Monitoring of Turnout Ballast Degradation Using Statistical Low-Complexity Behavioral Models

Barkhordari, Pegah; Galeazzi, Roberto; Blanke, Mogens

Published in: 
I E E E Transactions on Control Systems Technology

Link to article, DOI: 
10.1109/TCST.2020.2974697

Publication date: 
2020

Document Version 
Peer reviewed version

Link back to DTU Orbit

Citation (APA): 

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Monitoring of Turnout Ballast Degradation Using Statistical Low-Complexity Behavioral Models

Pegah Barkhordari, Roberto Galeazzi, Member, IEEE, and Mogens Blanke, Senior Member, IEEE

Abstract—The dependability of the railway infrastructure is paramount to guarantee safety, comfort, and network capacity. Turnouts are the key infrastructure element to enable the maximization of network capacity and minimization of transport delays. Their failure upsets the overall performance of the railway network; hence, infrastructure managers are interested in securing as high as possible uptime. The ballast layer provides the elastic support to the superstructure (sleepers and rail); thereby, it is crucial to ensure both safety and comfort in railway transport. This article presents a novel detection system for the monitoring of ballast degradation throughout its service life. A statistical model based on the generalized extreme value distribution is proposed to describe the behavior of the resonance frequency associated with the ballast. The generalized likelihood ratio test is then adopted to detect when the state of health of the ballast is out of control. The monitoring system is tested exploiting full-scale measurements of train-induced track vertical acceleration collected at a turnout of the Danish railway network over a two-year period, which includes a maintenance event. Results demonstrate the ability of the ballast monitoring system in detecting the progressive degradation of the ballast quality.

Index Terms—Condition monitoring, generalized extreme value (GEV) distribution, generalized likelihood ratio test (GLRT), low-complexity behavioral model, railway ballast degradation.

I. INTRODUCTION

RAILWAY transport has become a popular mode of modern transportation in Europe, playing a key role in conveying an enormous number of passengers and large amounts of freight. Increasing the capacity and reliability of the railway network is an essential goal for infrastructure managers. Such a goal may not be achieved without guaranteeing safety and continuous availability of railway track components. Switches and crossings (S&Cs), also known as turnouts, are important elements of the railway network, playing a vital role in providing flexibility for train operations. The safety of train operation is strongly affected by the performance of S&Cs. According to the 2014 RSSB Annual Safety Performance Report [1, Sec. 8.5], 31% of the track-related derailments were caused by S&Cs malfunctioning in the period 2009–2014 in Great Britain. A significant amount of the maintenance budget is, therefore, spent on maintenance and renewal actions of S&Cs. Banedanmark, the Danish railway infrastructure manager, estimates that each year one-third of the total track maintenance cost is attributed to turnouts [2].

The railway track consists of different elements: rail, sleepers, rail pad, ballast, and the underlying subgrade. The ballast is an important track element influencing not only railway track performance and safety but also the ride quality and passenger comfort. Particularly, the performance of S&Cs is significantly affected by the condition of the ballast layer. According to the failure data recorded in the U.K. in 2009 [3, Ch. 7], ballast degradation is the third most important component affecting turnout performance with a failing frequency of 7.9%. The periodic maintenance of the ballast, currently the main strategy adopted for assuring railway safety and reliability, is a major cost driver for infrastructure managers. Applying a predictive maintenance strategy based on condition monitoring of the infrastructure has the potential to significantly reduce the maintenance cost. Therefore, the development of an advanced condition-monitoring system for the timely detection of the degradation of railway components can play a significant role in actualizing the maintenance paradigm shift.

A. Literature Survey

Condition monitoring of the railway track, in its simplest form, can be performed through visual inspection [4]–[6]. This approach is effective for detecting surface defects on rails and sleepers but with limited applicability to ballast quality assessment. In recent years, different approaches have been employed to perform ballast degradation analysis. These include direct measurement methods, empirical predictive models, and model-based techniques.

In direct techniques, the quality of the ballast and subgrade layers is evaluated based on the data collected by different types of measurement systems. The ground-penetrating radar (GPR) [7]–[10], the cone penetration test (CPT) [9], and the track stiffness measurement vehicle [10], [11] are some well-known track inspection tools in this category widely used in the last decade for condition monitoring of the track substructure. These methods suffer from a number of drawbacks. CPT is a destructive and time-consuming test; the image quality of the GPR is strongly affected by the proper selection of the electromagnetic spectrum. By using the track...
stiffness measurement vehicle, it is not possible to provide early detection of track degradation, and the contribution of different components to track degradation cannot be isolated.

The development of predictive models is another approach for the investigation of railway track degradation. A number of empirical models have been proposed in the literature for predictive assessment of the long-term degradation. Main examples of such models can be found in [12]–[15]. Furthermore, prediction of track geometry degradation has been carried out by means of numerical models and an iterative procedure (see [16]–[18]).

Employing model-based techniques is the third approach to railway track degradation analysis. A notable research effort has been conducted by Lam et al. [19]–[21] for detection of ballast damage based on a combination of high complexity mechanical models and measured vibration data in model updating frameworks. The feasibility of quantifying ballast damage characteristics (location and level of stiffness reduction) in an artificially damaged system was studied using deterministic [19] and Bayesian [20] model updating techniques. A further investigation was carried out in [21] by employing the Markov chain Monte Carlo algorithm to ensure the applicability of the proposed methodology. The receptance function has been widely used in several research studies for the calibration of numerical model properties, such as ballast and rail-pad stiffness and damping coefficients [22]–[24].

The model-based approaches presented in the aforementioned articles feature high-dimensional models that lead to very complex diagnostic methods.

Approaches relying on reduced complexity models combined with impact excitation techniques have also been marginally investigated as a potential way to monitor the track structure. In [25], the structural condition of the ballasted railway track was evaluated using two-degrees-of-freedom (2-DOF) model. The 2-DOF model is obtained by employing the frequency response functions recorded during field test measurements. Prior work by Barkhordari et al. [26] and Barkhordari and Galeazzi [27] has also pursued the development of low-complexity behavioral models, exploiting both receptance test data and train-induced vibration data.

A fourth-order model representing key track components’ (ballast and rail-pad layer) behavior was identified using the eigenstructure realization algorithm (ERA) and unloaded vibration data recorded during receptance tests in [26]. Although the model was validated against the measured data induced by train passage, its identification relies on the execution of experimental campaigns that are rarely performed by track infrastructure managers. ERA is not suitable to perform system identification based on train-induced vibrations due to the following reasons that breach its working hypotheses: the vibration response of the infrastructure being loaded by a passing train is a forced response; in the train-operational speed range, the time interval between two consecutive excitations from two adjacent wheels of a bogie is shorter than the time required for the measured vertical acceleration to be completely damped. Therefore, the use of subspace system identification methods, such as N4SID, is proposed for the purpose of this study. In the continuation of the previous work in [27], statistical low-complexity models are estimated for the ballast layer at different locations along a railway turnout exploiting the loaded data, i.e., train-induced vertical accelerations. The estimation scheme is the combination of empirical mode decomposition (EMD) and numerical algorithms for subspace state-space system identification (N4SID). A statistical change detection algorithm is employed to develop a novel data-driven diagnostic tool for monitoring ballast degradation over time. Statistical change detection theory and algorithms have been successfully employed in a wide range of industrial applications for detecting faulty conditions as well as deviant system behaviors (see [28]–[33]).

