail squat defects," *Eng. Failure Anal.*, vol. 28, pp. 275–288, Mar. 2013.
[12] H. Thomas, T. Heckel, and G. Hanspach, "Advantage of a combined ultrasonic and eddy current examination for railway inspection trains," *Insight - Non-Destruct. Testing Condition Monitoring*, vol. 49, no. 6, pp. 341–344, Jun. 2007.
[13] A. Berry, B. Nejikovsky, X. Gibert, and A. Tajaddini, "High speed video inspection of joint bars using advanced image collection and processing techniques," in *Proc. 8th World Congr. Railway Res.*, 2008, pp. 1–13.
[14] F. Marino, A. Distante, P. L. Mazzeo, and E. Stella, "A real-time visual inspection system for railway maintenance: Automatic hexagonal-headed bolts detection," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 3, pp. 418–428, May 2007.
[15] C. Esveld, *Modern Railway Track,* 2nd ed. Zaltbommel, The Netherlands: MRT-Productions, 2001.
[16] P. F. Weston, C. S. Ling, C. J. Goodman, C. Roberts, P. Li, and R. M. Goodall, "Monitoring lateral track irregularity from in-service railway vehicles," *Proc. Inst. Mech. Eng. F, J. Rail Rapid Transit*, vol. 221, no. 1, pp. 89–100, Jan. 2007.
[17] M. Bocciolone, A. Caprioli, A. Cigada, and A. Collina, "A measurement system for quick rail inspection and effective track maintenance strategy," *Mech. Syst. Signal Process.*, vol. 21, no. 3, pp. 1242–1254, Apr. 2007.
[18] A. Caprioli, A. Cigada, and D. Raveglia, "Rail inspection in track maintenance: A benchmark between the wavelet approach and the more conventional Fourier analysis," *Mech. Syst. Signal Process.*, vol. 21, no. 2, pp. 631–652, Feb. 2007.
[19] A. Massel, "Power spectrum analysis—Modern tool in the study of rail surface corrugations," *NDT E Int.*, vol. 32, no. 8, pp. 429–436, Dec. 1999.
[20] J. S. Lee, S. Choi, S. S. Kim, C. Park, and Y. G. Kim, "A mixed filtering approach for track condition monitoring using accelerometers on the axle box and bogie," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, pp. 749–758, Mar. 2013.
[21] G. M. Shafiullah, S. A. Azad, and A. B. M. S. Ali, "Energy-efficient wireless MAC protocols for railway monitoring applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 649–659, Jun. 2013.
[22] Y. Sunaga, I. Sano, and T. Ide, "A practical use of axlebox acceleration to control the short wave track irregularities," *Railway Tech. Res. Inst., Quart. Rep.*, vol. 38, no. 4, pp. 1–6, 1997, World Congress on Railway Research.
[23] P. Weston, C. Ling, C. Roberts, C. Goodman, P. Li, and R. Goodall, "Monitoring vertical track irregularity from in-service railway vehicles," *Proc. Inst. Mech. Eng. F, J. Rail Rapid Transit*, vol. 221, no. 1, pp. 75–88, Jan. 2007.

