Delft University of Technology

Evaluating railway track stiffness using axle box accelerations

A digital twin approach

Shen, Chen; Zhang, Pan; Dollevoet, Rolf; Zoeteman, Arjen; Li, Zili

DOI
10.1016/j.ymssp.2023.110730

Publication date
2023

Document Version
Final published version

Published in
Mechanical Systems and Signal Processing

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Evaluating railway track stiffness using axle box accelerations: A digital twin approach

Chen Shen [a], Pan Zhang [a], Rolf Dollevoet [a], Arjen Zoeteman [b], Zili Li [a,*]

[a] *Delft University of Technology, Section of Railway Engineering, Stevinweg 1, 2628 CN, Delft, the Netherlands*
[b] *ProRail B.V., Utrecht, the Netherlands*

ARTICLE INFO

ABSTRACT

While various train-borne techniques have been developed for measuring railway track stiffness, differentiating stiffness at different track layers remains a challenge. This study proposes a digital twin framework for the vehicle–track interaction system, which enables track stiffness evaluations based on axle box accelerations (ABA). The digital twin consists of a physics-based model, a model library and data-driven models. Compared to existing techniques, the proposed method simultaneously evaluates the stiffness of the railpad, sleeper and ballast layers at a sleeper spacing resolution, while being robust to varying track conditions, such as track irregularities and vehicle speeds. This is accomplished by employing a localized frequency-domain ABA feature capable of distinguishing between the characteristics of different track layers. Furthermore, track stiffness is evaluated in near real-time. This is achieved using a model library derived from physics-based simulations of a range of track conditions. Two data-driven models that can quickly select or interpolate model instances contained in the library are developed. During operation, the data-driven models use the measured ABA features as input and then infer the stiffness for the different track layers. The proposed method is applied to evaluate the track stiffness of a downscale test rig in a case study. The track stiffness evaluated by the proposed method is compared with that obtained through hammer tests and with the observations of the track component conditions. These comparisons show that the proposed method can capture the stiffness variations due to periodically fastened clamps and substructure misalignments at different speeds. In addition, the proposed method is demonstrated to be superior to the commonly used hammer test method for evaluating track stiffness under loaded conditions.

1. Introduction

Railway track stiffness is an important track property that is closely related to track condition and maintenance. Track stiffness variations occur over time and space due to dynamic train loading and aging of track components, such as worn railpads, hanging sleepers, and ballast fouling. Track stiffness variations may further lead to geometric deterioration [1,2] and consequently, major maintenance costs. Therefore, it is essential to continuously monitor track stiffness variations and related track component degradations over time and space. This allows for preventive and targeted maintenance, which reduces the life-cycle cost of rail infrastructures.

---

* Corresponding author.
E-mail address: Z.Li@tudelft.nl (Z. Li).

The ballasted railway track is a multilayered structure that includes rails, fasteners, railpads, sleepers, ballast, and subgrades. Track stiffness is most commonly defined as the equivalent stiffness provided by all track components beneath the rail [3]. This is also known as the overall track stiffness or track modulus. However, it is desirable to differentiate the stiffnesses of the individual track layers because each layer has well-defined and distinct design/target stiffness values. This study aims to evaluate the stiffness of the individual track layers separately.

Track stiffness can be measured under loaded or unloaded conditions. Under unloaded conditions, external excitation forces are provided by impact hammers [4–8] or falling weights [9,10]. Another approach is to measure the track stiffness under a dynamic train load. Measurement techniques include pass-by [3,11,12] and train-born measurements [13–21]. Owing to the cost of sensor deployment, pass-by measurements are more suitable for discrete locations of special interest in rail networks. In comparison, train-borne measurements using in-service vehicles are more cost-effective and allow for more frequent surveys of the entire rail network. Among the various train-borne measurement techniques, axle box accelerations (ABA) show good potential for condition monitoring of railway tracks. For instance, ABAs have been used for detecting rail surface defects [22 23] and measuring track irregularities [24]. In this study, we further explored the potential of ABA in evaluating track stiffness.

The basis for measuring track stiffness using the ABA is to view the vehicle and track as an integrated system. Because of the vehicle–track interaction (VTI), the ABA is influenced by the underlying track properties, including track stiffness. Particularly, each layer of components in the vehicle–track system serves a specific function and possesses unique stiffness ranges, resulting in distinct vibration characteristics. These distinctive vibration characteristics can be effectively captured using ABA. Therefore, the ABA measured from a moving vehicle can be used to evaluate the stiffness of the track over which the vehicle travels.

A major challenge is to identify the features in the dynamic responses that can be correlated with the track stiffness at different layers. Several frequency-domain features have been proposed to assess the overall track stiffness [3,21,25] or local stiffness loss [19,26]. Alternatively, measured time-domain signals can be used directly as features [16,20]. However, these features cannot distinguish between track stiffnesses in different layers. Multiple features that can be distinctly correlated with different layers must be defined for multilayer track stiffness evaluations [7,14]. In our previous study [7], it was demonstrated that the track stiffness at different layers is correlated with different frequency response function (FRF) features. This opens the possibility of using ABA measurements for the same purpose.

Another issue with train-borne methods is the difficulty in obtaining measurement resolution at the sleeper spacing level. For example, Yang et al. [21] evaluated track stiffness variations at a resolution of 20 m. Quirke et al. [16] proposed a method for evaluating track stiffness variations with different sleeper supports. However, the stiffness values of each sleeper were not evaluated individually. Instead, a track stiffness profile template is used to represent the stiffness variation along the track length, thereby reducing the total number of stiffness variables. Zhu et al. [20] considered a track section with approximately 80 sleepers. Track stiffness can be evaluated at the sleeper spacing resolution when stiffness variations exist only in a single track layer.

Furthermore, a rapid track stiffness evaluation is required for train-borne measurement techniques. This is a large amount of data must be processed when the ABA is regularly and frequently measured using in-service vehicles. Evaluating track stiffness from measurements is essentially an inverse problem that involves fitting a VTI model to measurements. Two approaches can be used for this task. The first approach involves solving an optimization problem by iteratively minimizing the objective functions that define the difference between the modelled and measured dynamic responses [16,20]. However, conventional optimization-based iterative methods are time-consuming. For example, in [16], the computation time for evaluating the stiffness of a 140-m long track was 11 h. Alternatively, direct mapping between the measured dynamic responses and track stiffness can be established through simple regression [3] or more advanced data-driven methods [7,14].

Data-driven approaches rely on large quantities of labelled data for training, and such data are scarce for vibration measurements in railway applications. An accurate physics-based model is required to generate the training data. Therefore, a framework is required to integrate physics-based and data-driven models and enable frequent communication between the measured data and physics-based models. Emerging digital twin technology offers a potential solution to this problem. Although several conceptual physics-based digital twin frameworks have been proposed [27–29], they have not been used in rail applications.

Finally, there appears to be a lack of experimental and field validation of the train-borne track stiffness evaluation method. The trackside techniques in Refs. [3,11,12] were validated using field measurement data, whereas train-borne measurement techniques were demonstrated solely through numerical examples [16,18,20,21] or simple lab setups [19].

This study proposed a digital twin framework that enables multi-layer, high-resolution and near real-time track stiffness evaluations using ABA. The remainder of this paper is organized as follows. Section 2 describes the methodology of the proposed digital twin framework. Section 3 validates the individual components in the digital twin. Section 4 demonstrates the capability of the digital twin through numerical examples. Section 5 presents a case study on a test-rig in the lab. Section 6 discusses the possible improvement of the proposed method and Section 7 concludes this paper.

2. Methodology

2.1. Problem statement

The observed ABA data for *a track section* can be represented using a feature vector, denoted as $x^* \in X$, where X is the feature space (i.e., the space of all possible ABA features). The stiffness values of the ballast and railpad for the same track section are denoted by the vector $k^* = [k_b, k_p] \in K$, where K is the space of all possible stiffness values. The task is to use the observed ABA data x^* to infer the
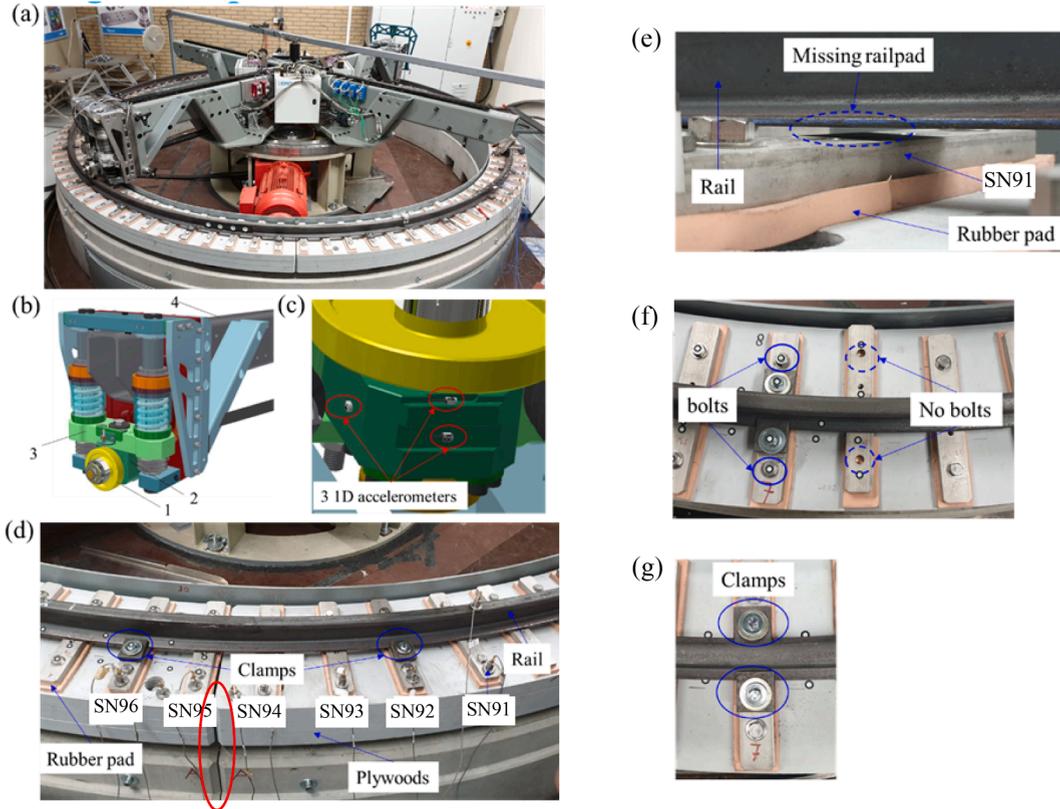
Fig. 1. V-Track and experimental setup for ABA measurement. (a) V-Track test rig; (b) structure of wheel assembly: 1) wheel, 2) axle box, 3) guiding block, 4) steel frame; (c) accelerometers for ABA measurement; (d) track structure overview for a section of V-track (the red oval indicates the gap between SN94 and SN95); (e) missing railpad at SN91;(f) missing sleeper bolts at SN94;(g) top view of fastening clamps. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stiffness values of track components k^*. This task can be framed as an inverse problem. The objective is to find an inverse mapping from feature space X to stiffness space K, i.e.,

$$G : X \rightarrow K. \tag{1}$$

2.2. Physical asset: V-Track

In this study, the proposed method is demonstrated and validated by measurements on a downscaled VTI test rig called V-Track. In this section, the V-track setup for ABA measurements is briefly introduced.

The V-Track consists of a maximum of four-wheel assemblies running over a ring-track system. In the configuration used in this study, a one-wheel assembly was used, as shown in Fig. 1(a). Fig. 1(b) shows the structure of the wheel assembly. A wheel (1) with a diameter of 130 mm was mounted on the guiding block (3) through the axle box (2). The wheel assembly was mounted on the arm of the steel frame (4) and vertically loaded using two springs. The combined stiffness of the two springs is 230 N/mm. A motor drives the steel frame, such that the wheel assemblies are pulled along the ring track. The speed range of the wheel assemblies is between 0 and 40 km/h. A more detailed description can be found in [30].

The ring track, with a radius of 2 m, consisted of four rails connected by four joints. The ring rail, which had a standard S7 profile, was supported on the railpads. The rail and railpads were fixed to the steel sleepers using fasteners, as shown in Fig. 1(d). Notably, clamps were used only every three or four sleepers to vary the track stiffness for the tests. A total of 100 steel sleepers were used, and the sleeper spacing was approximately 0.125 m. All sleepers were numbered from 1 to 100 for positioning (for instance, SN1 was designated as sleeper no. 1). Underneath the sleepers were rubber pads that were used to simulate the elasticity and damping of the ballast layer. The subgrade is simulated using plywood layers.

The track stiffness varied along the V-track in different track layers. For example, a missing railpad was set at SN91 (Fig. 1(e)), missing sleepers were set at SN94 (Fig. 1(f)), and fastening clamps (Fig. 1(d)) were used at SN92 and SN96. Gaps between plywood were also present (e.g., between SN94 and SN95 in Fig. 1(d), which is indicated by the red oval).

Three one-dimensional (1D) accelerometers were placed on the wheel-axle box to measure the ABA in three directions, as shown in Fig. 1(c). Vertical ABA was used to assess track stiffness. The sampling frequency of the ABA signals was 100 kHz. When recording the
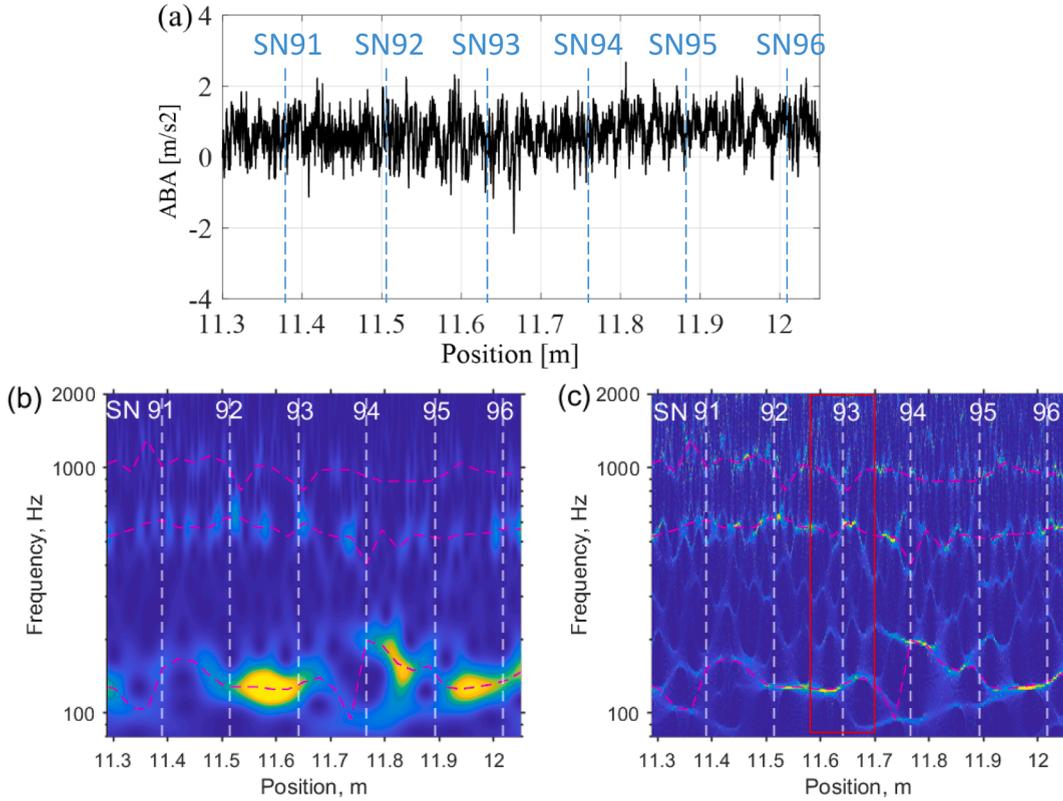
Fig. 2. WPS of the (a) ABA signal with (b) CWT; (c) SSWT. The dashed pink lines indicate major frequency components. The red rectangle represents a vertical sliced window that is used to calculate the LWPS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ABA, the running speeds of the wheel assembly on the ring track were 13, 10, and 7 km/h, which are equivalent to vehicle speeds of 65, 50, and 35 km/h on real tracks [30], respectively. The wheel preload was 4500 N.

2.3. ABA feature: Local wavelet power spectrum (LWPS)

To achieve the goal of evaluating track stiffness in multiple layers and sleeper spacing resolution, we propose a local frequency-domain representation of ABA (i.e., a local wavelet power spectrum (LWPS) based on the synchro-squeezed wavelet transform (SSWT)) [31].