B. Main Contribution

This article presents the design and validation of a monitoring system to supervise railway ballast degradation. Changes in ballast quality are addressed in terms of variation of ballast stiffness, which, in turn, relates to the first track resonance frequency.

Exploiting the modeling procedure presented in [27], a statistical model that describes the trend over time of the ballast resonance frequency for a turnout is introduced. The first track resonance frequency is estimated using the measured track vertical accelerations in loaded conditions. Measurements were collected over a period of almost two years at an S&C of the Danish railway infrastructure, and within that period, a ballast maintenance action (tamping) took place.

The generalized extreme value (GEV) distribution is proposed to model the estimated ballast resonance frequency, as it adequately explains the scatter of the estimates around the expected value. A ballast health monitoring system is then designed using the generalized likelihood ratio test (GLRT) for GEV distributed random processes. When tested on the full-scale data, the detector captures the sudden change in ballast quality around the occurrence of the tamping event and tracks the progressive deterioration of the ballast layer in the period following the maintenance action. This demonstrates that the proposed method is a valid approach to infer the quality of the ballast during its service life.

This article is structured as follows. Section II discusses the ballast monitoring problem and provides an overview of the railway turnout technology as well as the adopted measurement setup. Section III describes the method for estimating the track resonance frequencies. Section IV addresses the statistical modeling of the first track resonance frequency. Section V develops a detection method for ballast degradation monitoring. Section VI discusses the practical aspects of implementing the proposed monitoring method. Section VII analyzes the advantages and limitations of the proposed monitoring method. Finally, Section VIII draws some conclusions.

II. BALLAST MONITORING AND EXPERIMENTAL SETUP

A. Challenges in Ballast Monitoring

A railway track is an engineering structure consisting of different layers, as shown schematically in Fig. 1. The rails are fastened to the sleepers that are supported by a layer made
of crushed stones, namely, the ballast. This layer provides resistance to vertical, lateral, and longitudinal forces applied to the track. Therefore, the dynamic behavior of the ballast is an important factor that influences the overall track quality; hence, its maintenance is paramount [34]. Due to the complexity of the ballast layer and sensitivity of theoretical modeling to ballast material properties, it is extremely challenging to predict its dynamic behavior and long-term performance. In this study, monitoring the status of the ballast layer is performed by exploiting the train-induced track vertical acceleration response since features in the measured acceleration (frequencies and amplitudes) are directly connected to the quality of the ballast.

The consistent assessment of changes in ballast quality based on variations of features of interest in the measured vertical acceleration is arduous because there are several factors that may induce similar variations or hinder the retrieving of such features. These factors are categorized in intrinsic, such as curvature radius, type of sleepers, presence of undersleeper pads, and physical age of the components, and extrinsic, such as the quality of the measurement system/sensors, weather condition, train type, traveling speed of the trains, quality of wheels, and soil condition. This calls for the robust design of a monitoring system to achieve high insensitivity toward irrelevant temporary changes in one or more of the features connected to the ballast quality.

B. Railway Turnout Technology

A railway turnout is composed of three main elements: the switch panel, the closure panel, and the crossing panel. The switch panel refers to the part of the turnout through which a moving vehicle is guided to the desired direction. The switch panel consists of two half sets of switches, where the so-called diverging track begins to diverge from the main track. The half set of switches consists of a stock rail and a switch blade; the latter is a movable part directing the vehicle toward the main track or diverging track. The switch panel begins at the first rail joint in the turnout entry and ends at switch heels.

The closure panel is the part of the turnout that connects the switch panel and the crossing panel. The crossing panel is the part of the turnout separating the main track from the diverging track. It consists of a crossing, check rails, and stock rails. The crossing is a track component at which two opposite running edges of the turnout intersect each other. It is comprised of a crossing nose and two wing rails. The last set of rail joints located at the closure panel and the last set of rail joints after the crossing are considered as the front and end parts of the crossing panel, respectively.

C. Experimental Setup

Fig. 2 shows the sensors’ layout of the instrumented S&C located at Tommerup station in the Danish railway network. The track-side measurement system consists of three-wheel detectors, 12 two-axis accelerometers (measurement range: ±500 g), and three displacement sensors (measurement range: ±20 mm). All sensors are connected to a multichannel data acquisition system, where the signals are stored after being conditioned with bandpass filters and amplifiers.

The wheel detectors are used to trigger the automatic data recording during the passage of each train along the turnout. Magnetic bases are used to install the accelerometers on the rail web. Both vertical and lateral accelerations of the rail are recorded. The vertical displacement of the sleepers is measured using the displacement sensors. The data acquisition board samples all measurements at the sampling frequency $F_s = 20$ kHz. Pictures of the instrumented S&C and of selected sensors are shown in Fig. 3. An example data set, including the normalized vertical acceleration and

![Fig. 1. Railway track cross section with highlighted components.](image1)

![Fig. 2. Layout of the sensors’ location along the turnout at Tommerup station (Fyn, Denmark). The data utilized in this work refer to vertical accelerations measured by accelerometers A2, A7, and A11.](image2)

![Fig. 3. Pictures of the instrumented S&C nearby Tommerup station (Fyn). (a) Overview of the turnout in the proximity of the Tommerup station. (b) Accelerometer A5 magnetically connected to the rail web. (c) Displacement sensor resting on the sleeper. (d) Wheel detector mounted in proximity of the rail web.](image3)
time-shifted output of the wheel detector, used in the analysis is shown in Fig. 4. The normalized acceleration is calculated as \( \tilde{a} = a / \max(|a|) \). Measured data presented in this article are anonymized to comply with the policy of the Danish railway infrastructure manager.

To determine the time instants when the train wheels pass through the measurement locations A2 (switch panel), A7 (crossing panel), and A11 (transition zone between the turnout and open track), the time-shifted output of the wheel detector is exploited. The wheel passage signal \( WS(t) \) synchronized with the vertical acceleration \( a(t) \) allows the automatic slicing of the acceleration signal into subsignals corresponding to each train bogie, which provides the data needed for the model identification presented in Section III-D. The time delay \( \tau \) is computed based on the known train speed and distances between the wheel detection sensor and the considered accelerometers.

In addition to the on-site measurement system, an axle load checkpoint (ALC) system is installed in a nearby station located about 7 km from Tommerup station. The ALC system provides information about axle and wheel loads, which is exploited to select the trains with lower axle loads and minimum axle load variation. The track vertical accelerations induced by the passage of these trains will be used to carry out the model identification and subsequent statistical analysis. This is further discussed in Section IV.

III. PROPOSED METHODOLOGY FOR TRACK RESONANCE ESTIMATION

In [27], a method for the identification of low-complexity behavioral models of track dynamics was proposed. The block diagram shown in Fig. 5 provides an overview of the proposed methodology. The methodology can be employed to estimate both the first (ballast) and the second (rail pad) track resonance frequencies. However, this study is intended to analyze the ballast degradation, and therefore, only the first track resonance frequency is estimated.

To estimate the ballast resonance frequency, the measured accelerations induced by train passages are low-pass filtered using a cutoff frequency of 1 kHz. The obtained signal is then sliced at every two adjacent wheels. In order to extract the ballast vibration characteristics from the measured acceleration, it is essential to decompose the signal into its orthogonal components. The EMD algorithm is adopted for this purpose.

A subspace identification method is then used to find a model for the component describing the ballast dynamic behavior. The median over all the frequencies estimated by using signal slices is computed to improve the robustness of the estimation. A detailed description of the different steps of the proposed method is provided in the following.