[24] A. M. Remennikov and S. Kaewunruen, "A review of loading conditions for railway track structures due to train and track vertical interaction," *Struct. Control Health Monitoring*, vol. 15, no. 2, pp. 207–234, Mar. 2008.
[25] Z. Li and M. Molodova, "Method and instrumentation for detection of rail defects, in particular rail top defects," Eur. Patent WO 2 011 019 273 (A1), Feb. 17, 2011.
[26] Z. Li, X. Zhao, C. Esveld, R. Dollevoet, and M. Molodova, "An investigation into the causes of squats—Correlation analysis and numerical modeling," *Wear*, vol. 265, no. 9/10, pp. 1349–1355, Oct. 2008.
[27] M. Molodova, Z. Li, A. Nunez, R. Dollevoet, "Parameter study of the axle box acceleration at squats," *Proc. Inst. Mech. Eng., F, J. Rail Rapid Transit*., DOI: 10.1177/0954409714523583.
[28] I. Daubechies, "The wavelet transform, time–frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
[29] G. Bachman, L. Narici, and E. Beckenstein, *Fourier and Wavelet Analysis*. New York, NY, USA: Springer-Verlag, 2000.
[30] A. V. Ovanesova and L. E. Suárez, "Applications of wavelet transforms to damage detection in frame structures," *Eng. Struct.*, vol. 26, no. 1, pp. 39–49, Jan. 2004.
[31] K. M. Liew and Q. Wang, "Application of wavelet theory for crack identification in structures," *J. Eng. Mech.*, vol. 124, no. 2, pp. 152–157, Feb. 1998.
[32] Q. Wang and X. Deng, "Damage detection with spatial wavelets," *Int. J. Solids Struct.*, vol. 36, no. 23, pp. 3443–3468, Aug. 1999.
[33] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bull. Amer. Meteorol. Soc.*, vol. 79, no. 1, pp. 61–78, Jan. 1998.
[34] O. Rompelman and H. H. Ros, "Coherent averaging technique: A tutorial review Part 1: Noise reduction and the equivalent filter," *J. Biomed. Eng.*, vol. 8, no. 1, pp. 24–29, Jan. 1986.
[35] M. Molodova, "Detection of early squats by axle box acceleration," Ph.D. dissertation, Road Railway Eng. Section, Delft Univ. Technol., Delft, The Netherlands, Jan. 2013, 117 pp.

Maria Molodova received the M.Sc. degree in mechanical enginering from N.I. Lobachevsky State University of Nizhny Novgorod, Nizhni Novgorod, Russia, in 2005, with a thesis work on numerical modeling of the electrohydroimpulsive stamping process, and the Ph.D. degree from Delft University of Technology, Delft, The Netherlands, in 2013, with a thesis on the development of detection algorithms for tracking short-wave defects and their causes in the wheel–track system by means of axle box acceleration measurements.

She is currently with the Section of Road and Railway Engineering, Delft University of Technology. Her research interests include the implementation of coupled implicit–explicit finite-element methods for time-domain simulations of wheel–track systems and signal processing techniques for frequency-domain investigation of simulated and real-life measured data.

Zili Li received the B.Sc. and M.Sc. degrees in mechanical engineering from Southwest Jiaotong University, Chengdu, China, in 1988 and 1991, respectively. He received the Ph.D. degree from Delft University of Technology, Delft, The Netherlands, in 2002.

Between 1999 and 2005 he was with the Institute of Road Transportation of TNO, Delft, where he was engaged in research and software development of multibody dynamics and finite-element method for crash safety. In 2005 he joined the Faculty of Civil Engineering and GeoSciences, Delft University of Technology, where he taught and performed research on railway engineering. He is currently an Associate Professor with Delft University of Technology. His research interests include numerical solution of frictional rolling contact and its applications to analyses of wear, rolling contact fatigue, and vehicle dynamics; train–track interaction, particularly in the high-frequency/short-wave range and at switches and crossings; condition monitoring of tracks; and friction adhesion between wheel and rail.

Alfredo Núñez (M'10) received the Ph.D. degree in electrical engineering from Universidad de Chile, Santiago, Chile, in 2010.

He was a Postdoctoral Researcher with the Delft Center for Systems and Control. He is currently with the Section of Road and Railway Engineering, Delft University of Technology, Delft, The Netherlands. He authored the book *Hybrid Predictive Control for Dynamic Transport Problems* in the Series of Advances in Industrial Control (Springer-Verlag, 2013). His research interests include monitoring of railway infrastructure, modeling and control of traffic and transportation systems, model predictive control, and fuzzy systems.

Rolf Dollevoet received the M.Sc. degree in mechanical engineering from Eindhoven University of Technology, Eindhoven, The Netherlands, and the Ph.D. degree from the University of Twente, Enschede, The Netherlands, in 2010.

Since 2003 he has been with ProRail, Utrecht, The Netherlands. Since 2012, he has been appointed as a part-time Professor with the Section of Road and Railway Engineering, Delft University of Technology, Delft, The Netherlands. He is also currently the leader of the scientific research on rolling contact fatigue and wheel/rail interface within ProRail Civil Engineering, and the Chair of the International Union of Railways working group on wheel/rail conditioning and lubrication.

Dr. Dollevoet received the Jan van Stappen Spoorprijs 2010 award (a yearly prize for contributions to the travel quality and service for passengers in Netherlands) from the railway sector for his Ph.D. research and its huge potential to reduce track maintenance costs.