The SSWT is a time–frequency (TF) analysis method that can identify oscillatory components with time-varying frequencies and amplitudes. Compared to other TF transforms, such as continuous wavelet transform (CWT) and short-time Fourier transform (STFT), SSWT can compensate for the spreading in time and frequency by applying a post-processing "reassignment" map. This "reassignment" map focuses the spectrogram's energy on the instantaneous frequency curves and results in a sharpened TF representation. Detailed algorithms for calculating CWT and SSWT can be found in [31,32] and are described in Appendix A.

Fig. 2(a) shows a section of the ABA signal in the spatial domain measured from the V-track (track section SN91–SN96 in Fig. 1(d)). The horizontal axis represents the wheel position x_w relative to the first sleeper (SN1), which can be calculated using time b as follows:

$$x_w = v \times (b - b_0),   \tag{2}$$

where b_0 is the time at which the wheel passes the first sleeper (SN1). The sleeper positions are denoted by vertical dashed lines in Fig. 2(a).

Fig. 2 (b) and (c) show the wavelet power spectrums (WPS) of the CWT and SSWT of the ABA signal, which were calculated as the square magnitudes of the wavelet coefficients $|W_s^2(a,b)|$ (see Eq. (A.1)) and $|T_s^2(\omega,b)|$ (Eq. (A.4)). In the two plots, the scale a and angular frequency ω are converted to the frequency f in hertz as

$$f = \begin{cases} f_c/a \\ \omega/2\pi, \end{cases}   \tag{3}$$

where f_c is the center frequency of the mother wavelet ψ in Eq. (A.1).
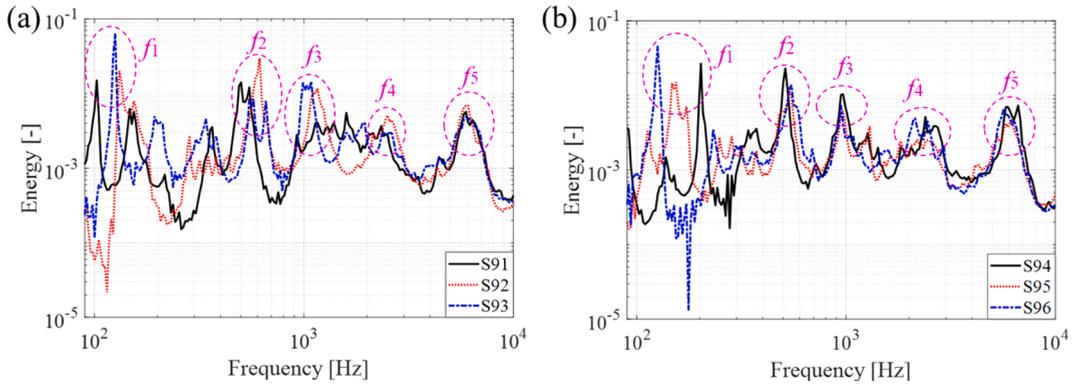
Fig. 3. Examples of LWPS of ABA at different sleepers: (a) SN91-SN93; (b) SN94-SN96. The dashed pink circles indicated five major characteristic frequencies f1–f5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
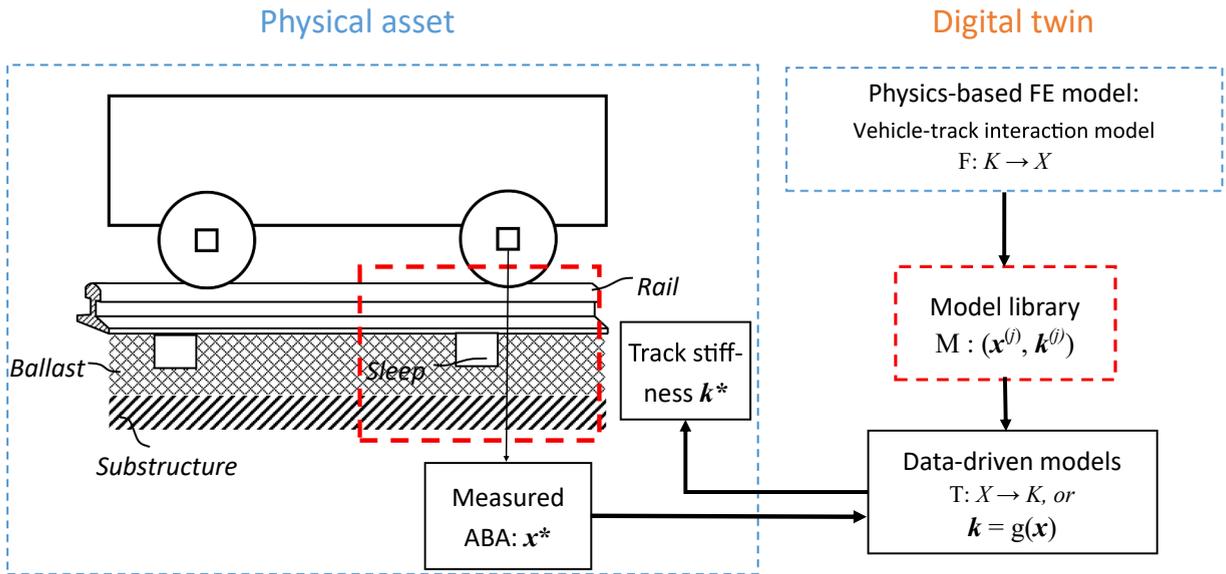
Fig. 4. Proposed digital twin framework for evaluating track stiffness.

This ABA signal has three major frequency components, i.e., 100–200, 500–620, and 950–1200 Hz, indicated by the dashed pink lines. Compared with the more smearing and distortion in the CWT spectrum, the time-varying frequency components were more clearly distinguished in the SSWT spectrum.

To obtain the frequency components of ABA at a local position d, a window was taken through the WPS, as indicated by the red rectangle in Fig. 2 (c). The window length is $\Delta d = d_2 - d_1$, where d_1 and d_2 are the starting and ending positions of the window, respectively. In this study, this window was centered at the sleeper positions with a length of one sleeper span ($\Delta d = 0.125$m). Subsequently, the LWPS was calculated at position d as the wavelet power averaged over the window length, which is expressed as

$$\overline{T}_d^2(f) = \frac{1}{n_{\Delta d}} \sum_{x_w=d_1}^{d_2} \left| T_s^2(f, x_w) \right|,$$ (4)

where $n_{\Delta d}$ is the number of position points within the window length Δd, and $|T_s^2(f, x_w)|$ is the wavelet power obtained by the SSWT in the spatial frequency domain, which can be calculated based on Eqs. (A.4), (2), and (3).

Fig. 3 shows the LWPS measured at six sleeper positions (SN91–SN96) in the frequency range of 100–10000 Hz. Five major frequency components were observed, as indicated by the dashed pink circles. Clear LWPS variations between different sleepers were also observed, which can be used to evaluate the spatial track stiffness variations (see Section 3.2).
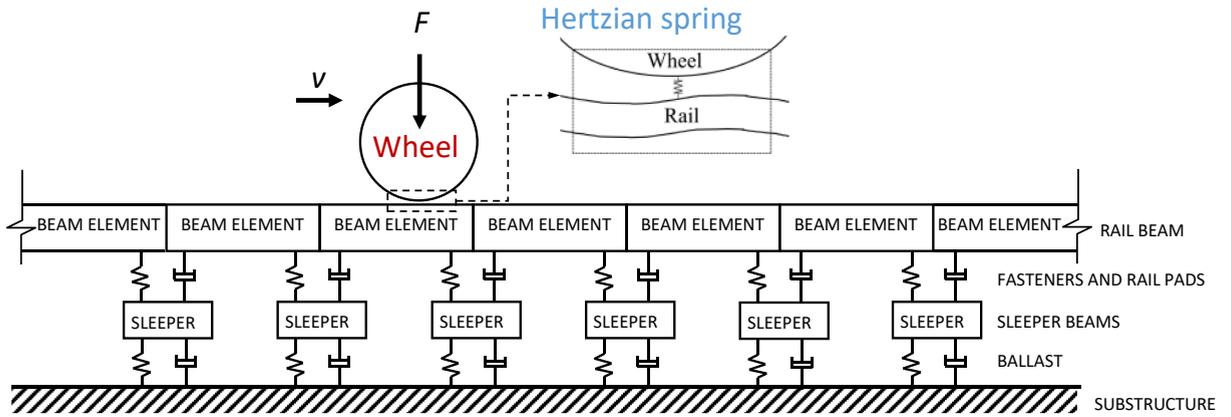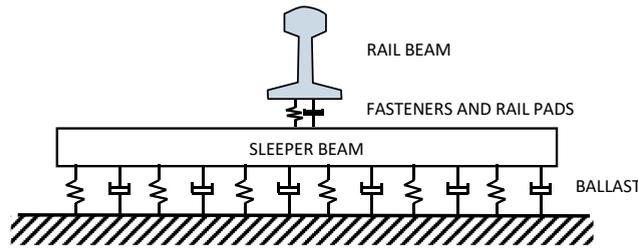
Fig. 5. FE vehicle–track interaction model.

2.4. Digital twin framework

To address the problem defined in Eq. (1), a digital twin framework was proposed consisting of three major components (i.e., a physics-based FE model, model library, and data-driven models associated with the library), as illustrated in Fig. 4.

2.4.1. Physics-based FE model

To model the V-track system, a two-layer discretely supported VTI model was developed using the finite element method [7,33], as shown in Fig. 5.

The rail and sleepers were meshed using Timoshenko beam elements [34]. This element has two nodes and four DOFs. At each node, only the vertical and in-plane rotational degrees-of-freedom were considered. The ballast and railpads were modelled as discrete spring-damper pairs. Note that other components of the fastening system were neglected, such as clamps and bolts. This is a widely accepted simplification of railway track models. Although clamps and bolts were not explicitly considered in the current model, their stiffness was reflected in the railpad stiffness. For example, a loose bolt reduces the clamping force, further reducing the stiffness of the railpad. The track length was 3.14 m, which is a quarter of the length of the test ring, and consisted of 25 sleeper spans.

The wheel was simplified as a rigid mass. The load from the primary suspension spring is applied as a vertical load on the wheel. The wheel-rail contact is modelled as a Hertzian spring. The equations of motion and solution techniques are described in detail in Appendix B and Appendix C.

In this study, we focus on examining the dynamic vehicle-track interaction (VTI) on straight tracks. Therefore, a straight track was modelled in the FE model. Regarding the V-track setup, despite a curved track being constructed, the dynamic VTI on a straight track was simulated by setting the angle of attack to be as small as possible [30,35]. The accuracy of the FE model was validated in Section 3.1 through comparisons with measured FRF and ABA on the V-track. This ensures that both the FE model and V-track system simulate the VTI on straight tracks.

The FE model establishes a forward mapping from the track parameter space P to the feature space X

$$F : P\left(k_{\mathrm{b}}, k_{\mathrm{p}}, c_{\mathrm{b}}, c_{\mathrm{p}}, EI_{\mathrm{s}}, A_{\mathrm{irr}}, \ldots\right), v \rightarrow X. \tag{5}$$

Note that the stiffness space K in Eq. (1) is a subspace of track parameter space P. This is because other track parameters may also influence the ABA features. In this study, a parameter space P that includes six uncertain track parameters was considered: ballast stiffness k_{b}, railpad stiffness k_{p}, ballast damping c_{b}, railpad damping c_{p}, sleeper bending stiffness EI_{s}, and track irregularity level A_{irr}. In
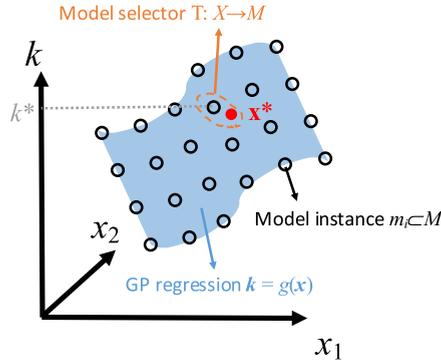
Fig. 6. Data-driven models for fast track stiffness evaluations: model library M, model selector T and GP regression model g demonstrated in a two-dimensional feature space $X = [x_1, x_2]$.

addition, vehicle speed v was considered to be an independent variable because it may also influence the feature space but is relatively convenient to measure with certainty. The other parameters of the FE model were kept constant. A detailed definition of the parameter space P is described in Appendix D.

2.4.2. Model library

Simulations can then be run with different input track parameters sampled from the parameter space P, representing different track conditions. Each simulation case using the FE model represents a possible state of the physical asset, called a model instance, denoted as $m_j \in M$, where M is a model library that includes all simulated model instances. Each model instance m_j is associated with a data pair $\left(x^{(j)}, k^{(j)}\right), x^{(j)} \in X, k^{(j)} \in K$, as shown in Fig. 4.

The model library was used to represent the states of a local track section, as illustrated by the red rectangular in Fig. 4. The track section covers the length of the sleeper spacing centered above a sleeper. A major advantage of such a local representation is that the dimensions of both the feature space X and track parameter space K are reduced. As shown in Section 2.3, the feature space is reduced by using a local frequency-domain representation of the ABA. The parameter space contains only the stiffness values of the railpad, sleeper, and ballast over a single span.

However, by using a local model library representation, an assumption was made that the ABA measured over a sleeper support is only related to the local track stiffness at that support. In reality, this assumption holds true only when the track stiffness is uniform across multiple sleeper spans. When stiffness variations exist, the evaluations by the local model library usually indicate an averaged stiffness across adjacent sleeper spans. Nonetheless, such an assumption is used in many applications. It allows for more efficient analysis of the large volumes of data required to understand variation over a long length of track. This effect is investigated in Section 4.3.2.

2.4.3. Data-driven models

With the model library, the goal is to perform fast evaluations of track stiffness k^*, given an observed ABA feature vector x^*, as shown in Fig. 4. To accomplish this, two data-driven methods were proposed, i.e., a model selector and a Gaussian process (GP) regression model. Fig. 6 illustrates the principle of the two methods in a two-dimensional feature space $X = [x_1, x_2]$.

2.4.3.1. Model selector. The model selector searches directly within the model library M for one or several model instances $m_j \sim \left(x^{(j)}, k^{(j)}\right)$ that best match observation x^*. The model selector is denoted by

$$T : X \rightarrow M. \tag{6}$$

The basic principle is to search for the nearest neighbours of x^* in $x^{(j)}$ for $j = 1, 2, ..., N$, as illustrated by the dashed oval in Fig. 6. The algorithm is outlined as follows.

1 The Euclidean distance is calculated between x^* and $x^{(j)}$

$$d = \{d_j\} = \left\| x^* - x^{(j)} \right\|_2, j = 1, 2, ...N. \tag{7}$$

2. The Euclidean distance d is sorted in ascending order to obtain the sorted vector $p = \{p_j\}$.
3. A ranking vector r_j is returned for each d_j such that $d(r_j) = p_j, j = 1, 2, ..., N$.
4. The top $n < N$ matches of x^* in $m_j \sim \left(x^{(j)}, k^{(j)}\right)$ is returned, denoted $m_i^* \sim \left(x^{*(i)}, k^{*(i)}\right)$, where $m_i^* = m_{r_i}, i = 1, 2, ..., n$.
5. The prediction for k^* is calculated by taking the average of the top n matches
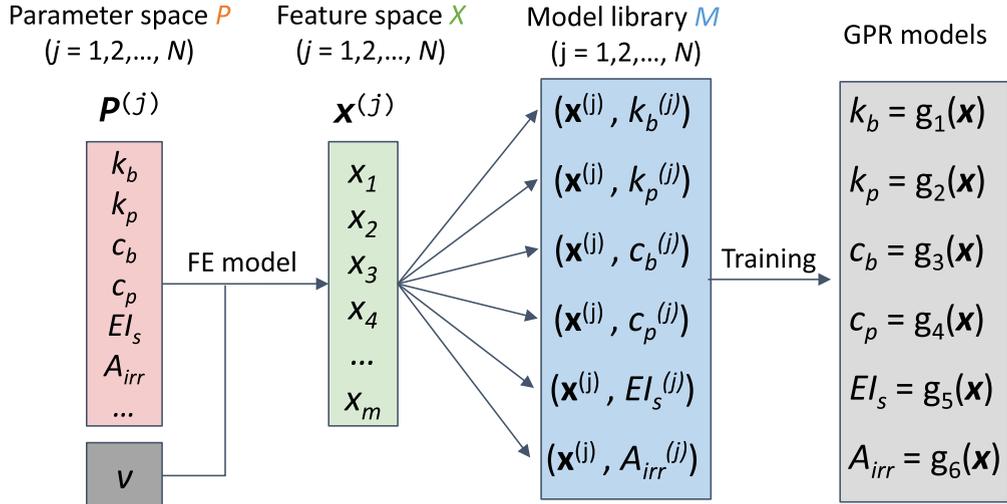
Fig. 7. Workflow of the digital twin framework.

$$\overline{\boldsymbol{k}}^* = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{k}^{*(i)}. \tag{8}$$

The model selector has the advantage of being directly applicable to a model library without the need to train a data model. It can be used independently as a track stiffness evaluation method. Alternatively, it can serve as an efficient first step in evaluating the quality of the model library before being used for building data-driven models. The capabilities of the model selector are demonstrated in Sections 3.1.2 and 3.3.