A. Filtering

The signal to be decomposed through EMD is obtained by preprocessing the measured vertical acceleration. The first step of preprocessing is signal filtering to remove unwanted frequencies. According to the literature (see [35]) and the analysis of the receptance test presented in [26], the frequency interval containing the first and second resonance frequencies of the track is \([0, 1000]\) Hz. Therefore, the measured vertical acceleration in response to train passages is filtered using a third-order Butterworth low-pass filter with the cutoff frequency of 1000 Hz.

B. Signal Slicing

Let \( \tilde{a}(t) \) for \( t \in [t_0, t_1] \) (where \( t_0 \) and \( t_1 \) are the time instants when the train enters and leaves the turnout, respectively) be the filtered normalized acceleration response of a specific location of the turnout to the passage of a train \( T \) with \( N_b \) bogies, and \( B = \{b_1, \ldots, b_{N_b}\} \) the set of train bogies. \( \tilde{a}(t) \) is sliced to \( N_b \) segments representing different time intervals corresponding to the passage of bogies at the measurement location. This is carried out based on the time-shifted output of the wheel detector. In other words, \( \tilde{a}_k(t) \) is the acceleration response to the passage of the \( k \)th bogie measured in the time interval \( t \in [t_{01,k}, t_{02,k} + \delta] \), where \( t_{01,k} \) is the time instant the first wheel passes over the accelerometer, \( t_{02,k} \) is the time instant at which the second wheel passes over the sensor, and \( 0 < \delta < t_{01,k+1} \). The slicing of the signal is carried out to attenuate the effect of the parametric frequencies and their harmonics.

The parametric frequencies are the excitation frequencies induced by the periodic characteristics of the track and train. The sleeper passing frequency (\( \nu/d_s \)) is a track-related parametric frequency present in the frequency response of a track with the sleeper distance \( d_s \) under wheel loads moving at speed \( \nu \). Other parametric frequencies are train-related frequencies (\( \nu/d_w \) and \( \nu/d_B \), where \( d_w \) and \( d_B \) are the distance between wheels and the distance between bogies, respectively), which are induced by moving wheel loads and bogies periodically distributed along the train [36]. For the type of train considered in this study (IC3) and for the chosen traveling speed (160 km/h), the sleeper passing frequency and the train-related parametric frequencies induced by moving bogies and wheels are calculated as 74.07, 3, and 17.5 Hz, respectively.

The presence of the parametric frequencies in the track response is discussed in detail in [36]. Fig. 6 illustrates the...
C. Empirical Mode Decomposition

In order to analyze the measured data and to extract the information required in the identification step, an advanced signal processing method is employed. The signal processing technique adopted in the framework of the proposed monitoring method should preferably characterize dynamic components of the measured signals in the time domain and relate them to the dynamic behavior of track layers within certain frequency intervals. Therefore, EMD is chosen in this study.

EMD is a signal processing method developed by [37] suitable for decomposing nonstationary and nonlinear signals in the time–frequency domain. The method has been applied successfully to structural fault diagnosis problems [38], [39]. The conventional EMD algorithm has some drawbacks, such as “mode mixing” and “end effects” [40], the latter of which significantly affects the decomposition of the signals considered in this study. Moreover, the stopping criterion in the sifting process of the conventional EMD is met when the standard deviation of two consecutive sifting components is smaller than an empirical threshold that has no physical meaning. These problems have been addressed thoroughly in [40], and a modified EMD algorithm has been proposed to overcome such drawbacks. Such an algorithm is employed in this study, and it is summarized in Algorithm 1.

By applying the algorithm to \( \tilde{a}_k(t) \), the signal is decomposed through the sifting process into its so-called intrinsic mode functions (IMFs). IMFs represent the fundamental oscillatory modes of the original signal. The sifting process stops when the energy ratio of the original signal to the residual \( r(t) \), \( Q_{\text{residual}} \), is smaller than a predetermined value \( \epsilon_1 = 40 \) dB. The stopping criterion for the sifting process associated with \( j \)th IMF is created based on a resolution factor defined as the ratio of the energy of the signal at each iteration to the energy of the average of the envelopes \( Q_{\text{resolution}} \). This criterion is met when the current energy ratio is larger than a resolution factor \( \epsilon_2 = 45 - 50 \) dB. The extracted IMFs corresponding to the dynamic behavior of the ballast layer are utilized in the identification step for identifying a low-complexity behavioral model.

D. N4SID Subspace Identification Method

The N4SID algorithm was proposed for LTI systems in order to identify state-space models capable of representing the system input–output behavior [41]. Since each IMF represents one dominant oscillatory mode of the original signal, a second-order model is proposed to describe each IMF behavior.

The canonical realization of the second-order discrete-time state-space model describing the \( j \)th IMF can be defined as

\[
\begin{align*}
\hat{x}_j(k + 1) &= \hat{A}_j \hat{x}_j(k), \quad \hat{x}_j \in \mathbb{R}^2 \\
\hat{y}_j(k) &= \hat{C}_j \hat{x}_j(k), \quad \hat{y}_j \in \mathbb{R}.
\end{align*}
\]

For the \( j \)th IMF extracted from the EMD algorithm, the past and future output data in the N4SID algorithm are given as

\[
\hat{x}_j(k) = \left[ \begin{array}{c}
\hat{x}_{j1}(k) \\
\hat{x}_{j2}(k)
\end{array} \right], \quad \hat{y}_j(k) = \left[ \begin{array}{c}
\hat{y}_{j1}(k) \\
\hat{y}_{j2}(k)
\end{array} \right]
\]
Algorithm 1 EMD Adopted From [40]

Data: Track vertical acceleration $\dddot{a}_k(t)$

Result: Set of intrinsic mode functions

$\mathcal{X}_k = \{\text{IMF}_1, \ldots, \text{IMF}_m\}$

1. Residual $r(t) \leftarrow \dddot{a}_k(t)$;
2. $i \leftarrow 1$;
3. while $Q_{\text{residual}} < \epsilon_1$ do
    4. $h_{10}(t) \leftarrow r(t)$;
    5. $j \leftarrow 1$;
    6. repeat
        7. Find all local extrema $h_{1(j-1)}(t)$;
        8. Extrapolate the extremum;
        9. Find the upper and lower envelopes by spline interpolation of the local extrema;
        10. Compute the mean value $m(t)$ between the lower and upper envelopes;
    11. Compute $Q_{\text{resolution}}$:
        $$Q_{\text{resolution}} = 10 \log_{10} \left( \frac{\|h_{12}(t)\|^2}{\|m(t)\|^2} \right)$$
    12. Subtract the mean value from the original signal $h_{1j}(t) = h_{1(j-1)}(t) - m(t)$;
    13. $j \leftarrow j + 1$;
    14. until $Q_{\text{resolution}} > \epsilon_2$;
    15. $\text{IMF}_i(t) \leftarrow h_{1(j-1)}(t)$;
    16. $r(t) \leftarrow r(t) - \text{IMF}_i(t)$;

Compute $Q_{\text{residual}}$:

$$Q_{\text{residual}} = 10 \log_{10} \left( \frac{\|\dddot{a}_k(t)\|^2}{\|r(t)\|^2} \right)$$

$i \leftarrow i + 1$;

end

$Y_p = (\text{IMF})_{0:k-1} \in \mathbb{R}^{k \times N}$ and $Y_f = (\text{IMF})_{k:2k-1} \in \mathbb{R}^{k \times N}$, where $N$ is assumed to be sufficiently large and $k$ is strictly larger than $n = 2$. The past and future input vectors in the identification process (i.e., $U_f$ and $U_p$) are set to zero since the input signal (i.e., train excitation) is unknown. Defining $W_p$ as $[U_p^T, Y_p^T]^T$, the N4SID identification algorithm consists of the following steps [42].