2.4.3.2. Gaussian process (GP) regression. Because feature space *X* is a potentially high-dimensional space, the feature vectors $\boldsymbol{x}^{(j)}$ in the model library are likely to be sparsely distributed. Consequently, it is possible that the target ABA feature is not sufficiently close to any of the features in the model library. In this case, the GP regression method was used to interpolate between model instances, as shown by the blue plane in Fig. 6. The GP regression theory can be found in [36]. A detailed implementation of a similar task was described in our previous work [7]. The training of the GP models is described in Section 2.5.5.

2.5. Digital twin workflow

2.5.1. Sample from the parameter space
The workflow of the digital twin framework begins by defining parameter space *P*, as shown in Fig. 7. The value ranges for the six track parameters were determined based on the nominal or design values of the track components. In addition, parameter values that represent potential component degradations were also considered.

Subsequently, the Sobol sequence sampling technique [37] was used to obtain $N = 10,000$ sets of space-filling input parameters $\boldsymbol{p}^{(j)}$ for the VTI model. The sampling was performed over the entire parameter space *P* rather than just one parameter. This implies that each sample $\boldsymbol{p}^{(j)}$ contains simultaneous variations in different track parameters, representing different fault conditions in different track components. A detailed definition of the parameter space *P* and sampling techniques are described in Appendix D.

2.5.2. Run simulations
N simulations were run with different input parameters $\boldsymbol{p}^{(j)}$ for *j = 1, 2, …, N*. The LWPS were calculated for each simulation case according to Eq. (4).

2.5.3. Extract feature vectors
The LWPS in different frequency ranges were used as the feature vectors. They can be calculated as
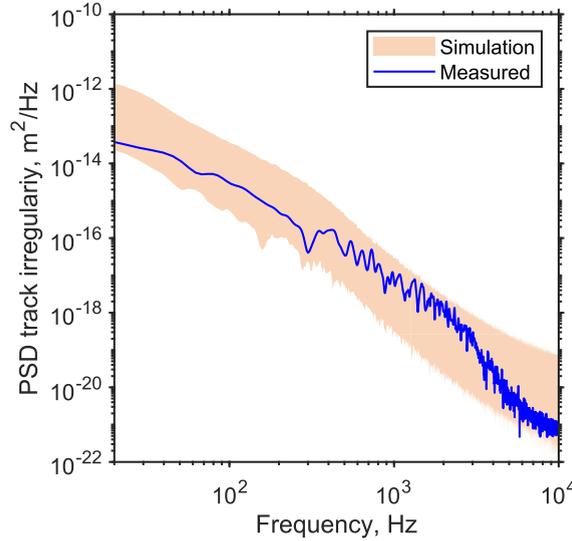
Fig. 8. Comparison of measured and simulated track irregularities. The frequency is calculated based on a speed of 13 km/h.

$$x = \left\{ w_1 \log\left(\overline{T}_d^2(\omega_1)\right), w_2 \log\left(\overline{T}_d^2(\omega_2)\right), ..., w_N \log\left(\overline{T}_d^2(\omega_n)\right) \right\}^T$$

$$\left. \begin{array}{l} x_{\mathrm{GF}} : for\ \omega_i \in [50, 4000] \\ x_{\mathrm{LF1}} : for\ \omega_i \in [50, 200] \\ x_{\mathrm{LF2}} : for\ \omega_i \in [400, 700] \\ x_{\mathrm{LF3}} : for\ \omega_i \in [800, 1500] \\ x_{\mathrm{LF4}} : for\ \omega_i \in [1500, 4000] \end{array} \right\} w_i = 1, otherwise, w_i = 0 \tag{9}$$

where ω_i is the frequency point in Hz, $\overline{T}_d^2(\omega_i)$ is the LWPS at frequency ω_i and position d (Eq. (4)), and w_i is the binary weighting factor for $\overline{T}_d^2(\omega_i)$. A global feature (GF) vector x_{GF} was defined as the log magnitude of the LWPS between 50 and 4000 Hz. The LWPS above 4000 Hz was excluded due to the inadequate agreement between the measured and simulated LWPS in that frequency range (Fig. 10). Furthermore, four local feature (LF) vectors ($x_{\mathrm{LF1}} - x_{\mathrm{LF4}}$) were defined in four local frequency ranges. They were approximately centered on the first four characteristic frequencies of the LWPS (Fig. 3). This is because the stiffness of a particular track layer is affected only by certain local characteristics (Fig. 12). The inclusion of irrelevant features can degrade the performance of the data-driven models.

2.5.4. Build model library

Based on the extracted feature vectors from simulations, discrete data pairs were generated representing the relationships between the ABA features and each track parameter, such as $\left(x^{(j)}, k_{\mathrm{b}}^{(j)}\right)$ and $\left(x^{(j)}, k_{\mathrm{p}}^{(j)}\right)$, as shown in Fig. 7. These data pairs are essentially discrete forms of the mapping in Eq. (1), which were used to build a model library.

2.5.5. Train GP models

For each combination of a feature vector $x \in \{x_{\mathrm{GF}}, x_{\mathrm{LF1}}, x_{\mathrm{LF2}}, x_{\mathrm{LF3}}, x_{\mathrm{LF4}}\}$ and a track parameter $p \in \{k_{\mathrm{b}}, k_{\mathrm{p}}, c_{\mathrm{b}}, c_{\mathrm{p}}, EI_{\mathrm{s}}, A_{\mathrm{irr}}\}$, a data set was created, denoted by $D = \{x^{(j)}, p^{(j)}, j = 1, 2, ..., N\}$. The data set D with size $N = 10,000$ was randomly divided into two sets, each with 5000 samples. One was used as the training set and the other for testing. Five-fold validation was used during the training process to prevent overfitting. The performance of the GP models on the test set is evaluated in Section 3.4.

Six track parameters were included to account for multiple fault conditions in different components. A separate GP model was trained for each parameter. This enabled the simultaneous evaluation of the track parameters or multiple fault conditions. As shown in Fig. 7, the same feature vectors, $x^{(j)}$, were used to train the different GP models. Consequently, the effects of other track parameters on a GP model are not explicitly excluded but are inherently considered as uncertainties. This ensures that a trained GP model (e.g., for railpad stiffness) is robust to changes in other track parameters, such as ballast stiffness and track irregularities. The robustness and accuracy of the GP models are evaluated through numerical examples in Section 4 and a case study in Section 5.

3. Validation of the digital twin

In this section, each component in the digital twin framework was validated separately.
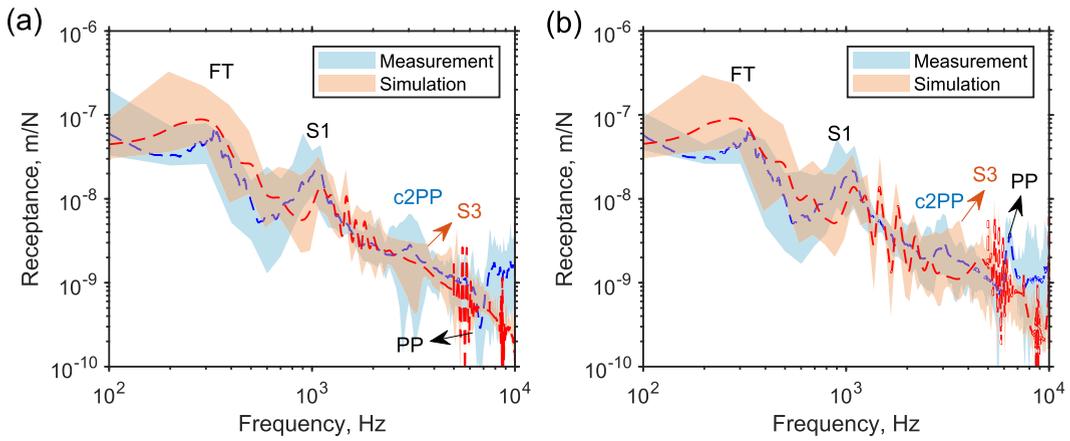
Fig. 9. Validation of the FE model based on FRF. (a) and (b) show the rail's point receptances above sleeper support and at mid-span, respectively.
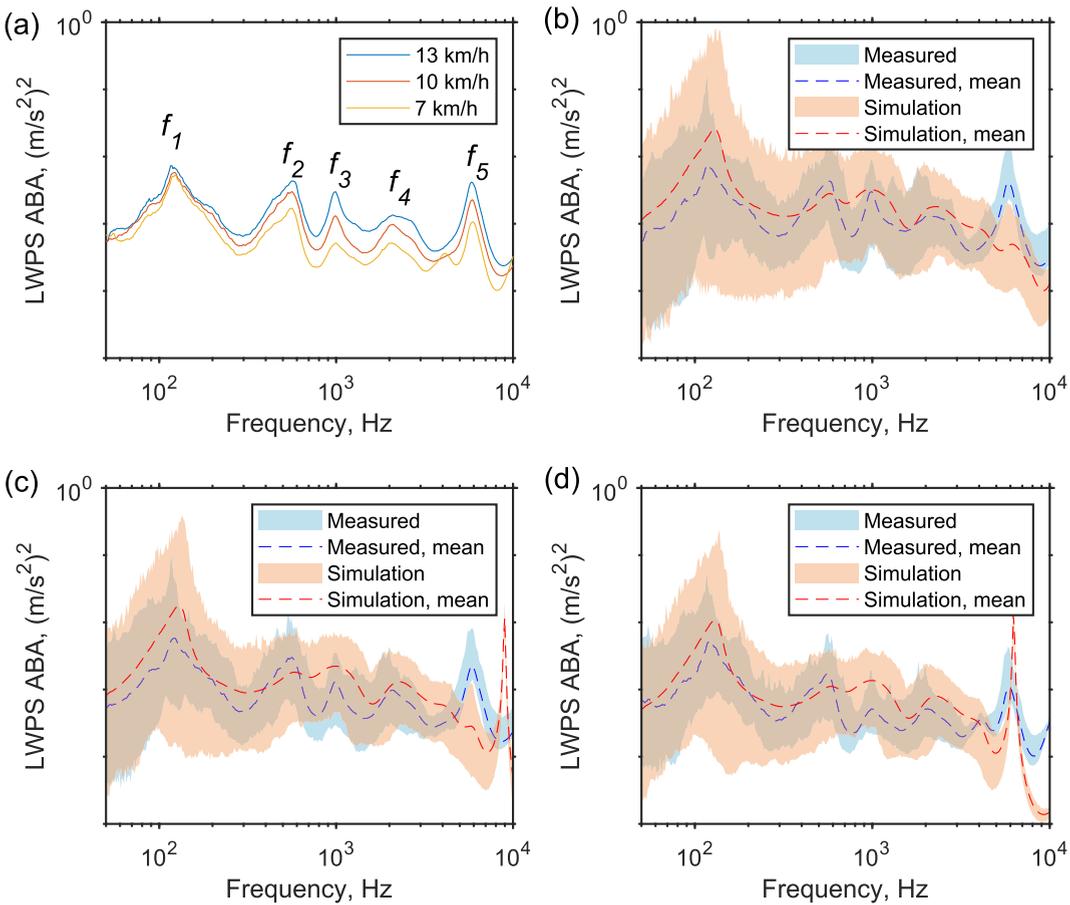
Fig. 10. Validation of the FE model based on ABA measurement at different vehicle speeds. (a) Measured LWPS at different speeds. Each LWPS curve represents the average results at the 100 sleepers in the V-track. (b), (c) and (d) compare the LWPS from measurements and simulations at speeds of 13 km/h, 10 km/h, and 7 km/h, respectively.

3.1. VTI model

3.1.1. Model library

The simulation cases in the model library were compared with the measurements. The objectives are to ensure that the parameter space and value ranges are well defined and that the numerical model adequately captures the primary dynamic characteristics of the
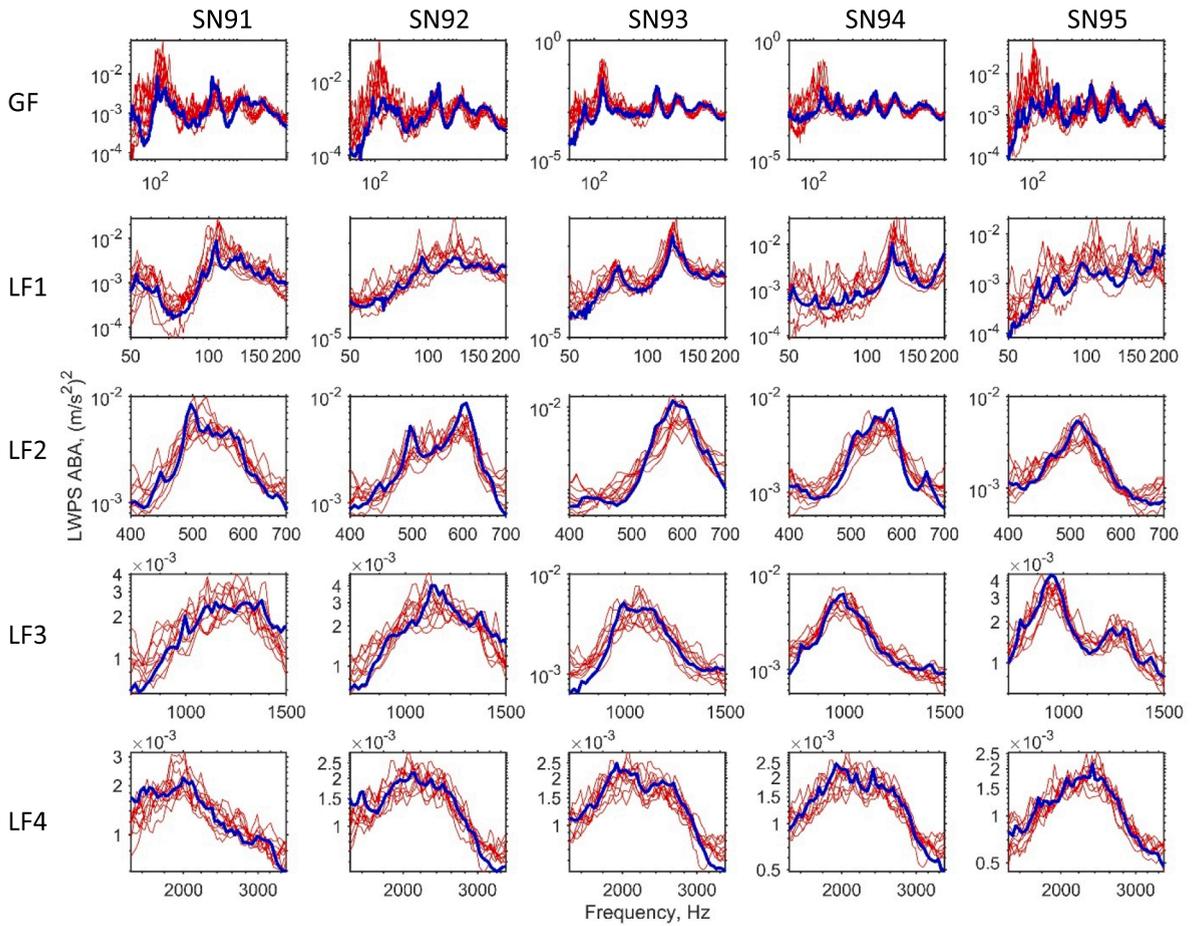
Fig. 11. Examples of the measured ABA LWPS (blue lines) compared to their top ten matches selected from the model library (red lines). Each row shows the models selected based on a feature vector. Each column shows the LWPS for a sleeper. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

wheel-track system.

Fig. 8 compares the simulated and measured PSD of the track irregularity. The measured PSD was based on the track irregularity measured on the V-track between SN91 and SN95. The simulated PSDs did not perfectly match the measured values. The simulated PSD tended to overestimate the track irregularity below approximately 300 Hz while underestimating it between 300 and 3000 Hz. However, the simulated PSDs cover the measured PSD in the frequency range of $10–10^4$ Hz. This indicated that the value range of the track irregularity level A_{irr} was correctly determined.

The simulated and measured FRFs of the V-track are compared in Fig. 9. Hammer tests were used to obtain the FRFs at 46 locations along the V-track (23 at the midspan and 23 above the sleepers). Hammer tests were simulated at the center of the tack model to eliminate the effect of boundary conditions. In general, the simulated and measured FRFs agree well.