1) Compute the oblique projection of $Y_f \in \mathbb{R}^{k \times N}$ onto $W_p \in \mathbb{R}^{2k \times N}$ along $U_f \in \mathbb{R}^{k \times N}$ (i.e., $\zeta$) by employing (2) and the LQ decomposition in (3)

$$\zeta = \hat{E}_{[U_f^T,Y_f^T,W_p]} = \text{R}_{22}R_{22}^\dagger W_p = O_2X_f$$

where the oblique projection is expressed as $\hat{E}_{[\cdot]}$, $O_2 \in \mathbb{R}^{k \times 2}$ is the extended observability matrix, and $X_f$ is the future state vector

$$\begin{bmatrix} U_p \\ W_p \\ Y_f \end{bmatrix} = \begin{bmatrix} \text{R}_{11} & 0 & 0 \\ \text{R}_{21} & \text{R}_{22} & 0 \\ \text{R}_{31} & \text{R}_{32} & 0 \end{bmatrix} \begin{bmatrix} \text{Q}_1^T \\ \text{Q}_2^T \\ \text{Q}_3^T \end{bmatrix}$$

where $\text{R}_{11} \in \mathbb{R}^{k \times k}$ and $\text{R}_{22} \in \mathbb{R}^{2k \times 2k}$ are the upper triangular matrices, and $\text{Q}_1 \in \mathbb{R}^{N \times k}$, $\text{Q}_2 \in \mathbb{R}^{N \times 2k}$, and $\text{Q}_3 \in \mathbb{R}^{N \times k}$ are the orthogonal matrices. $\text{R}_{22}$ is the pseudoinverse of $\text{R}_{22}$.

2) Compute the state vector $X_f$ using the singular value decomposition of $\zeta$ as

$$\zeta = U_1\Sigma_1V_f^T = O_2X_f$$

$$O_2 = U_1\Sigma_1^2T$$

$$X_f = T^{-1}\Sigma_1^2V_1^{-1}$$

where $X_f = [x_j(k), x_j(k+1), \ldots, x_j(k+N-1)] \in \mathbb{R}^{2k \times N}$.

3) Consider the regression equation

$$\begin{bmatrix} \tilde{X}_{k+1} \\ \bar{Y}_{k|k} \end{bmatrix} = \begin{bmatrix} A_j & B_j \\ C_j & D_j \end{bmatrix} \begin{bmatrix} \hat{X}_k \\ \bar{U}_{k|k} \end{bmatrix}$$

with $\tilde{X}_{k+1}, \hat{X}_k, \bar{U}_{k|k}$, and $\bar{Y}_{k|k}$ defined as

$$\tilde{X}_{k+1} = [x_j(k+1), \ldots, x_j(k+N-1)] \in \mathbb{R}^{2N-1}$$

$$\hat{X}_k = [x_j(k), \ldots, x_j(k+N-2)] \in \mathbb{R}^{2N-1}$$

$$\bar{Y}_{k|k} = [x_j(k), \ldots, x_j(k+N-2)] \in \mathbb{R}^{1 \times N-1}$$

4) Compute the matrices $\hat{A}_j, \hat{B}_j, \hat{C}_j, \hat{D}_j$ through solving the regression equation by using the least-squares technique

$$\begin{bmatrix} \hat{A}_j & \hat{B}_j \\ \hat{C}_j & \hat{D}_j \end{bmatrix} = \left( \begin{bmatrix} \tilde{X}_{k+1} \\ \bar{Y}_{k|k} \end{bmatrix} \begin{bmatrix} \hat{X}_k \\ \bar{U}_{k|k} \end{bmatrix} \right)^{-1}$$

where $\tilde{X}_{k+1} \in \mathbb{R}^{2N-1}$ is the estimated future states, $\bar{Y}_{k|k} \in \mathbb{R}^{1 \times N-1}$ is the future output data, and $\bar{U}_{k|k} \in \mathbb{R}^{1 \times N-1}$ is the future input data.

E. Frequency Estimation

The system dynamical matrix $\hat{A}_j$ and the output matrix $\hat{C}_j$ are expressed in terms of the natural frequency ($\omega_{nj}$) and damping factor ($\zeta_j$) of the equivalent continuous-time model in the following equation:

$$\hat{A}_j = \exp \left( \begin{bmatrix} 0 & 1 \\ -\omega_{nj}^2 & -2\zeta_j\omega_{nj} \end{bmatrix} \times T_s \right), \quad \hat{C}_j = [1 \ 0]$$

where $T_s$ is the sampling time. The natural frequency and the damping factor in the model identified for the $j$th IMF can be obtained by means of eigenvalue analysis

$$\omega_{nj} = \frac{\ln(\lambda_1(\hat{A}_j))}{2\pi T_s}, \quad \zeta_j = \frac{-\Re(\ln(\lambda_1(\hat{A}_j)))/T_s}{|\ln(\lambda_1(\hat{A}_j))|}$$

where $\lambda_1(\hat{A}_j)$ is the first eigenvalue of the complex pair associated with the matrix $\hat{A}_j$. 

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.
F. Low-Complexity Behavioral Model

The method introduced in Section III-D is now applied to the acceleration slice \( \tilde{a}_k \) to find a low-complexity behavioral model representing the ballast layer dynamics. Fig. 7 shows an example of the acceleration response recorded for a passage of two adjacent wheels and the four extracted IMFs that are considered as identification data sets. The N4SID method is then applied to each identification data set to identify a state-space model representing the dominant dynamic behavior of the ballast. The MATLAB function \texttt{n4sid} is used to estimate the models. It can be shown that only the excited frequencies calculated for the fourth IMF are within the frequency range in which the first track (ballast) resonance frequency is expected according to the literature [27], [35].

The discrete-time state-space realization of the identified model for IMF4 associated with the data set shown in Fig. 7 is

\[
\begin{bmatrix}
\hat{A}_4 & \hat{B}_4 \\
\hat{C}_4 & \hat{D}_4
\end{bmatrix} =
\begin{bmatrix}
0.9983 \pm 6.3 \cdot 10^{-4} & 0.0443 \pm 6.1 \cdot 10^{-4} \\
-0.0443 \pm 6.1 \cdot 10^{-4} & 0.9983 \pm 6.3 \cdot 10^{-4}
\end{bmatrix}
\begin{bmatrix}
0.7073 \pm 0.65 \cdot 10^{-4} & -0.0209 \pm 0.63 \cdot 10^{-4}
\end{bmatrix}.
\]

The confidence interval of the estimated parameters of the model \( \mathcal{M}_4 \) is given by the Cramer–Rao lower bound covariance matrix [43], which is the lowest attainable mean square error. As shown in [44], the subspace identification does not converge to this bound, which then represents the best case scenario. However, the precise computation of the asymptotic variance for subspace estimates is still an open and much researched problem [45]. This estimate of the confidence interval is directly provided by the MATLAB function \texttt{n4sid}. The confidence interval of the parameter estimates is provided to solely evaluate the validity of the identified model.