The FRF peaks correspond to track resonances [7]. The measured FRFs exhibited four peaks at approximately 300, 1000, 3000, and 6000 Hz. Modal analysis of both the rail and sleepers was performed to determine the mode shapes at these frequencies. The four frequencies correspond to the full track (FT), sleeper 1st bending (S1), clamp 2nd pinned–pinned (c2PP), and pin-pin (PP) resonances. The c2PP resonance is caused by the special constraint of the fastening clamps in the V-Track, which are used every three or four sleepers to vary the track stiffness, as shown in Fig. 1. Consequently, this is unique to the V-track system.

The simulated FRFs can reproduce the FT and S1 resonances, but not the c2PP resonance. This is because the FE model does not account for the stiffness variations along the track. However, the simulated FRFs exhibited a peak representing the sleeper 3rd bending mode (S3) near the c2PP resonance.

The PP resonance cannot be well reproduced by the simulated FRFs. This could be because the Timoshenko beam element is incapable of modelling the S7 rail profile adequately at high frequencies.

Fig. 10 compares the measured and simulated LWPS of ABA at 13, 10, and 7 km/h. As shown in Fig. 10(a), the overall magnitude of the LWPS decreased as the vehicle speed decreased, whereas the characteristic frequencies remained constant. In Fig. 10 (b)–(d), the simulations reproduce the first four characteristic frequencies (f_1–f_4) relatively well compared with the measurements. The simulated
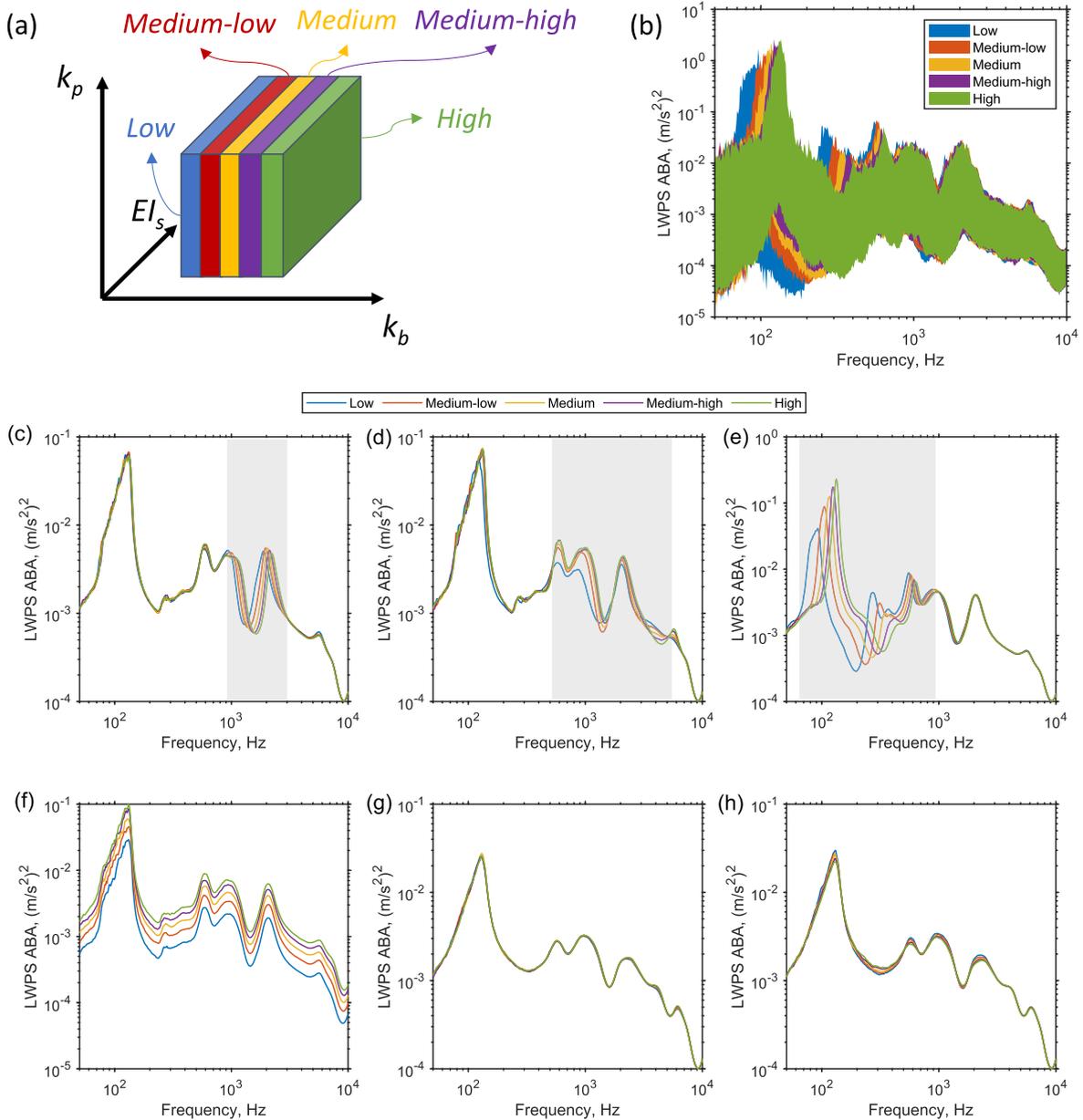
Fig. 12. Global sensitivity of LWPS to the change of track parameters. (a) The value range of ballast stiffness is divided into five sub-ranges. (b) Sensitivity of LWPS to ballast stiffness. (c)–(h) show the sensitivity of LWPS to different parameters with each curve representing the average LWPS within each value group. (c) Sleeper bending stiffness; (d) railpad stiffness; (e) ballast stiffness; (f) track irregularity level; (g) railpad damping; (h) ballast damping.

magnitude of the LWPS differed from the measurements at f_5. This can be attributed to the inaccurate high-frequency performance of the Timoshenko beam element or simplified fastening system models.

Based on the results shown in Fig. 9 and Fig. 10, it can be concluded that the parameter space and value ranges for the VTI model are well-defined. This is supported by the overlap envelops of the simulated and measured results. In addition, the VTI model adequately captures the primary dynamic characteristics of the wheel-track system. This is evident from the fact that the general trend in magnitude and major resonances can be captured by the VTI model.

Major differences were observed above 4000 Hz in both the FRF (Fig. 9) and LWPS (Fig. 10). One possible explanation could be that the Timoshenko beam element used in the simulation may not adequately model the S7 rail profile at high frequencies. However, these differences do not affect the evaluation of track stiffness, as LWPS above 4000 Hz were not used as ABA features.
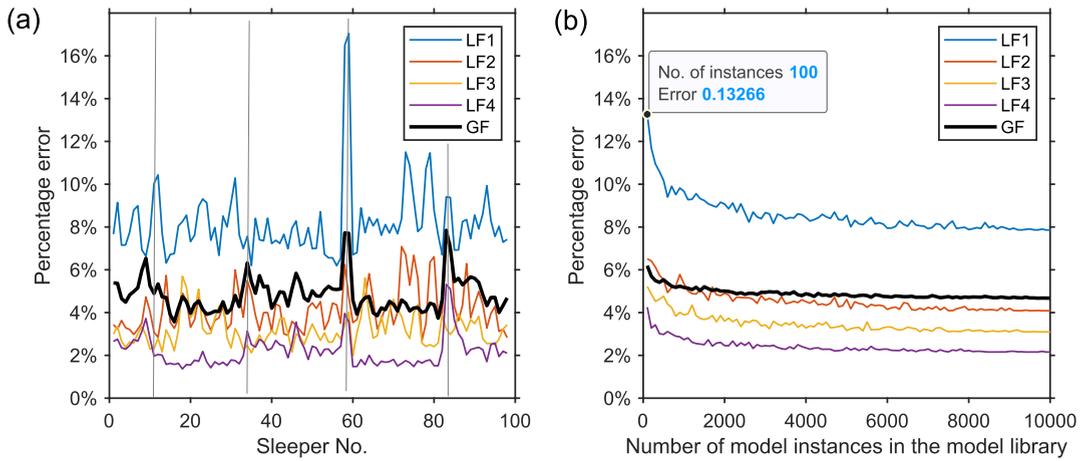
Fig. 13. Performance of the model selection algorithm in matching the measured ABA features. Vertical grey lines indicate the location of rail joints. (a) Averaged MAPEs of the top ten matches at each sleeper. (b) Relationships between the MAPE and the number of model instances. The MAPEs are calculated by averaging the MAPEs in (a) over all the sleepers with different numbers of model instances in the model library.

3.1.2. Case-by-case comparison

In the previous section, we compared multiple cases using envelopes. In this section, we will be comparing them on a case-by-case basis. Fig. 11 shows a comparison of the top ten matches selected from the model library and the measured LWPS using the model selector (Section 2.4.3.1). Each row corresponds to a feature vector defined in Eq. (9), where each column corresponds to a sleeper (SN91–SN95). Although the measured LWPS varied between sleepers, the simulations could consistently match the measured LWPS. When the GF was used (first row), the largest discrepancies were observed below approximately 400 Hz, and the LWPS selected from the model library exhibited larger magnitudes than the measurement. This could be because the track irregularities used in the simulations had a greater magnitude than the measured irregularities below approximately 400 Hz (Fig. 8). Such discrepancies were reduced by selecting models using LF1 rather than GF. However, a better local match did not guarantee a good match for the remaining frequencies of the LWPS.

3.2. Feature definition

To robustly infer the stiffness in each track layer, there should be unique ABA features that are sensitive to each stiffness. These features should also be insensitive to other track parameters and/or modelling errors. In this section, the feature definition was validated through a global sensitivity analysis.

The analysis was based on the 10,000 simulation cases. First, the value range of each track parameter was divided into five sub-ranges. Subsequently, for each parameter, 10,000 sets of input parameters were divided into five groups based on the five subranges. Fig. 12 (a) shows an example of dividing the input parameters into five groups based on ballast stiffness (k_b). Within each group, other track parameters such as the railpad stiffness (k_p) and sleeper bending (EI_s) stiffness vary over their entire range of interest. The effect of the ballast stiffness on the LWPS is shown in Fig. 12 (b). Despite the varying railpad stiffness and sleeper bending stiffness within each group, there were distinct changes in the LWPS between the different groups.

Fig. 12 (c)–(h) show the sensitivity of the LWPS to different track parameters, where each curve represents the average LWPS of all simulation cases within a group. Overall, the LWPS was more sensitive to the ballast stiffness (Fig. 12(c)) and track irregularity (Fig. 12 (d)). The sensitivity of the LWPS to the railpad stiffness is non-linear (Fig. 12(b)). It is sensitive to the rail pad stiffness only when it is low (less than 100 MN/m in the V-track, equivalent to 500 MN/m in a real track) and becomes less sensitive as the rail pad stiffness increases. A similar non-linear relationship was observed on track FRF [7]. This may result in a greater degree of uncertainty when evaluating the railpad stiffness variations for stiff railpads.

Each track parameter shown in Fig. 12 (c)–(h) corresponds to a unique characteristic of the LWPS. Variations in ballast stiffness result in changes in the magnitude and characteristic frequencies of the LWPS below approximately 1000 Hz. The track irregularity level had the greatest influence on the LWPS magnitude in a wide frequency range. The sleeper bending stiffness and railpad stiffness have overlapping frequency ranges. However, the sleeper bending stiffness is more related to the change in the characteristic frequencies, whereas the railpad stiffness is more related to the change in the LWPS magnitude. These distinct characteristics imply that the LWPS can be used to concurrently evaluate track stiffness in multiple layers.

Fig. 12 (g) and (h) show the sensitivity of the LWPS to railpad and ballast damping, respectively. It is evident that the LWPS is insensitive to the railpad and ballast damping within the value ranges specified in this study. There were several reasons for this observation. First, capturing track vibrations using a moving axle box is a transient process in which damping has a limited effect. This is because multiple cycles of a vibration signal are required to define a meaningful damping value based on decay. More cycles are required for lower damping values. The railpad and ballast damping on the V-track were scaled to 1/25 of the damping on the real tracks. Therefore, it is anticipated that railpad and ballast damping on real tracks will have a greater impact on the LWPS. Second, the
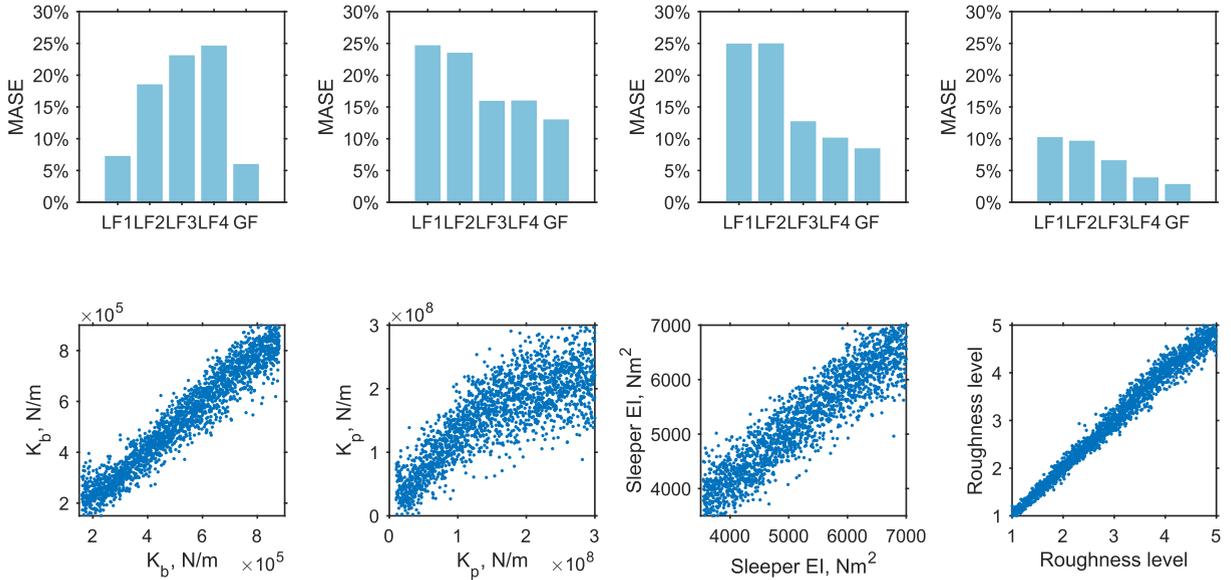
Fig. 14. Performance of the GP models on the test dataset. The four columns correspond to four track parameters, namely ballast stiffness (k_b), railpad stiffness (k_p), sleeper bending stiffness (*EI*), and track irregularity level. The first row shows the MASEs of the predictions made by the GP models with various feature vectors. The second row shows the scatter plots of the true and predicted values obtained using the global feature (GF).

effect of track irregularity may cancel out the damping effect. This is because both damping and track irregularity influence the magnitude of the LWPS but not its characteristic frequencies. For the value ranges specified in this study, the most influential parameter for the magnitude of the LWPS was track irregularity. The effects of railpad and ballast damping may be more apparent if their value ranges were expanded.

3.3. Model selector

To evaluate the performance of the model selector shown in Fig. 11, the mean absolute percentage error (MAPE) was used to quantify the differences between the simulated and the measured ABA features, defined as

$$\text{MAPE} = \frac{1}{N} \sum_{k=1}^{N} \left| \frac{\hat{x}_k - x_k}{\hat{x}_k} \right|, \tag{10}$$

where $\hat{x}_k$ and x_k ($k = 1, 2, \ldots N$) are the elements of the measured and simulated feature vectors (see Eq. (9)) and N denotes the length of the feature vector.

For each ABA feature, the MAPEs were calculated for the top ten matches for each sleeper in the V-track (see Fig. 11 for SN91–SN96). The results are shown in Fig. 13(a). The MAPEs between the best-matching models and the measurements are approximately 4%–6% when the GF is used. Low-frequency local features (LF1) are more difficult to match than high-frequency features (LF2–LF4). The MAPEs significantly increased at the rail joints (indicated by the vertical grey lines). Because rail joints are not considered in the VTI model, the model library is less capable of matching the LWPS caused by local vehicle–joint interaction.

The influence of the model library size on the model selector was also investigated, as shown in Fig. 13(b). The averaged MAPEs over all sleepers were calculated for different library sizes. The MAPEs decreased as the number of model instances in the model library increased. When 100 model instances were used, the largest MAPE was still less than 15% (LF1). The MAPE slightly decreased from 6.2% to 4.7% as the number of model instances increased from 100 to 10,000 when the global feature (GF) was used. These results indicate that the model library can accurately represent the actual physical system. This is because a significant amount of prior information was incorporated into the model library during its construction. This information includes prior knowledge of the design values of the track parameters and the physics of the vehicle track system encoded in the VTI model.