Similarly, a model representing the dynamic behavior of the rail pad can be obtained using the second extracted IMF [27]. These two models can then be combined through the output into a unified model able to predict the dominant behavior of the measured track vertical acceleration in relation to the ballast and rail-pad components. The robustness of the overall model and its capability in predicting the measured accelerations in response to trains excitation were evaluated in [27].

IV. STATISTICAL CHARACTERIZATION OF THE BALLAST RESONANCE FREQUENCY

For each train passage, a robust estimate of the ballast resonance frequency \( \omega_n,1 \) is obtained applying the median to the \( \hat{\omega}_n \) estimated frequencies. Due to the simplicity of the chosen modeling approach, many sources of uncertainty can affect the estimation process—wheel defects, track defects, pollution of the ballast layer, temperature gradients, and so on. Therefore, to account for this, a statistical model is built by considering all estimates within a month period.

In [46], the effect of different train parameters (i.e., type, speed, and axle load) on the predictive capability of the low-complexity behavioral model was investigated. It was concluded that the train speed and axle load significantly affect the estimation accuracy of the track resonance frequencies. It was shown that the low-complexity behavioral models estimated based on the receptance test have their best performance in predicting the dominant behavior of the track vertical acceleration for high-speed trains with low axle load. Table I presents the mean value and 1\( \sigma \) standard deviation of the axle loads calculated by considering all IR4, IC3, and freight trains passing through Tommerup station during one month period. For the IC3 trains, the measured axle load has the lowest variance and a relatively small mean value compared with other train types. Therefore, the IC3 trains with the speed range of [158–160] km/h are selected for further analysis in this study.

Fig. 8 shows the time period when the installed track-side measurement system collected the data, the quality of the measured data, and a tamping event. Three-time periods are distinguished in this time chart: TP1 in which poor-quality data...
were collected by the measurement system; TP2 in which data are missing due to maintenance of measurement system; and TP3 in which good quality data are collected.

A. Statistical Model for Ballast Resonance Frequency

The measured data collected over monthly periods at different locations along the turnout are utilized to obtain a statistical characterization of the ballast resonance frequency. Fig. 9 shows probability plots of the first track resonance at locations A2, A7, and A11 obtained for the month right before (September 2016) and the month right after (November 2016) the tamping event in TP1. A distinct change in the statistical characteristics of the first track resonance before ($\omega_{n1, BT}$) and after ($\omega_{n1, AT}$) tamping is evident at all locations. This suggests that the estimated ballast resonance frequency can be employed as an indicator of changes in the ballast layer stiffness due to degradation and maintenance actions.

To develop a ballast degradation detection scheme, it is important to determine which distribution better represents the statistical behavior of the first track resonance. The Gaussian distribution ($N$), the Weibull distribution ($W$), and the GEV distribution ($G$) have been selected to describe the behavior of the ballast frequency estimates. Figs. 10–12 show the probability plots for two months: a month before the tamping and a month after the tamping in TP1. It can be observed that the GEV distribution fits the data much closer than either Gaussian or Weibull distributions, especially in correspondence of the tails. This is paramount for detecting the change in the status of the ballast from healthy (after tamping/renewal) to degrade, and vice versa.

Estimates of the parameters of the considered distributions for the data presented in Figs. 10–12 are listed in Table II together with the $p$-value obtained using the Kolmogorov–Smirnov (K–S) goodness-of-fit test with a confidence level of 95%. Data at locations A2 and A7 are well approximated by the GEV distribution; however, none of the three distributions gives a good fit to the data at location A11 before the tamping event. This is due to the lack of good quality data at this location before the tamping event. To overcome this problem and to provide a sufficient statistic for the ballast resonance frequency characterization, a pool of nonidentical trains (i.e., IC3 and IR4) is taken into account at A11. Since representing the behavior of the data at different locations by using a unique distribution is of interest, the GEV distribution was selected as the best in comparison to the Normal and Weibull distributions. The probability density function of the GEV distribution is given by

$$G(x) = \frac{1}{\sigma} \exp \left[-\left(1 + \frac{x - \mu}{\xi \sigma}\right)^{-\frac{1}{\xi}}\right] \left(1 + \frac{x - \mu}{\xi \sigma}\right)^{-1-\frac{1}{\xi}}$$

(9)
TABLE II

<table>
<thead>
<tr>
<th>Location</th>
<th>Ballast resonance</th>
<th>Distribution</th>
<th>Parameters</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>( \omega_{n1,BT} )</td>
<td>( \mathcal{G} )</td>
<td>( \xi = -0.18, \sigma = 6.59, \mu = 158.85 )</td>
<td>0.05</td>
</tr>
<tr>
<td>A2</td>
<td>( \omega_{n1,AT} )</td>
<td>( \mathcal{W} )</td>
<td>( \alpha = 2.89, \beta = 21.10, \mu = 142.78 )</td>
<td>0.02</td>
</tr>
<tr>
<td>A2</td>
<td>( \omega_{n1,AT} )</td>
<td>( \mathcal{N} )</td>
<td>( \mu = 6.99, \mu = 161.62 )</td>
<td>0.02</td>
</tr>
<tr>
<td>A7</td>
<td>( \omega_{n1,BT} )</td>
<td>( \mathcal{G} )</td>
<td>( \xi = -0.61, \sigma = 8.36, \mu = 187.41 )</td>
<td>0.04</td>
</tr>
<tr>
<td>A7</td>
<td>( \omega_{n1,AT} )</td>
<td>( \mathcal{W} )</td>
<td>( \alpha = -0.07, \beta = 2.03, \mu = 87.01 )</td>
<td>0.06</td>
</tr>
<tr>
<td>A7</td>
<td>( \omega_{n1,AT} )</td>
<td>( \mathcal{N} )</td>
<td>( \mu = 1.80, \beta = 9.35, \mu = 85.08 )</td>
<td>0.001</td>
</tr>
<tr>
<td>A11</td>
<td>( \omega_{n1,BT} )</td>
<td>( \mathcal{G} )</td>
<td>( \xi = -0.14, \sigma = 3.32, \mu = 91.17 )</td>
<td>0.69</td>
</tr>
<tr>
<td>A11</td>
<td>( \omega_{n1,AT} )</td>
<td>( \mathcal{W} )</td>
<td>( \alpha = 3.67, \beta = 9.26, \mu = 92.68 )</td>
<td>0.12</td>
</tr>
<tr>
<td>A11</td>
<td>( \omega_{n1,AT} )</td>
<td>( \mathcal{N} )</td>
<td>( \mu = 10.77, \beta = 76.24, \mu = 104.01 )</td>
<td>5.30 ( \times 10^{-7} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \sigma = 8.67, \mu = 174.41 )</td>
<td>1.75 ( \times 10^{-10} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \xi = -0.27, \sigma = 9.92, \mu = 181.87 )</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \alpha = 2.56, \beta = 27.11, \mu = 161.50 )</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \mu = 10.23, \mu = 185.54 )</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Fig. 13. Probability plot obtained over a month before and after tamping in TP1 and during TP3 period for A2 location.

defined on the set \( \{ x : 1 + \xi(x - \mu)/\sigma > 0 \} \), where \( \mu, \sigma, \) and \( \xi \neq 0 \) are the location, scale, and shape parameters, respectively. The GEV distribution consists of three different subfamilies: Type I (\( \xi = 0 \)), Type II (\( \xi > 0 \)), and Type III (\( \xi < 0 \)).

Fig. 13 shows the distributions of monthly estimates of the first track resonance frequency over a period of two years at location A2. A similar trend has been observed at locations A7 and A11. At all locations and in different measures, it is evident the gradual change in the probability distribution addressing possible variations in the quality of the ballast layer stiffness. To identify which parameters of the GEV distribution play a key role in explaining the changes in the distribution, the maximum likelihood estimation (MLE) is applied to the available data.

Fig. 14 shows the MLE of the GEV parameters for the estimated ballast resonance frequency at location A2. The MLE was computed every two-week interval during TP1 (before and after the tamping event) and TP3. The location parameter \( \mu \) decreases gradually, indicating the progressive degradation of the ballast layer due to changes of stiffness. The erratic behavior of the scale and the shape parameters over time is due to the presence of uncertainties imposed by, e.g., environmental conditions and dynamic loading components caused by wheel defects. In order to embrace the effect of both the gradual variation of the location parameter and the uncertainties, the mean value (first moment) of the GEV distribution is used to describe the total variation in the data.

The mean value calculated using the estimated parameters is given by [47, Ch. 6]

\[
\mu_{\text{prob}} = \left\{ \begin{array}{ll}
\mu + \frac{\sigma (\Gamma(1-\xi) - 1)}{\xi}, & \xi < 1, \xi \neq 0 \\
\mu + \sigma \gamma, & \xi = 0 \\
\infty, & \xi \geq 1
\end{array} \right.
\]

where \( \gamma \) is the Euler constant and \( \Gamma(\cdot) \) is the gamma function.
developed first. The test statistic is developed applying the GLRT approach based on the GEV distribution. The GLRT is chosen because it enables the design of a flexible monitoring tool able to address a priori unknown changes.

A. $G$-GLRT Detector Design

Assume that the estimated first track resonance frequency $\hat{\omega}_{n,1}$ obtained for faulty and healthy cases (before and after tamping) is a realization of random GEV “type III” process. The distribution of $N$ independent identically distributed (i.i.d.) samples, denoted as $x = [\hat{\omega}_{n,1}(k - N), \ldots, \hat{\omega}_{n,1}(k)]^T$ for the sake of simplicity, is characterized using the following probability density function:

$$G(x; \theta) = \frac{1}{\sigma} \prod_{i=1}^{N} \exp\left(-\frac{1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)^{-\frac{1}{\xi}}}{\theta}\right) \times \prod_{i=1}^{N} \left(1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)^{-1}$$  \hspace{1cm} (11)

where $N$ is the total number of IC3 trains in the speed range [158, 160] km/h passing through the monitored S&C over monthly periods. In Section VII-C, the i.i.d. assumption will be verified and discussed considering the estimates of the first track resonance frequency obtained over a month.

The GLRT detector decides $H_1$ if

$$L_G(x) = \frac{G(x; \hat{\theta}_1, H_1)}{G(x; \theta_0, H_0)} > \gamma$$  \hspace{1cm} (12)

where $\hat{\theta}_1$ is the MLE of the GEV parameters under the hypothesis $H_1$ and $\gamma$ is the threshold calculated to obtain the desired probability of false alarms. To compute $L_G(x)$, the unknown parameters are estimated by maximizing $G(x; \hat{\theta}_1, H_1)$ with respect to the parameter vector. The log-likelihood for the GEV parameters when $\xi_1 \neq 0$ is [50]

$$-N \ln \sigma - \left(1 + \frac{1}{\xi_1}\right) \sum_{i=1}^{N} \ln \left[1 + \xi_1 \left(\frac{x_i - \mu_1}{\sigma_1}\right)\right] - \sum_{i=1}^{N} \left[1 + \xi_1 \left(\frac{x_i - \mu_1}{\sigma_1}\right)\right]^{-1}$$  \hspace{1cm} (13)

provided that

$$\left[1 + \xi_1 \left(\frac{x_i - \mu_1}{\sigma_1}\right)\right] > 0.$$

From (13), the MLE $\theta_1$ is computed using standard numerical optimization algorithms since an analytic solution does not exist.

By substituting (11) into (12), taking the natural logarithm of both sides of the resultant equation and replacing $\theta_1$ parameters with their estimates and $\ln \gamma$ with $\gamma$, it is possible to derive the explicit form of the detector. For the $k_{th}$ window containing, a batch of data the test statistic $L_G(x)$ is

The estimated mean value and its 95% confidence interval for different GEV distributions at A2, A7, and A11 are shown in Fig. 15. The rate of change of $\mu_{prob}$ at location A7 is lower than at the other two locations along the turnout; this suggests that the degradation rate at the crossing panel is slower. This finding is in line with observations reported in [48] and [49] that indicated a high rate of settlements in the switch blade area (location A2) and the lowest rate in the area where the longer sleepers are (just before and around the crossing’s frog).

V. BALLAST DEGRADATION MONITORING SYSTEM

To detect changes in ballast quality, a statistical test is designed considering two alternative hypotheses: $H_0$ corresponds to the ballast status right after a tamping or renewal event, and $H_1$ describes the ballast status between the maintenance events. Since the mean value of the GEV distribution captures the total variation, both in terms of trend and dispersion, the test statistic is set up as follows:

$$H_0 : \mu_{prob} = \mu_{prob,0}$$
$$H_1 : \mu_{prob} \neq \mu_{prob,0}$$

where $\mu_{prob,0}$ is the mean value under the hypothesis $H_0$. Equation (10) shows that the mean value is function of the GEV parameters; therefore, the detection problem can be reformulated as parameter test, that is

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

where $\theta = [\mu, \sigma, \xi]^T$ is the vector of the GEV parameters and $\theta_0$ is the value it takes under the hypothesis $H_0$, which is assumed to be known.

To carry out such a test with the highest possible performance, i.e., the lowest number of false alarms and highest probability of detection, a proper test statistic should be
After the temperature drops below 0, the natural frequency of a ballasted bridge takes place. In [51], it is shown that a significant change in the estimate of the first track resonance frequency show the trend that correlates with variations in air temperature or the weather phenomena; therefore, it is meaningful to evaluate if the parameter vector \( \theta_0 \) needs to be estimated after every maintenance/renewal action in order to capture the actual stiffness characteristics of the ballast layer. In between the maintenance events, \( \theta_0 \) is known and constant.

The operational profile of the ballast monitoring system distinguishes two phases.

**Estimation Phase:** Let \( T_m \) be the time at which a ballast maintenance or renewal action takes place. Then, the MLE of the GEV parameter vector under the hypothesis \( H_0 \) is computed exploiting data collected during the passage of \( M_E \) trains of the chosen type (low axle load and high traveling speed). Once \( \theta_0 \) is available, the \( \mathcal{G} \)-GLRT detector switches into the detection phase.

**Detection Phase:** Using (14), the detector monitors the status of the ballast until a new maintenance action occurs. A detection window that is a function of the number of trains \( M_D \) passing through the turnout within a given time span is used to perform the periodic monitoring. The number of trains \( M_D \) and \( M_E \) influences the performance of the monitoring system in its ability to deliver a timely and accurate detection of ballast degradation. As soon as a tamping event has occurred, there is a settling time of three to six months when the ballast gravel settles under the loading of the passing trains. In this period, the ballast stiffness reaches a new steady-state value. Therefore, \( M_E \) should preferably be the number of trains corresponding to a minimum of three-month period after tamping. The size \( M_D \) of the detection window is a tradeoff between the promptness and the accuracy of the detection; a large \( M_D \) will induce a more monotonic behavior of the \( \mathcal{G} \)-GLRT index, whereas a small \( M_D \) will enable a more timely detection.