3.4. GP models

The performance of the GP models on the test set (see Section 2.5.5) was evaluated. The trained GP models are denoted as $g(\mathbf{x})$, and the test set is denoted as $\widetilde{D} = \left\{ \widetilde{\mathbf{x}}^{(j)}, \widetilde{p}^{(j)}, j = 1, 2, \ldots, N \right\}$. Trained GP models were employed to predict the track parameter values in the test set given the corresponding feature vectors (i.e., $\widetilde{p}^{(j)*} = g\left(\widetilde{\mathbf{x}}^{(j)} \right) j = 1, 2, \ldots, 5000$). The mean absolute scaled error (MASE) [38] was
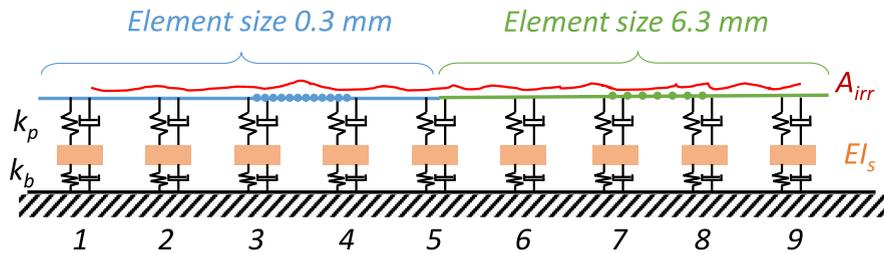
Fig. 15. Track models used for numerical examples.
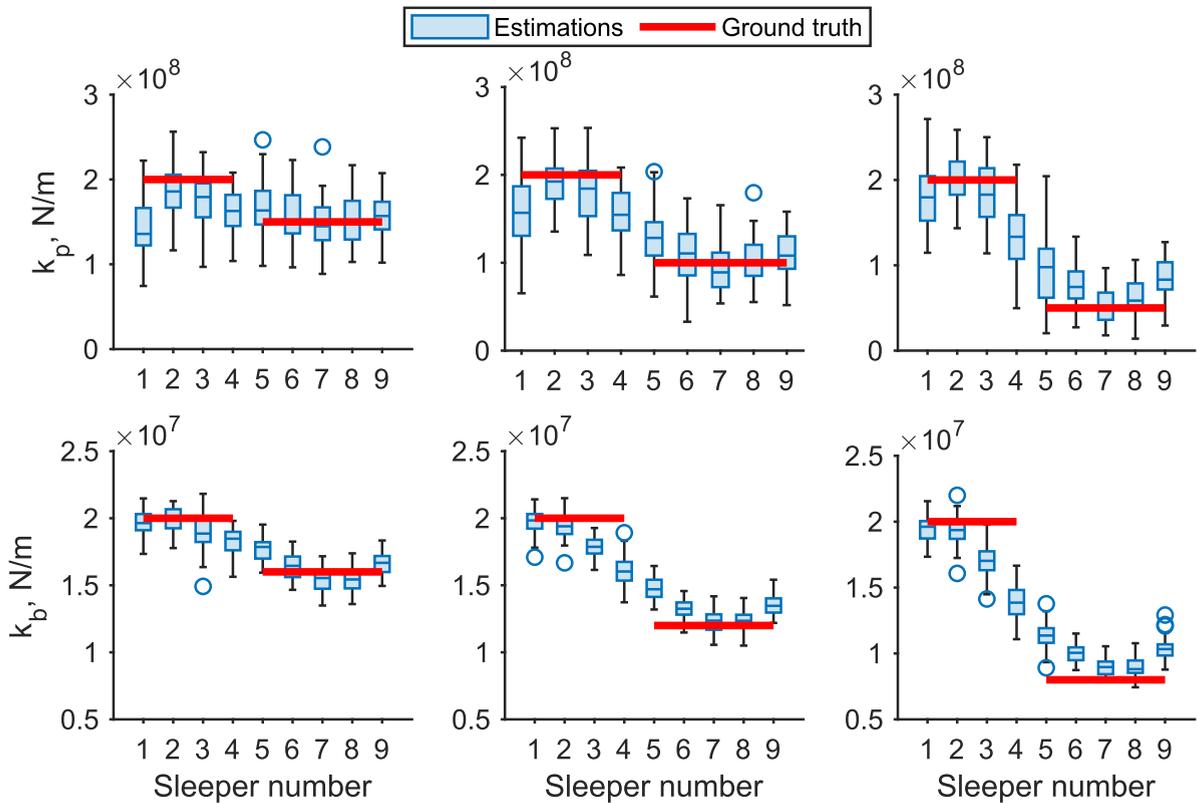
Fig. 16. Robustness of track stiffness evaluations to simultaneous changes of ballast and railpad stiffness. Three cases are considered as shown in the three columns. First column: parameter set A0 at sleepers 1 to 4 and A1 at sleepers 5 to 9. Second column: parameter set A0 at sleepers 1 to 4 and A2 at sleepers 5 to 9. Third column: parameter set A0 at sleepers 1 to 4 and A3 at sleepers 5 to 9. See Table 4.1 for the values of the parameter sets.

used for the evaluation, which defines the mean absolute error of the predicted values of a parameter divided by the value range of the parameter, i.e.,

$$MASE = \frac{1}{N} \sum_{j=1}^{N} \left| \frac{\widetilde{p}^{(j)} - \widetilde{p}^{(j)^*}}{\max(p) - \min(p)} \right|. \tag{11}$$

MASE is useful for comparing the GP model performance across parameters with different scales. The MASEs for the four track parameters are presented in the first row of Fig. 14. The track irregularity level can be accurately predicted with an MASE of less than 5 %. The MASE was still less than 10% even when only local features were considered. Ballast stiffness can also be predicted with high accuracy using GF or LF1 (approximately 5% MASE). The MASEs for the predictions of sleeper bending stiffness using LF3, LF4, and GF were approximately 10%, whereas those for the railpad were approximately 15%. This means that for the railpad stiffness with a value range of approximately 300 MN/m considered in this study, the expected absolute error of the prediction is approximately 45 MN/m.

Table 4.1
Parameter values used in the numerical examples.

Parameter set No.	k_b, N/m	k_p, N/m
A0	2e7	2e8
A1	1.6e7	1.5e8
A2	1.2e7	1e8
A3	8e6	5e7
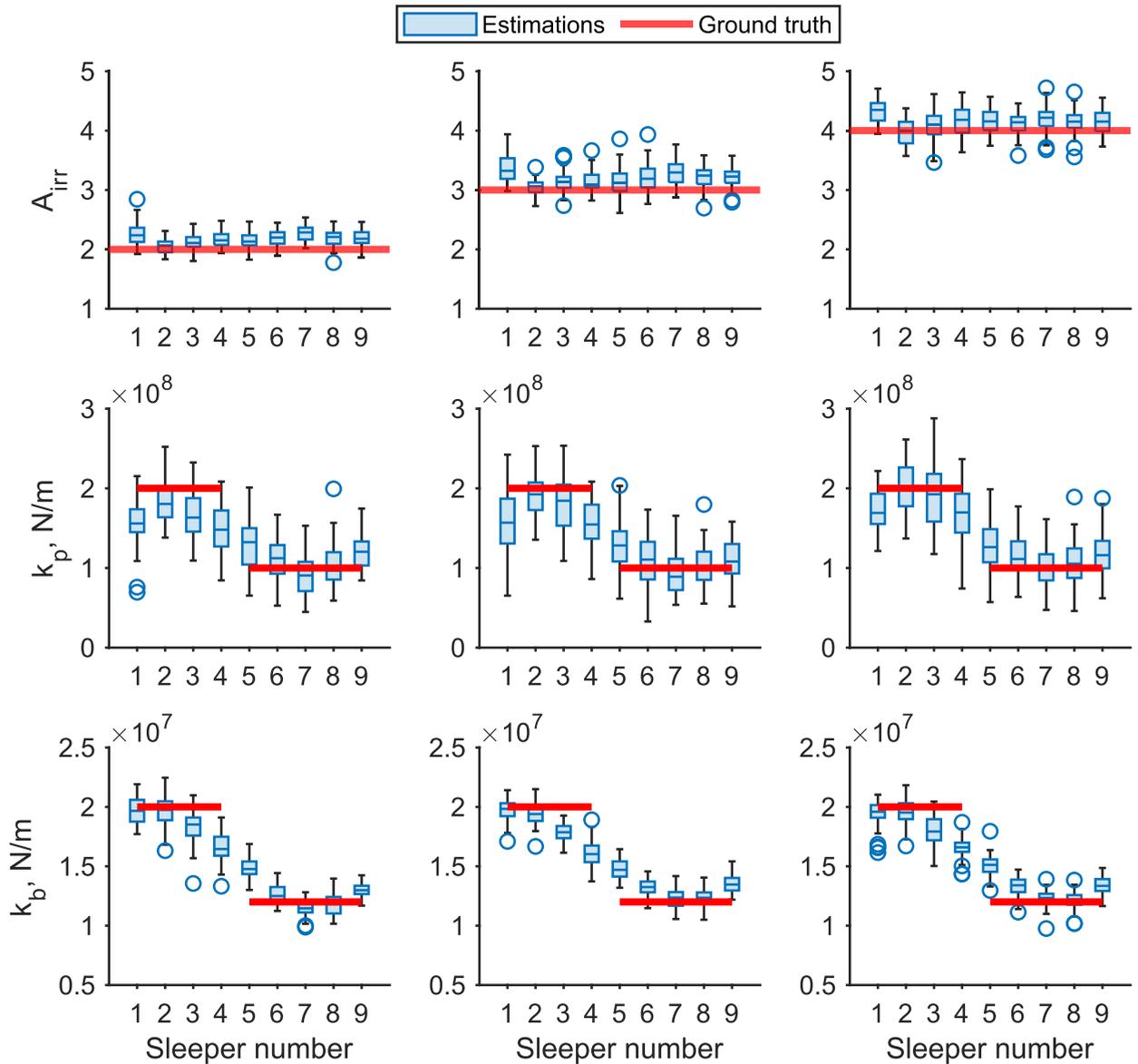
Fig. 17. Robustness of track stiffness evaluations to different track irregularity levels A_{irr}. Three cases are considered as shown in the three columns. In all three cases, the railpad and ballast stiffness are kept constant with the parameter set A2 in Table 4.1. The track irregularity levels are 2, 3 and 4 for the three cases, respectively.

Notably, the feature vectors with relatively low prediction errors for each track parameter correspond well with the sensitive features of the track parameters shown in Fig. 12. This implies that the GP model inferred the correct physics underlying the data evaluations. For instance, when evaluating ballast stiffness, the GP model automatically weighs more on LF1, which is consistent with the sensitivity analysis. In summary, the GF outperformed the LFs for all four track parameters. Certain LFs can produce MASE levels similar to those of the GF for a particular track parameter, because they are sensitive to changes in that track parameter.
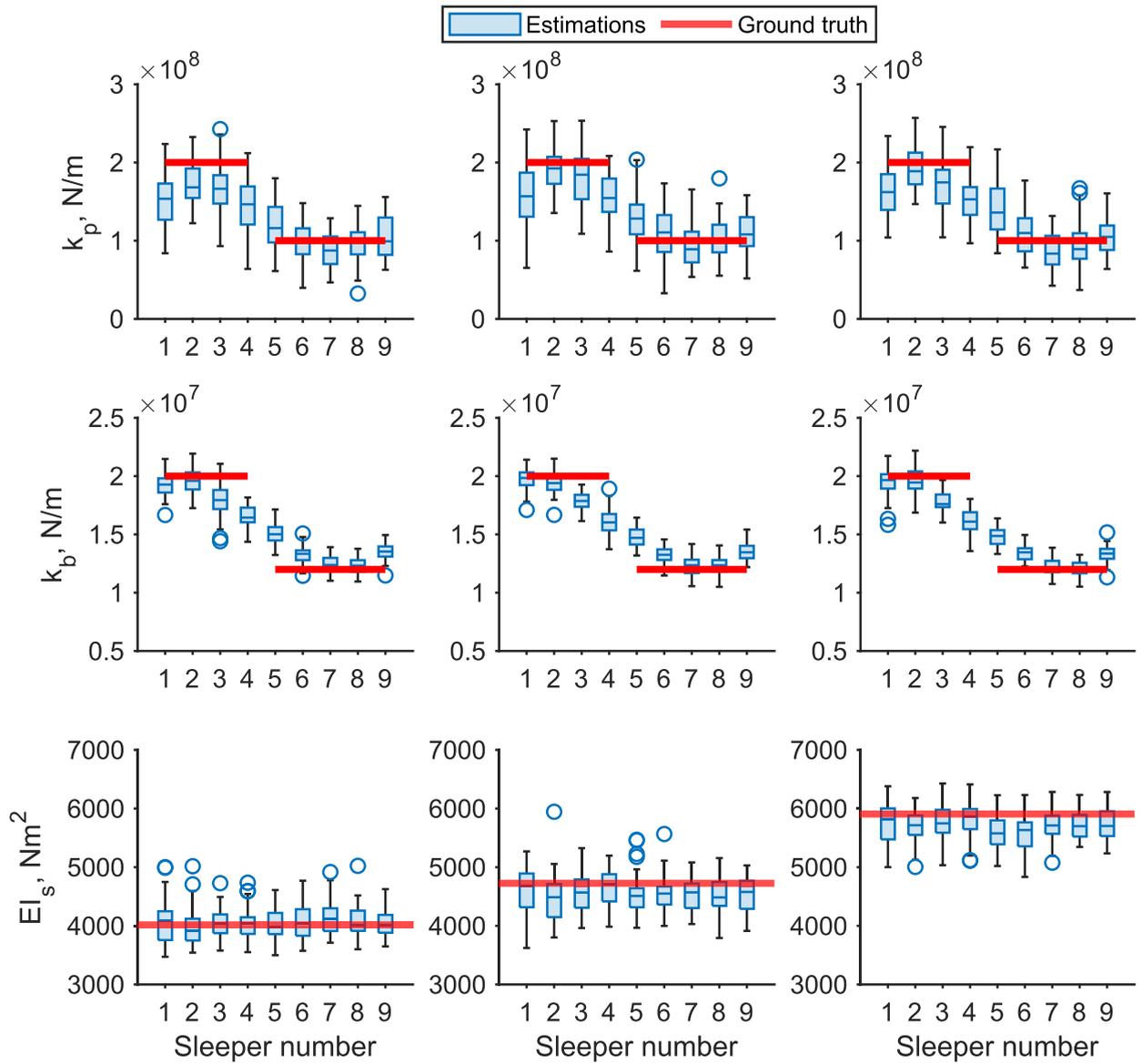
Fig. 18. Robustness of track stiffness evaluations to different sleeper bending stiffness EI_s. Three cases are considered as shown in the three columns. In all three cases, the railpad and ballast stiffness are kept constant with the parameter set A2 in Table 4.1. The sleeper bending stiffness for the three cases is indicated by the ground truths shown in the last row.

The second row shows scatter plots of the true and predicted values for each track parameter calculated using the GF. The prediction errors were smaller when the railpad stiffness was low and increased as the railpad stiffness increased. This is because the ABA features became less sensitive to high railpad stiffness values, as shown in Fig. 12. The prediction errors for the remaining three track parameters were constant across the range of values.

4. Numerical examples

In this section, the robustness and accuracy of the proposed method are demonstrated using numerical examples. Fig. 15 shows the track model used for the numerical examples. The track section contains nine sleepers. The rail was discretized between sleepers 1 and 5 with a fine mesh (element size: 0.3 mm) and with a course mesh between sleepers 5 and 9 (element size: 6.3 mm).