**B. Performance Evaluation**

The designed condition-monitoring system for ballast degradation is tested to verify its ability to track gradual and systematic variations of the ballast quality as the cumulative loading increases over time. In the performance evaluation,
the emphasis is put on the impact of different choices of $M_E$ and $M_D$ on achieving a monotonic behavior of the $G$-GLRT, thereby reducing the uncertainty about existing degradation trends. The performance evaluation does not address aspects of detection, which would require the characterization of ballast quality levels with respect to ranges of variation of the $G$-GLRT and the consequent selection of the threshold $\gamma'$ by trading off between desired probabilities of detection and false alarms.

Since there is a long time gap between the periods TP1 and TP3, the monitoring signal $x$ built in these two periods is analyzed separately. To evaluate the ability of the $G$-GLRT to address a change in ballast quality due to tamping, the first monitoring signal is constructed as $x = [\omega_{T,n,1BT}, \omega_{T,n,1AT}]^T$, where $\omega_{T,n,1BT}$ and $\omega_{T,n,1AT}$ are obtained using data of sufficient quality recorded during the two one-month periods right before and right after the tamping event in TP1. Noteworthy that due to generally reduced quality of the data collected during TP1, data collected in months other than September–November 2016 were not processed by the $G$-GLRT detector.

The effect of using three different window sizes (two weeks, one month, and three months) on the monitoring results has been evaluated in this study. In the analysis, it is assumed that $M_E = M_D$. It has been found that a window size of three months with two months overlap between two consecutive data windows provides a good tradeoff between the time resolution and accuracy. This configuration of the monitoring system is adopted only for the period TP3 since the low-quality data of the period TP1 prevent it. Hence, the data gathered during the month right before and right after tamping are evaluated individually with no overlap.

Figs. 17–19 show the obtained results for locations A2, A7, and A11. The color scale provided in Figs. 17–19 shows the value of $G$-GLR (status of the ballast) with respect to the $G$-GLR value calculated before the tamping event (%BT). A clear monotonic degradation trend is achieved for locations A2 and A7, despite the growth rate of the $G$-GLRT index is significantly different; however, some fluctuations of the index are still displayed for location A11.

VII. DISCUSSION

A. Ballast Degradation

The developed ballast monitoring system successfully tracks changes in ballast stiffness due to the ability of statistically discriminating variations of the first track resonance frequency. The analysis carried out in this study shows that the detection window size has a significant effect on reducing the uncertainty on the estimated level of ballast quality as seen from the output of the $G$-GLRT. The introduction of the window overlap enables the conjugation of contrasting needs as high time resolution (achieved through a small data batch size) and low uncertainty (obtained through a large data batch size). The tuning of the monitoring system with three months window size and two months overlap provides a detection window of one month and displays rather monotonic trends, thereby fulfilling the different demands.

The current practice for ballast quality monitoring adopted by railway infrastructure managers utilizes the standard deviation $\sigma_H$ of the longitudinal track level, which is measured every three months by a track geometry measurement car.
Such an indicator is computed over 200-m-long track sections in order to mitigate the undesired effects of spatial misalignment of the measured data introduced by the uncertainty of the GNSS-based geotagging, as recommended by the European standard [53]. An example of the $\sigma_H$ indicator for the track line passing through the Tommerup station is shown in Fig. 20. Zones C and D in Fig. 20 correspond to different track quality levels. There are five track quality classes (classes A–E) defined in the European standard based on the line speed and values of $\sigma_H$. Classes A and E represent the good and poor quality of the track, respectively. Three different sections (Sections 1–3) are considered by shifting the 200-m-long track window around the position of the turnout. $\sigma_H$ has been calculated for eight consecutive measurement campaigns over the considered period. Such a long time interval between the measurement campaigns is partly due to the slow degradation of the ballast layer as well as the large cost associated with the measurement campaigns. However, different turnouts degrade at different rates in relation to the capacity of the railway line where the S&C is placed and the environmental factors such as the subgrade condition.

A key advantage of the proposed ballast monitoring method, compared with the industry practice, is the capability of performing a localized analysis of ballast degradation at different sections along the turnout. Fig. 20 shows that considering a 200-m-long track section leads to an over or underestimation of the track quality along the turnout. This is caused by the inclusion of multiple turnouts and/or open track located within a 200-m-long section. Therefore, the spatial resolution of the method currently used in practice is not sufficient for the purpose of local monitoring of different short sections along the turnout (e.g., switch panel and crossing panel).

**B. Plug-and-Play Ballast Monitoring**

Due to its data-driven nature, the proposed ballast monitoring system is potentially highly portable across the entire railway network. A limiting factor to its network-wide applicability is the cost associated with the installation of a track-side measurement system. Despite the accelerometers are becoming increasingly cheap with the widespread of MEMS technologies, the deployment cost will, in general, be high because of a number of turnouts to be monitored and the necessary ruggedness that the sensors should have to withstand the harsh operational conditions. A cost-wise viable alternative to the track-side measurement system is the use of vibration sensors installed onboard in-service trains. However, this may require the modification of the ballast frequency estimation procedure since the measured vibrations would contain both train and track dynamics.

**C. Operational Tool for Infrastructure Managers**

The adoption of the ballast monitoring system from infrastructure managers for integration into the legacy systems devoted to maintenance planning requires the provision of actionable information from the monitoring system itself. As shown in Figs. 17–19, the $G$-GLRT index clearly shows the presence of a trend related to the change in ballast stiffness; however, the index value, per se, does not provide any immediate insight about the actual quality of the ballast.

Railway infrastructure managers plan maintenance actions when standard indicators chosen to characterize the track geometry quality are about to cross either operational or safety limits. A way toward industrial adoption of the ballast monitoring system would then be the identification of the frequency range of the first track resonance corresponding to a “sufficiently stiff” ballast from an operational/safety standpoint, thereby determining the threshold that distinguishes this scenario from the “insufficiently stiff” one.

Theoretically, if $\mathbf{x}$ is i.i.d., then based on $G(\mathbf{x}; \theta_0)$, the threshold $\gamma'$ for the desired probability of false alarms $P_{FA}$ is computed using the Neyman–Pearson theorem [54, Ch. 3] as

$$P_{FA} = \int_{L_{G}(\mathbf{x}) > \gamma'} G(\mathbf{x}; \theta_0) \, d\mathbf{x}. \quad (15)$$

Furthermore, if the dimension of $\mathbf{x}$ tends to infinity, then the modified test statistics $2L_{G}(\mathbf{x})$ asymptotically converges to a chi-squared distribution with $r$ degrees of freedom [54], thereby leading to a constant threshold.