4.1. Changing stiffness in railpad and ballast

Three simulation cases were considered, as shown in the three columns of Fig. 16. For all three cases, the ballast and railpad
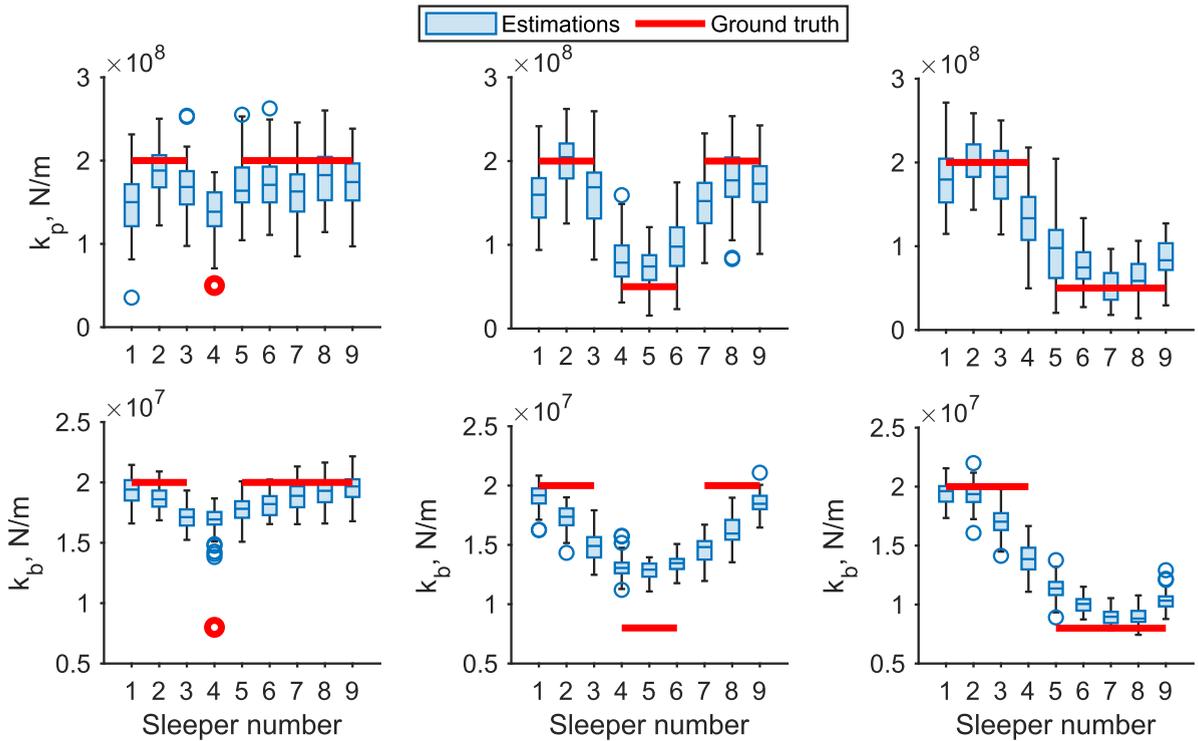
Fig. 19. Effect of local stiffness changes. Three cases are considered as shown in three columns. First column: stiffness reduction at sleeper 4, which is indicated by the red circle. Second column: stiffness reductions at sleepers 4, 5, and 6. Third column: stiffness reductions at sleepers 5 to 9. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stiffness were kept constant at sleepers 1–4 using the parameter set A0, as shown in Table 4.1. The stiffness of the railpad and ballast at sleepers 5–9 were simultaneously reduced to three different levels (i.e., parameter sets A1, A2, and A3). This was performed to simulate the presence of multiple faults with different severities.

The stiffness values evaluated using the proposed method agreed well with the ground-truth values. The GP models predicted gradual stiffness changes between sleepers 4 and 5 instead of the abrupt changes indicated by the ground truths. This is because we assumed a uniform stiffness across different sleepers in the FE model. The modelling error is discussed in detail in Section 4.3.

4.2. Effect of track irregularity level

The robustness of the track stiffness evaluations at different track irregularity levels was investigated. Three cases were considered, as shown in Fig. 17. The railpad and ballast stiffness values were kept constant using the parameter set A2 in Table 4.1. The track irregularity levels were set to 2, 3, and 4 for the three cases. The results show that the track stiffness values and track irregularity levels can be correctly evaluated using the proposed method. This demonstrates the robustness of the method against changing track irregularities.

4.3. Effect of modelling errors

The effects of three types of modelling errors are investigated.

4.3.1. Modelling of bolts in sleepers

As mentioned in Section 2.4.1, sleeper bolts (Fig. 1 (f)) were not considered in the FE model. This modelling error was accounted for by treating the sleeper bending stiffness (EI_s) as an uncertain parameter. Fig. 18 shows the evaluation of the track stiffness with varying sleeper bending stiffnesses. The railpad and ballast stiffness values were kept constant using parameter set A2 in Table 4.1. Sleeper bending stiffness was changed, as shown in the last row of Fig. 18. It can be observed that the proposed method can correctly differentiate between railpad, ballast, and sleeper bending stiffness. This demonstrates the robustness of the method to this modelling error.

4.3.2. Uniform stiffness assumption

The FE model assumes a uniform stiffness distribution across multiple sleepers. This assumption carries over to the GP models
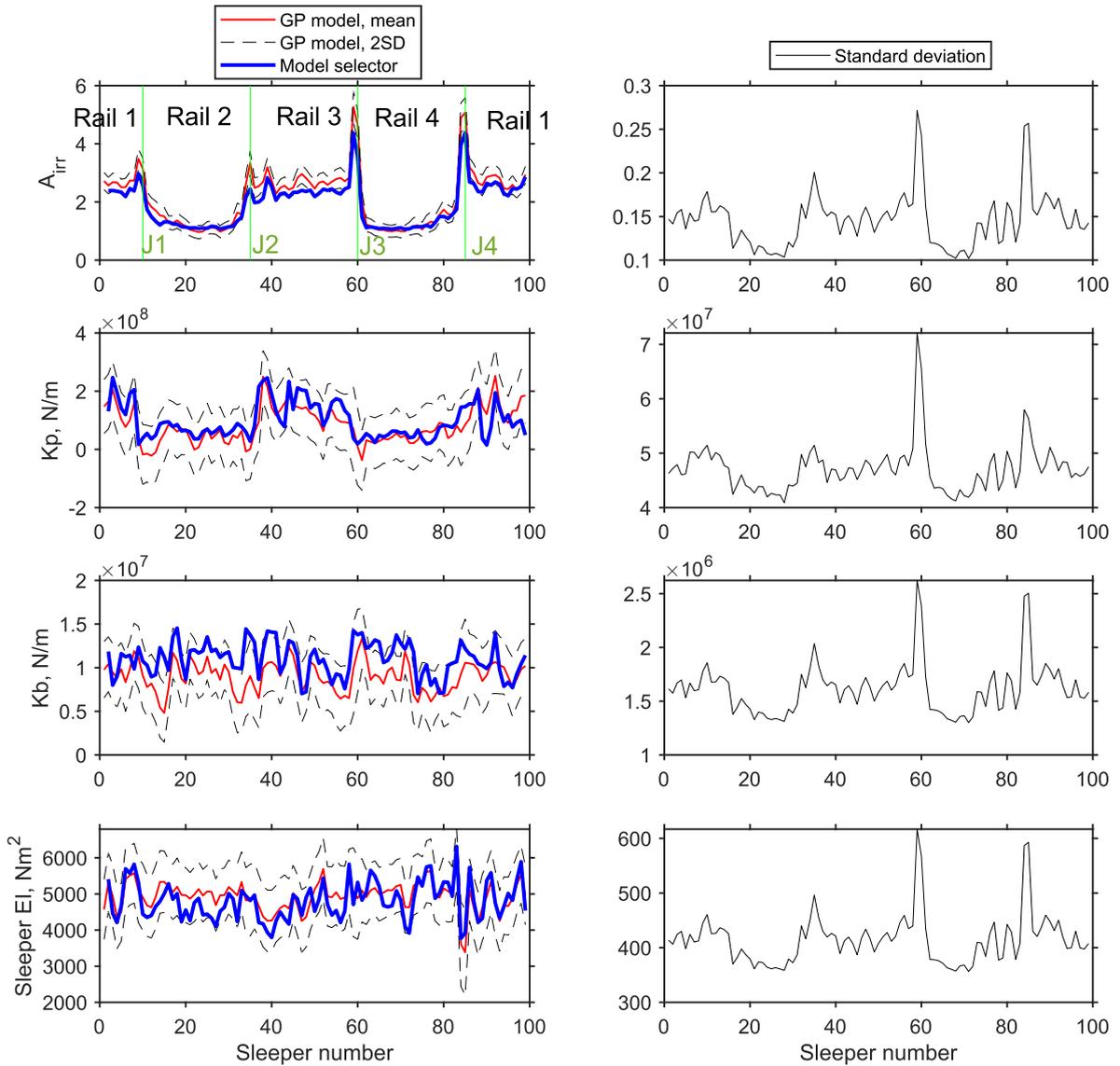
Fig. 20. Track irregularity level (first row) and stiffness variations (second to fourth rows) for the V-track. The left column shows the comparisons of the GP models and the model selector. The results of the model selector represent the averaged values of the top ten matches. The vertical green lines indicate the position of rail joints. To demonstrate more clearly the prediction uncertainty, the right column shows the standard deviations for the evaluations at each sleeper. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

trained using the data generated by the FE model. Simultaneous local stiffness reductions for the railpad and ballast were simulated at different numbers of sleepers, as shown in Fig. 19. It can be observed that the estimated stiffness values are more accurate when stiffness reduction occurs at more sleepers. The GP models exhibit a local stiffness reduction when the stiffness is reduced for only one sleeper (first column in Fig. 19), albeit with larger prediction errors. This is because the evaluations by the GP models represent an averaged stiffness across adjacent sleeper spans, owing to the assumption of uniform stiffness. Nonetheless, such an assumption is used in many applications [3,12] because it significantly lowers the computational cost.

4.3.3. Discretization error

The rail was meshed with two different elemental sizes, as shown in Fig. 15. The GP models used for the stiffness evaluations were trained on the track section with a finer mesh (i.e., between sleepers 1 and 5). However, these models also provided accurate evaluations of the track stiffness on the track section with the courser mesh (i.e., between sleepers 5 and 9). This indicates that the discretization error was negligible for the cases considered in this study.
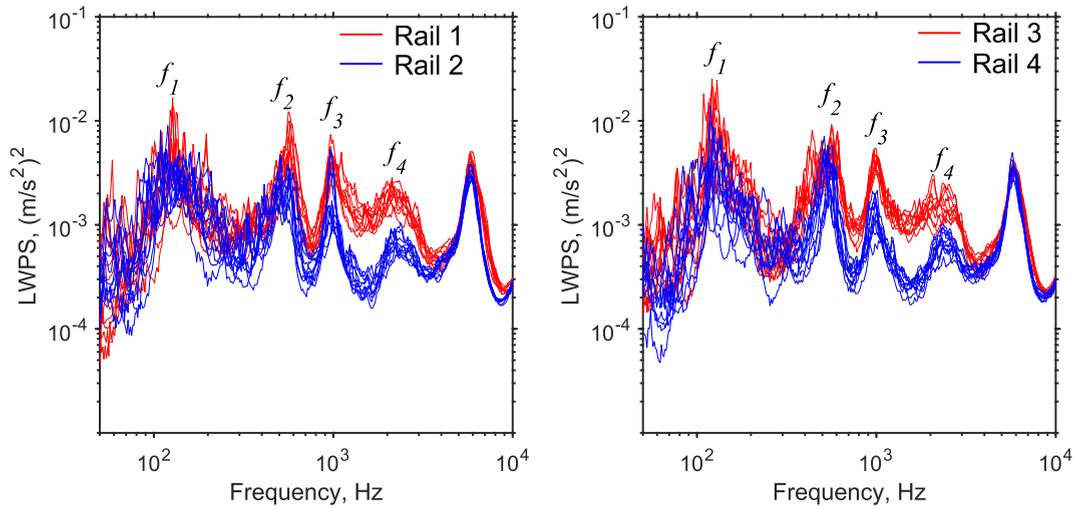
Fig. 21. Comparisons of LWPS measured at different rail segments.
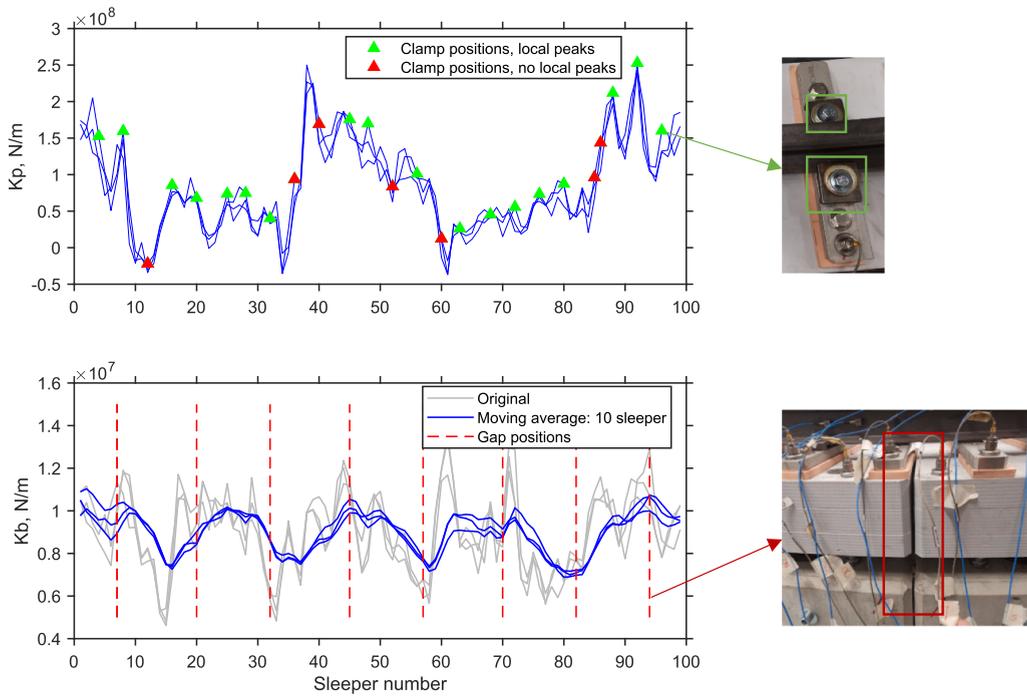
Fig. 22. Correlations between track stiffness variations and structural changes in the V-track. The upper plot shows the correlation between the clamp positions and the local peaks of the evaluated railpad stiffness. The lower plot shows the correlation between the gap positions and variations in ballast stiffness. The results of three repeated measurements are shown in both plots.

5. Case study

The proposed method was applied to evaluate the track stiffness of the V-track using measured ABA. The results of the two proposed data-driven models were first compared. Then, the evaluated stiffness variations were compared with the observed track structure variation. Subsequently, track stiffness evaluated by ABA was compared with those obtained using hammer tests [7]. The effects of vehicle speed and time efficiency were also discussed.
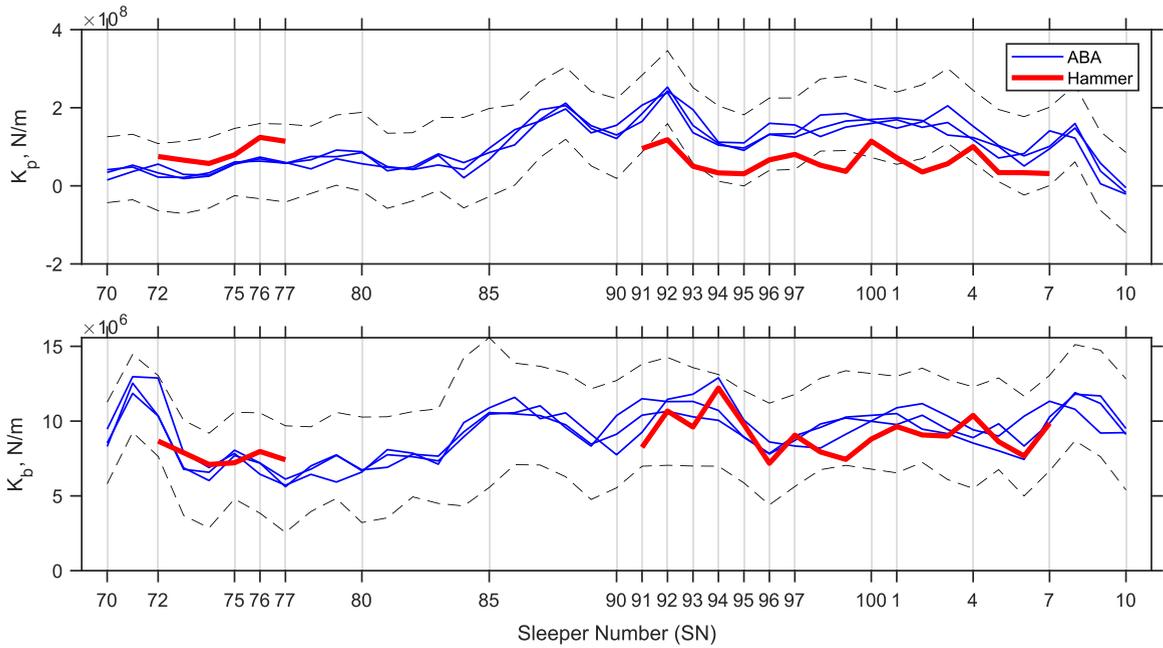
Fig. 23. Comparisons of track stiffness evaluations by ABA and hammer tests. First row: railpad stiffness; second row: ballast stiffness.

5.1. Comparison of the model selector and GP model

Fig. 20 shows the results obtained using both the GP models and model selectors. The general trends of the evaluations using the two methods were in good agreement. The entire test ring comprised four rail segments (Rails 1–4), which were connected via four rail joints (J1–J4), as illustrated in the first subplot of Fig. 20. Rails 1 and 3 exhibited higher levels of track irregularity and railpad stiffness than Rails 2 and 4. In addition, there is a local increase in the level of track irregularity at the rail joints. These increases at J3 and J4 were greater than those at J1 and J2, implying that J3 and J4 were in worse conditions. This is confirmed by visual inspection. The sleeper bending stiffness (EI) fluctuates at approximately 5000 Nm^2, which is more than twice the nominal value (2300 Nm^2). This reflects the effect of the sleeper bolts used in the V-track (see Fig. 1(f)). The bolts constrained the sleeper bending movement, which was not considered in the VTI model. Furthermore, as shown in the right column, the degree of evaluation uncertainty for all the parameters significantly increased at the four rail joints, indicating larger modelling errors. This is because the rail joints were not considered in the VTI model.

Direct measurements were not available to verify the high irregularity level and railpad stiffness at Rails 1 and 3, as shown in Fig. 20. However, this can be justified by comparing the LWPS measured at different rail segments, as shown in Fig. 21. The LWPS magnitudes were larger for Rails 1 and 3 than those for Rails 2 and 4, most notably between f_2 and f_4 (700–4000 Hz). According to the sensitivity analysis shown in Fig. 12, a local magnitude change at f_2 and f_3 indicates a railpad stiffness variation. In addition, magnitude changes in a wide frequency range indicate a variation in the track irregularity level. This explains why the data-driven models predicted higher irregularity levels and railpad stiffnesses for Rails 1 and 3.

Furthermore, the magnitude differences were not evenly distributed. The differences were larger between approximately 700 and 4000 Hz. This could be due to the unevenly distributed PSD for the track irregularity, as shown in Fig. 8. The track irregularity was measured on Rail 1 (between SN91 and SN96), which has a higher PSD magnitude between 400 and 4000 Hz. This frequency range overlaps with that where larger LWPS differences exist.

5.2. Comparison with observed track structure variations

The track stiffness variations were compared with the track structure variations observed on the V-track to further verify the evaluation results. The rail was periodically clamped to the sleepers in the V-track, as shown in Fig. 1(d). Such periodicity in the track structure was well reflected by the local peaks in the railpad stiffness variations, as shown in the upper plot in Fig. 22. The green triangles denote the clamp positions corresponding to the local peaks.

Another type of track structure variation considered here is the gap between the plywood pieces in the substructure, as shown in the lower plot in Fig. 22. Although the substructure was not modelled in the VTI model, such changes could still be captured as ballast stiffness variations. The blue curves in the lower plot of Fig. 22 represent the moving-averaged ballast stiffness with a window of ten sleeper spans. The gap positions correspond well to the locations of local ballast stiffness changes.
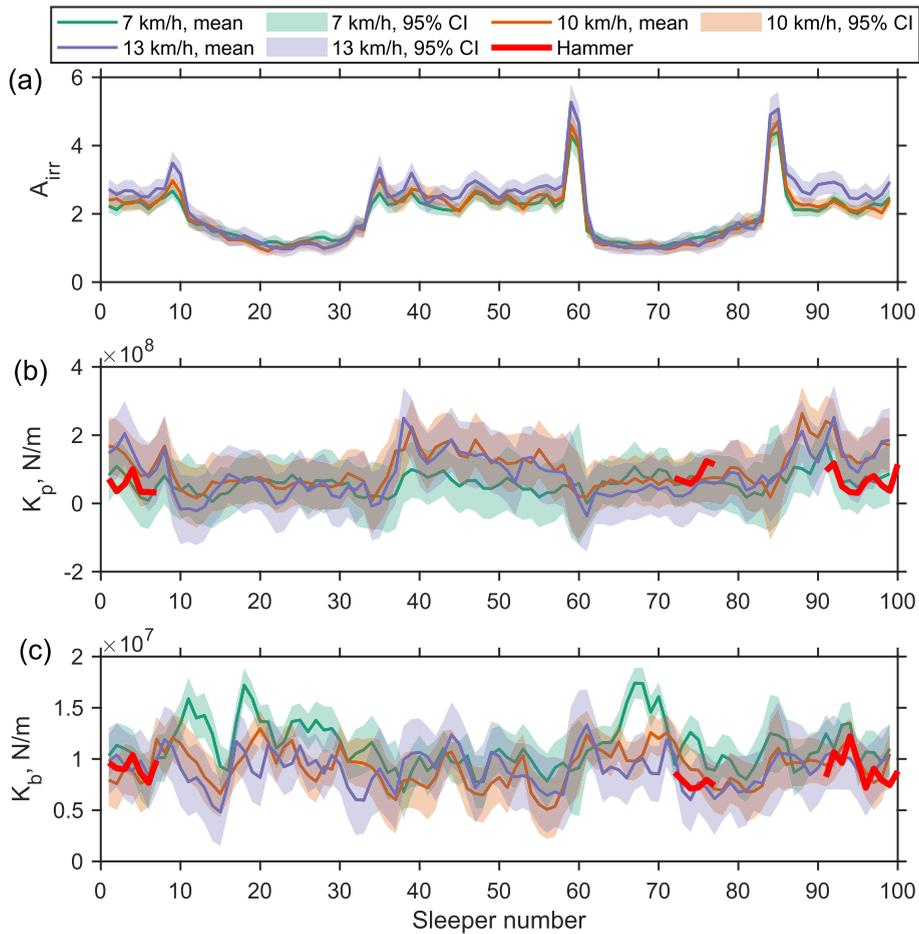
Fig. 24. Comparison of track parameter evaluations by ABA at different vehicle speeds. The 95% confidence intervals (CI) for the evaluations are indicated by the shaded areas. Hammer test results are also shown in bold red lines. (a) Track irregularity level. (b) Railpad stiffness. (c) Ballast stiffness. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.3. Comparison of the evaluations by ABA and hammer test

The track stiffness evaluated by the ABA was compared with that obtained by hammer tests [7]. Point FRFs were measured using hammer tests on the rail top above 23 sleepers (SN72–SN77 and SN91–SN7). The railpad and ballast stiffnesses were subsequently obtained based on the measured FRF features [7], as indicated by the red lines in Fig. 23.

The railpad stiffness values evaluated using the two methods followed similar trends for each segment. The clamps at SN76, 92, 96 and 4 were captured using both methods as local peaks in the stiffness variations. The local peaks may not exactly correspond to the clamp locations. For example, the railpad stiffness evaluated by ABA and hammer tests exhibited local peaks at SN3 and SN97, respectively. This is a shift of one sleeper relative to the clamp position. A large local dip was observed at SN90 for the railpad stiffness evaluated by ABA. This corresponds to a missing railpad observed at SN91 (Fig. 1 (e)) also with a one-sleeper shift. The reason for these shifts could be that the adopted track model assumes uniform track stiffness. Therefore, the track stiffness evaluated by the proposed method represents the averaged stiffness values within a few sleeper spans. As a result, the proposed method may not always detect precise locations with considerably local stiffness changes.

Additionally, the railpad stiffness values evaluated by ABA were higher between SN91 and SN7 than those evaluated by hammer tests. This is because the loading conditions for the two methods were different. The rail and sleepers may not come into direct contact under unloaded conditions because the rail is not fastened to each sleeper. For the hammer test, the hammer loads were relatively small. They could not close any possible gaps between the track layers. This led to lower railpad stiffness evaluations, as shown in the track section between SN91 and SN7. In contrast, when the track was loaded, the large wheel loads enabled firm contact between the various track layers, resulting in increased track stiffness. Considering this, the evaluation of track stiffness using ABA is more accurate than that using a hammer test because it represents the actual loaded condition.

The ballast stiffness values evaluated using the two methods are in good agreement. Both methods predicted a large stiffness change between SN94 and SN95 because of the gaps in the substructure, as discussed in Section 6.2. Additionally, the hammer test method

showed local peaks at the clamp locations (e.g., SN76, SN92, SN97, and SN4) in the ballast stiffness variations, despite the clamps being in the fastening system. This implies that the effect of fastening stiffness on the ballast layer is difficult to separate using the hammer test method. On the other hand, the ABA method does not show such changes. This can be attributed to different loading conditions. Under the loaded conditions, the wheel load was significantly greater than the clamp force. Consequently, the stiffness differences between the clamped and unclamped conditions were not significant enough to cause a difference in the ballast layer. This demonstrates that the ABA can discriminate track stiffness among different layers under loaded conditions.

5.4. Effect of vehicle speed

Fig. 24 compares the track irregularity level, railpad stiffness, and ballast stiffness evaluated at different speeds using the ABA. In general, the proposed method is robust to varying vehicle speeds. The track irregularity levels evaluated at different speeds agreed well. For evaluating the railpad and ballast stiffness, the proposed method is more robust at higher speeds of 13 and 10 km/h. However, it leads to some deviations at a speed of 7 km/h. This is because the track stiffness is load dependent, as discussed in Section 5.3. Dynamic loading decreases as the speed decreases, which may result in lower stiffness evaluations. For instance, the railpad stiffness evaluated at 7 km/h was lower than that at 13 km/h and 10 km/h at several track sections (i.e., among sleeper numbers 0–10, 40–60, and 90–100), while agreeing better with the hammer test results.

The proposed method was validated in the laboratory with vehicle speeds up to 13 km/h, which is equivalent to 65 km/h on real tracks. The running speed of an actual train can exceed 65 km/h. However, prior studies [20,21] have shown that a high sampling rate combined with a low to moderate vehicle speed is necessary to ensure accurate and high-resolution track stiffness evaluations. This ensures that sufficient data points are collected within a short track segment, such as a sleeper span.

Future research should include field tests to investigate the effects of higher vehicle speeds. A key finding of this study is that the characteristic frequencies of the ABA remain constant with changing vehicle speeds (see Fig. 10 (a)). This ensures the robustness of the proposed method to varying vehicle speeds. Therefore, it is critical to verify whether the same observations can be obtained during field tests at higher vehicle speeds.

5.5. Time efficiency

The GP model can perform rapid assessments of track stiffness. The evaluation time for a single sleeper span was 0.0093 s in the case study with an 8-core CPU (Intel Xeon E5-2687 W v2). This was calculated by dividing the total amount of time spent on the evaluation by the number of sleeper spans. As demonstrated in [7], inverting the covariance matrix is the most time-consuming step in a single evaluation. The prediction procedure is almost instantaneous after the inverse covariance matrix is available because it requires only simple matrix multiplications. In practice, inversion is performed only once at the start of the procedure and is retained in memory for the duration of the evaluation. Consequently, the proposed approach can be easily scaled to long-track sections.

The SSWT procedure is a bottleneck when applied in real-time applications. The SSWT was performed on a signal length of 3485 sampled points for each sleeper in the current case, resulting in an average calculation time of 0.32 s. By comparison, halving the signal length to 1742 points reduces the calculation time to 0.15 s. However, this may come at the expense of decreased precision.

6. Discussions

6.1. Further improvement of the proposed method

The SSWT was used to implement a local time–frequency (TF) feature. GP models were used to establish quantitative relationships between the ABA feature and track stiffness. Other TF analysis and machine learning (ML) techniques are available. Exhaustive research on the optimal TF analysis technique and the most precise ML technique is beyond the scope of this study. Nonetheless, these are potential research areas for further enhancement of the proposed method and are discussed in this section.

6.1.1. Time-frequency analysis methods

The goal of the TF analysis is to transfer the time-domain ABA signal to the frequency domain at a local position. This local frequency-domain representation should be sparse such that a low-dimensional feature vector can be used to train the GP models. In this study, the SSWT was used to obtain the LWPS as the feature vector. In Section 2.3, the advantages of the SSWT method were demonstrated over the conventional CWT method. The SSWT method focuses on the spectrogram energy of instantaneous frequency curves and results in a sharpened TF representation. Considering this, SSWT is a sparser representation than CWT.

Many other TF analysis techniques can lead to sparse representations of ABA signals. A promising technique that can be applied to the problem presented in this study is the family of sparse signal decomposition methods [39,40]. This technique decomposes signals using a linear combination of atoms in an overcomplete dictionary. Compared to the Fourier transform (FT) or CWT/SSWT, sparse decomposition methods can result in an even sparser representation of a signal because of the overcomplete dictionary. For example, the resonance-based sparse decomposition method [41] proposes the use of overcomplete atoms constructed based on a signal's oscillation (resonance) property.

It is beyond the scope of this study to construct overcomplete atoms for the ABA signals and decompose the ABA signals using sparse decomposition methods. In future research, this technique will be explored to improve the performance of the proposed method because it can result in a sparser feature vector of the ABA signals.
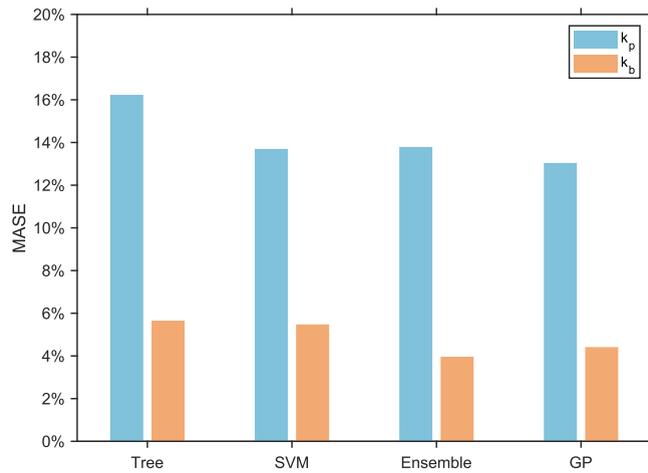
Fig. 25. Comparison of different machine learning methods for evaluating track stiffness. The model types for each method are the Medium Tree, Medium Gaussian SVM, Bagged Trees (Ensemble), and GP with the exponential kernel. The models are tested on the test set shown in Section 2.5.5 and are evaluated based on the mean absolute scaled error (MASE) (Eq. (11)).

Table D1
Parameters and value ranges used in simulations.

Parameter	Design value			Value Range
	Scaled track	Real track	Scale factor	
Railpad				
Stiffness, k_p (MN/m)	100–312	500–1560	1/5	10–300
Damping, c_p (kN.s/m)	1.2–2.7	30–68	1/25	4–9
Ballast (rubber, half-track)				
Stiffness, k_b (MN/m)	8–14	40–70	1/5	4–22
Damping, c_b (kN.s/m)	0.4–1.3	10–32	1/25	0.4–1.3
Sleeper (steel, half-track)				
Spacing, L_s (m)	0.12–0.13	0.60–0.65	1/5	0.125
Weight, m_s (kg)	1–1.36	120–170	1/125	1.12
Length (m)	0.24	–	–	0.24
Density, ρ (kg/m^3)	7800	–	–	7800
Cross section area, A(m^2)	6e-4	–	–	6e-4
Young's modulus, E (GPa)	210	–	–	–
Cross section modulus, I (m^4)	1.125e-8	–	–	–
EI (Nm^2)	2362.5	–	–	3543–7086
Rail (Steel, S7 profile)				
Density, ρ (kg/m^3)	7800	–	–	7800
Cross section area, A (m^2)	8.65e-4	–	–	8.65e-4
Young's modulus, E (GPa)	210	–	–	210
Cross section modulus, I (m^4)	5.15e-7	–	–	5.15e-7
Track irregularity level, A_{irr}	–	–	–	1–5
Wheel				
Radius (m)	0.065	–	1/5	0.065
Weight (kg)	44	–	–	44
Wheel load (N)	4500	–	–	4500
Speed, v (km/h)	7, 10, 13	35, 50, 65	1/5	7, 10, 13

6.1.2. Machine learning methods

The GP regression method was compared with three other commonly used non-linear ML regression methods: decision-tree regression, support vector machines, and tree ensembles. The three methods were trained using the same training data as the GP models. The MATLAB regression learner app was used for training. All submodel types available within each method were trained. The model with the smallest training error was selected for each method. The submodel types for each method were medium tree, medium Gaussian SVM, bagged trees (ensemble), and GP with an exponential kernel.

The performances of the ML models on the test set are compared in Fig. 25. The GP and ensemble models showed the most accurate results for the evaluation of the railpad stiffness (k_p) and ballast stiffness (k_b), respectively. However, a clear advantage of the GP regression method is that it can provide an uncertainty quantification for the evaluated track stiffness. For example, as shown in Fig. 20, the degree of evaluation uncertainty for all parameters significantly increased at the four rail joints, indicating larger modelling errors.