Fig. 21 shows an example of the autocorrelation $C_{xx}(\kappa)$ of $\mathbf{x}$, where $\kappa$ denotes the $k$th estimation of first track resonance. The autocorrelation stays within the confidence bounds most of the time; this suggests that the signal $\mathbf{x}$ could be practically white. To further verify this hypothesis, a whiteness test [55]...
is applied. Given the autocorrelation function
\[ C_{xx}(\kappa) = N^{-1} \sum_{k=1}^{N-\kappa} x(k + \kappa)x(\kappa), \quad (\kappa \geq 0) \]
(16)
then the test statistic of the whiteness test is given by
\[ R(m, N) = \frac{N}{C_{xx}(0)} \sum_{i=1}^{m} C_{xx}^2(i). \]
(17)
Let \( T_a \) be the whiteness threshold selected to obtain the probability of false alarms \( \alpha \), i.e., \( \alpha = Pr\{R > T_a\} \). If \( R \leq T_a \), then \( x \) is a white and serially uncorrelated sequence over a range of \( m \) data points. Setting \( \alpha = 0.05 \), the whiteness test returns \( p \)-values equal to 0.95, 0.62, and 0.001 for A2, A11, and A7 locations, respectively. It can then be concluded that \( x \) is i.i.d. at locations A2 and A11, whereas it is not at location A7. Despite the i.i.d. assumption is verified in two out of three cases, the data record is not large enough for the test statistics \( 2L_G(x) \) to converge to its asymptotic limit. Hence, the determination of the threshold would be better addressed based on the empirical analysis of the output of the test statistics (14).

Although theoretically well-founded and empirically implementable, the aforementioned approach is challenged by the possibility of quantitatively identifying the “sufficiently stiff” ballast condition either on theoretical considerations or operational know-how. A more practicable approach could be the creation of an empirical mapping between the \( G \)-GLRT index and the track quality index already adopted by infrastructure managers in agreement with standard norms. The projection of the proposed ballast quality index into an existing one would also facilitate acceptance from industry as its meaning would be transparent. The identification of such mapping is not considered in this article, and it will be a matter of future research.

D. Comparative Analysis With Alternative Methods

There are several factors influencing the effectiveness and practicability of ballast monitoring methods within the industrial framework, thereby their adoption by railway infrastructure managers. The following performance metrics have been selected to carry out a comparison with other methods proposed in the literature, such as direct methods (DM) and high-fidelity model-based methods (HFMM) \([7, 9, 19]–[21]\), as well as with the current industrial practice based on the track geometry measurements (TGM) \([56]\):

- **Continuous Monitoring**: The method allows for continuous monitoring with variable time resolution.
- **No Traffic Disruption**: The method does not require the disruption of the normal train traffic for data acquisition and/or calculation of the quality index.
- **Local Monitoring**: The method enables the monitoring of a single turnout and/or individual sections of a single turnout.
- **Simplicity**: The method is easy to understand, and the output of the method is interpretable.
- **Plug & Play**: The method is portable across the railway network and can be readily used at new locations.

**Practicality**: The method takes into account practical constraints and implementation aspects and is thereby suitable for use in practice.

**Nondestructive**: The method does not require destructive testing for assessing the quality of the ballast.

**Stand-Alone Solution**: The method provides actionable information without the need for supplementary inputs provided by other systems/methods.

Table III shows the outcome of the comparative analysis. Noteworthy that the integration of the proposed method with the current industry practice may result in a superior turnout monitoring technology with variable time and spatial resolution and capable to produce actionable information aligned with the industry operational standards.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>DM</th>
<th>HFMM</th>
<th>TGM</th>
<th>( G )-GLRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous monitoring</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>No traffic disruption</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Local monitoring</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Simplicity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Plug &amp; play</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Practicality</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Non destructive</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Stand-alone solution</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

VIII. CONCLUSION

A condition-monitoring system for the assessment of ballast quality in railway turnouts was proposed in this article. The monitoring system consists of a frequency estimator and a change detection algorithm based on the GLRT specifically designed for GEV distributed random processes. The frequency estimator combines the EMD and numerical algorithms for subspace state-space system identification to provide a robust estimate of the track first resonance frequency exploiting track vertical accelerations induced by train passage. The estimated frequency is then utilized to create the statistical model at the core of the monitoring algorithm.

The ballast monitoring system was developed and verified using full-scale measurements collected over a period of almost two years in a turnout of the Danish railway infrastructure. The performed testing demonstrates the clear ability of the monitoring system to address the degradation of the ballast layer due to the accumulated loading of the track. The tradeoff between the time resolution and accuracy of the monitoring index was discussed in connection with the selection of the detection window. Based on results and the discussion, it is concluded that the ballast monitoring obtains timely condition information with sufficient spatial resolution to support precise and timely maintenance actions.

Future work could investigate methods for integration of the proposed system with the track/ballast quality index \( \sigma_H \) currently used by infrastructure managers. Furthermore, the monitoring system could be extended to include other track...
components, such as the rail pad and the undersleeper pad. The frequency estimator is already capable of estimating track resonance frequencies related to other components; therefore, the creation of new statistical models for such estimates might make component-specific diagnostics possible.

Finally, to pursue the network-wide applicability of the monitoring system, the monitoring algorithm could be altered to work based on accelerations measured by measurement setup installed onboard in-service trains. This would require additional signal processing to separate the train bogie dynamics from the track dynamics and to repeat the characterization of estimated resonance frequencies.

**ACKNOWLEDGMENT**

This research study has been carried out as part of the INTELLISWITCH project. They also very much appreciate the help from the Danish Meteorological Institute that made weather data freely available for this research.

**REFERENCES**


Mogens Blanke (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from the Technical University of Denmark (DTU), Lyngby, Denmark, in 1974 and 1982, respectively.

He was a Systems Analyst with the European Space Agency, Noordwijk, The Netherlands, from 1975 to 1976 and DTU from 1977 to 1984, the Head of Division at Lyngso Marine, Harsholm, Denmark, from 1985 to 1989, and a Professor with Aalborg University, Aalborg, Denmark, from 1990 to 1999. Since 2000, he has been a Professor of automation and control with DTU. He was an Adjunct Professor with the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, from 2005 to 2017. His areas of special focus are the diagnosis, prognosis, fault-tolerant control, and autonomous systems. His research leadership included the control system for the Ørsted satellite from 1992 to 1999 and the National Danish Effort in Surface Ship Autonomy since 2017.

Prof. Blanke received various international recognitions, recently, the ASME DSCD 2018 Rudolf Kalman Best Paper Award. He has served as a Technical Editor for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS from 2006 to 2016. He is also an Associate Editor of the Control Engineering Practice and the Deputy Editor of the Ocean Engineering.

Roberto Galeazzi (Member, IEEE) received the M.Sc. degree in electronic engineering from the Polytechnic University of Marche, Ancona, Italy, in 2005, and the Ph.D. degree in automation and control from the Technical University of Denmark (DTU), Lyngby, Denmark, in 2010.

He was a Post-Doctoral Fellow with the Center of Excellence for Ships and Ocean Structures, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, from 2009 to 2010 and DTU from 2010 to 2011. He is currently an Associate Professor with the Department of Electrical Engineering, DTU. His research interests include fault diagnosis and prognosis, fault-tolerant and reconfigurable control, nonlinear and adaptive control, and state estimation. His focus areas are safety-critical systems and autonomous systems. Dr. Galeazzi is also a member of the IEEE Technical Committee on Aerospace Control. He has been serving as the Chair of the IFAC Technical Committee on Marine Systems since 2018. He is also an Associate Editor of the Control Engineering Practice.

Pegah Barkhordari received the M.Sc. degree in electrical engineering from Iran University of Science and Technology (IUST), Tehran, Iran, in 2015, and the Ph.D. degree in automation and control from the Technical University of Denmark, Lyngby, Denmark, in 2019.

She is currently a Research and Development Scientist with MAN Energy Solutions, Copenhagen, Denmark. Her research interests include fault diagnosis and prognosis, fault-tolerant control, system identification, data mining, and signal processing.

Dr. Barkhordari received the Best Application Paper Award from the PHM Conference in 2017.