Notably, the comparison serves only as a rough benchmark because all models were trained using default settings without any form of optimization. In future research, prediction errors can be reduced by optimizing the hyperparameters of the ML models or by fusing the prediction results obtained by different methods [7].

6.1.3. Effects of rail surface defects and discontinuities

Rail-surface defects and discontinuities, such as squats and joints, can cause sudden variations in ABA signals. They may affect the accuracy of track stiffness evaluations. Fig. 20 demonstrates that the proposed method exhibits reduced accuracy when rail joints are present. However, the proposed method cannot detect rail defects or discontinuities. Neither can the method differentiate them from track stiffness evaluations. This is because the rail defects and discontinuities were not considered in the VTI model. Therefore, their corresponding ABA features were not included in the model library used to train the GP models.

In future research, the proposed method can be combined with ABA techniques that detect rail surface defects [22]. This allows for identifying the locations where rail defects are present. Moreover, it may be beneficial to augment the parameter space defined in Eq. (5) with the parameters describing the rail defect geometries. Thus, a global sensitivity analysis can be performed to identify the unique ABA features that can differentiate between rail surface defects and track stiffness variations. For instance, increased magnitudes in the WPS within certain frequency ranges could indicate singular rail surface defects [22], whereas shifts in characteristic frequencies may be useful for evaluating spatial track stiffness variations.

6.2. Application of the proposed method in the field

The proposed digital twin framework was validated in the lab. In future research, it is intended to validate this framework using real tracks. Difficulties and issues may arise when the proposed method is applied in the real world.

First, one or more track sections with variations in spatial or temporal stiffness variations must be selected. In a controlled lab environment, stiffness variations can be artificially created on the test rig by setting missing railpads, clamps, and bolts, as shown in Fig. 1. This is not possible in operational lines. Instead, specific track characteristics and conditions observed during track inspections must be examined to identify track sections with stiffness variations. For example, measured track geometry variations may be linked to variations in stiffness. Different fastening systems, transition zones, and property changes in embankments or subsoil may signify spatial stiffness variations. Track maintenance work, such as tamping, may cause stiffness variations over time (i.e., before and after maintenance).

Second, the track stiffness evaluated using the ABA must be validated using other methods. In this study, a hammer test method was used for validation. It was demonstrated that the loading conditions can influence the evaluated track stiffness. The load-dependent stiffness must also be considered for real tracks. This requires a larger impact load, such as that of a falling weight [9,10]. Alternatively, the track stiffness evaluated by the ABA can be compared with that obtained under train loads, such as by pass-by measurements [3,11,12].

Third, it is necessary to investigate the effects of a real vehicle on ABA features. The vehicle was simplified as a wheel mass and primary suspension on the V-track. In the real world, a vehicle is in contact with a track by using multiple wheelsets. Rail bending waves between wheelsets are generated due to the VTI [42], which may affect the ABA features. In addition, vehicle dynamics and flexible wheelsets may add additional features to ABA signals.

Finally, the proposed DT approach aims to provide a generic solution for evaluating the dynamic stiffness of ballasted tracks without focusing on a specific track type. In the real world, different track types exist due to different rail types, fastening systems, sleeper types, and sleeper spacings. These different track types can be considered by adjusting the input value ranges of the corresponding track parameters in the VTI model. In addition, different track types may exhibit different baseline ABA characteristics. It is important to validate the VTI model for a different track type before using it to generate training data for the GP models (see examples of model validations in Section 3). Other track types, such as slab tracks, are beyond the scope of this study.

7. Conclusions

This study developed a digital twin framework to rapidly and robustly evaluate track stiffness at different track layers. The framework comprised a physics-based VTI model, a model library, and two data-driven methods. The capability of the digital twin was demonstrated using numerical examples and lab experiments.

Both numerical examples and laboratory experiments showed that the proposed framework can evaluate the stiffness of multiple track layers simultaneously at a sleeper spacing resolution. The track stiffness evaluated using the proposed method was robust to different input track parameters, such as track irregularities, damping, sleeper bending stiffness, and vehicle speeds.

The proposed method was compared to an FRF-based hammer test method to evaluate the stiffness of three track sections for 23 sleepers on the V-track. The average values of the ballast and railpad stiffnesses evaluated using the two methods agreed well, indicating the validity of the proposed methods. Additionally, the proposed method is more accurate than the hammer test method, because it represents track stiffness under operational load conditions.

Controlled lab measurements and observations showed that the developed method can evaluate the stiffness of the fastening system and ballast, as well as their variations in space. When ABA is regularly measured from tracks, the fastening and ballast stiffness and their changes over time can be assessed. This can provide a better understanding and prediction of the track geometry changes in space and time. Thus, efficient and effective track maintenance can be performed on the right components (fastening or ballast) at the right places and at the optimal time.

The proposed digital twin framework is highly scalable and computationally efficient. This is beneficial because a large amount of data is generated due to the high sampling rate required for high-resolution evaluations. In the case study, it was demonstrated that the proposed method compressed 3485 data points of a time-domain ABA signal into a sparse four-element vector within approximately 0.33 s per sleeper span. Because of the simple and fast calculation process, it can be implemented on the edge of the infrastructure, such as an onboard computer, without the need to transfer a large amount of data to the cloud or operational centers.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

Appendix A. Algorithms for CWT and SSWT

The algorithms [31,32] and parameters used to calculate the CWT and SSWT are briefly introduced. For a section of the ABA signal $s(t)$, the wavelet coefficients $W_s(a, b)$ of the CWT can be calculated as

$$W_s(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(t) \overline{\psi\left(\frac{t - b}{a}\right)} \, dt, \tag{A.1}$$

where a is the wavelet scale, b is the time, ψ is the mother wavelet, and the Morlet wavelet is used as the mother wavelet. The Morlet wavelet is a type of tapered cosine wave that resembles structural and mechanical vibration signals. Compared to other commonly used complex wavelets, the Morlet wavelet shows better time–frequency localization [43]. As a result, it has been extensively employed for analyzing structural and mechanical vibration signals [44–47] and, in particular, for ABA signals [22,48–50].

The scale parameter a is discretized into $N_j + 1$ scales, and each scale is defined as

$$a_j = s_0 2^{j/N_v}, j = 0, 1, ..., N_j, \tag{A.2}$$

where s_0 is the base scale and its value is $2/f_s$ (f_s is the sampling frequency of the signal $s(t)$). N_v is a parameter that controls the resolution of scales, referred to as the number of "voices per octave". The larger the N_v, the finer the scales are discretized. In this study, 48 voices per octave were used. This means that 48 scales are required to increase the scale by an octave (a double), for example, from $2s_0$ to $2^2 s_0$. N_j is the number of scales minus one and determined by $N_j = N_v \log_2(N_s/2)$, where N_s is the length of the signal $s(t)$. If N_s is not a power of 2, the signal $s(t)$ is padded to the next higher power of 2.

For the SSWT, instantaneous frequencies need to be extracted from $W_s(a, b)$. In the scenario satisfying $W_s(a, b) \neq 0$, a candidate instantaneous frequency can be calculated via the following phase transform:

$$\omega_s(a, b) = \frac{-i}{2\pi W_s(a, b)} \frac{\partial W_s(a, b)}{\partial b}. \tag{A.3}$$

The wavelet coefficients $W_s(a, b)$ can then be transformed from the time-scale plane (b, a) to the TF plane $(b, \omega_s(a, b))$.

In the next step, $W_s(a, b)$ is calculated at discrete values a_k, with $\Delta a_k = a_k - a_{k-1}$. Correspondingly, the synchrosqueezed transform $T_s(\omega, b)$ of $W_s(a, b)$ can be calculated at discrete values ω_l, with $\Delta \omega = \omega_l - \omega_{l-1}$, as

$$T_s(\omega_l, b) = \frac{1}{\Delta \omega} \sum_{a_k : |\omega_s(a_k, b) - \omega_l| \leq \Delta \omega/2} \left(W_s(a_k, b) a_k^{-3/2} \Delta a_k \right). \tag{A.4}$$

Appendix B. Equations of motion for the VTI model

Track model

The equation of motion for the track system is expressed as

$$M_t \ddot{U}_t + C_t \dot{U}_t + K_t U_t = F, \tag{B.1}$$

where M_t, C_t, and K_t are the mass, damping, and stiffness matrices of the track, respectively. $\ddot{U}_t$, $\dot{U}_t$, and U_t are the acceleration, velocity, and displacement vectors, respectively, for all the degrees of freedom (DOFs) of the track system. F is a force vector that can be calculated as

$$F = \left[N_{r,s}^w \right]^T F_C(x_w), \tag{B.2}$$

where $F_C(x_w)$ is the wheel-rail contact force at wheel position x_w (see Eq. (B.4)). $N_{r,s}^w$ is the shape function of the rail element that distributes the contact force to the adjacent nodes of the rail beam (see Eq. (C.6)). Appendix B provides the element matrices and shape functions.

Vehicle model

The equation of motion for the vehicle is given by

$$m_w \ddot{u}_w = F_C(x_w) + F_w, \tag{B.3}$$

where m_w and $\ddot{u}_w$ denote the mass and vertical acceleration of the wheel, respectively, and F_w denotes the vertical load applied to the wheel.

Contact model

The wheel-rail contact is modelled as a Hertzian spring, with the contact force at the wheel position x_w calculated as

$$F_C(x_w) = \begin{cases} C_H \big(Z_w(x_w) - Z_r(x_w) - Z_{irr}(x_w) \big)^{3/2} & \text{if } Z_w(x_w) - Z_r(x_w) - Z_{irr}(x_w) \langle 0 \\ 0 & \text{if } Z_w(x_w) - Z_r(x_w) - Z_{irr}(x_w) \geq 0 \end{cases}, \tag{B.4}$$

where $Z_w(x)$, $Z_r(x)$, and $Z_{irr}(x)$ are the vertical coordinates of the wheel, rail, and track irregularity, respectively, and C_H is the Hertzian coefficient and can be approximated as [33]

$$C_H = \frac{2ER^{1/2}}{3(1 - v^2)}, \tag{B.5}$$

where E and v are Young's modulus and Poisson's ratio of the wheel and rail (assumed to be equal for both materials), respectively, and R is the radius of the railhead in the lateral direction.

Track irregularity $Z_{irr}(x)$ is generated as random variables following a power spectral density (PSD) function [51] in the form

$$S(\Omega) = k A_{irr} \frac{\Omega_c^2}{(\Omega^2 + \Omega_c^2)\Omega^2}, \tag{B.6}$$

where $k = 0.25$, $\Omega_c = 0.8242 \text{rad/m}$, and A_{irr} is the track irregularity level that controls the PSD amplitude.

Equations (B.1), (B.3) and (B.4) are solved in the time domain using Newmark integration. The Newton-Raphson iteration was adopted within each time step to ensure the convergence of the contact force. The solution process is implemented in MATLAB.

Appendix C. Element matrices and shape function for the VTI model

Timoshenko beam element for rail and sleepers

The rail and sleepers are meshed with the Timoshenko beam element developed in [34]. Each element has two nodes and four DOFs. The displacement vector for an element is denoted as

$$u^e = [w_1, \theta_1, w_2, \theta_2]^T. \tag{C.1}$$

At each node, only the vertical (w_1, w_2) and in-plane rotational (θ_1, θ_2) degrees of freedom are considered. The subscripts (1,2) indicate the two nodes in an element.

The element stiffness matrix for the rail and sleepers is

$$K_{r,s}^e = \frac{EI}{(1 + \phi)L^3} \begin{bmatrix} 12 & 6L & -12 & 6L \\ & (4 + \phi)L^2 & -6L & (2 - \phi)L^2 \\ & & 12 & -6L \\ \text{symmetric} & & & (4 + \phi)L^2 \end{bmatrix}, \tag{C.2}$$

where L is the length of the element, E is the Young's modulus, I is the area moment of inertia of the cross-section, and $\phi = \frac{12}{L^2}\left(\frac{EI}{\kappa GA}\right)$,

where κ is a shear coefficient, G is the shear modulus and A is the area of the cross-section.

The element mass matrix is given by

$$\boldsymbol{M}_{\mathrm{r,s}}^{\mathrm{e}} = \boldsymbol{M}_{\mathrm{r,s}}^{\mathrm{e},\rho A} + \boldsymbol{M}_{\mathrm{r,s}}^{\mathrm{e},\rho I}, \tag{C.3}$$

where the first part is related to translational inertia

$$\boldsymbol{M}_{\mathrm{r,s}}^{\mathrm{e},\rho A} = \frac{\rho A L}{210(1+\phi)^2} \begin{bmatrix} (70\phi^2 + 147\phi + 78) & (35\phi^2 + 77\phi + 44)\frac{L}{4} & (35\phi^2 + 63\phi + 27) & -(35\phi^2 + 63\phi + 26)\frac{L}{4} \\ & (7\phi^2 + 14\phi + 8)\frac{L^2}{4} & (35\phi^2 + 63\phi + 26)\frac{L}{4} & -(7\phi^2 + 14\phi + 6)\frac{L^2}{4} \\ & & (70\phi^2 + 147\phi + 78) & -(35\phi^2 + 77\phi + 44)\frac{L}{4} \\ \text{symmetric} & & & (7\phi^2 + 14\phi + 8)\frac{L^2}{4} \end{bmatrix}, \tag{C.4}$$

and the second part is related to rotational inertia

$$\boldsymbol{M}_{\mathrm{r,s}}^{\mathrm{e},\rho I} = \frac{\rho I}{30(1+\phi)^2 L} \begin{bmatrix} 36 & -(15\phi - 3)L & -36 & -(15\phi - 3)L \\ & (10\phi^2 + 5\phi + 4)L^2 & (15\phi - 3)L & (5\phi^2 - 5\phi - 1)L^2 \\ & & 36 & (15\phi - 3)L \\ \text{symmetric} & & & (10\phi^2 + 5\phi + 4)L^2 \end{bmatrix}. \tag{C.5}$$

The shape functions for w and θ in Eq. (C.1) are

$$\left[\boldsymbol{N}_{\mathrm{r,s}}^{w}\right]^T = \begin{bmatrix} \frac{1}{(1+\phi)}\left\{2\left(\frac{x}{L}\right)^3 - 3\left(\frac{x}{L}\right)^2 - \phi\left(\frac{x}{L}\right) + (1+\phi)\right\} \\ \frac{L}{(1+\phi)}\left\{\left(\frac{x}{L}\right)^3 - \left(2 + \frac{\phi}{2}\right)\left(\frac{x}{L}\right)^2 + \left(1 + \frac{\phi}{2}\right)\left(\frac{x}{L}\right)\right\} \\ -\frac{1}{(1+\phi)}\left\{2\left(\frac{x}{L}\right)^3 - 3\left(\frac{x}{L}\right)^2 - \phi\left(\frac{x}{L}\right)\right\} \\ \frac{L}{(1+\phi)}\left\{\left(\frac{x}{L}\right)^3 - \left(1 - \frac{\phi}{2}\right)\left(\frac{x}{L}\right)^2 - \frac{\phi}{2}\left(\frac{x}{L}\right)\right\} \end{bmatrix}, \tag{C.6}$$

$$\left[\boldsymbol{N}_{\mathrm{r,s}}^{\theta}\right]^T = \begin{bmatrix} \frac{6}{(1+\phi)L}\left\{\left(\frac{x}{L}\right)^2 - \left(\frac{x}{L}\right)\right\} \\ \frac{1}{(1+\phi)}\left\{3\left(\frac{x}{L}\right)^2 - (4+\phi)\left(\frac{x}{L}\right) + (1+\phi)\right\} \\ -\frac{6}{(1+\phi)L}\left\{\left(\frac{x}{L}\right)^2 - \left(\frac{x}{L}\right)\right\} \\ \frac{1}{(1+\phi)}\left\{3\left(\frac{x}{L}\right)^2 - (2-\phi)\left(\frac{x}{L}\right)\right\} \end{bmatrix}. \tag{C.7}$$

Spring and damper element for ballast and railpad

Ballast and railpads are modelled as discrete spring-damper pairs. The element stiffness and damping matrices are

$$\boldsymbol{K}_{\mathrm{p,b}}^{\mathrm{e}} = k_{\mathrm{p,b}} \begin{bmatrix} 1 & 0 & -1 & 0 \\ & 0 & 0 & 0 \\ & & 1 & 0 \\ \text{symmetric} & & & 0 \end{bmatrix}, \tag{C.8}$$

$$\boldsymbol{C}_{\mathrm{p,b}}^{\mathrm{e}} = c_{\mathrm{p,b}} \begin{bmatrix} 1 & 0 & -1 & 0 \\ & 0 & 0 & 0 \\ & & 1 & 0 \\ \text{symmetric} & & & 0 \end{bmatrix}. \tag{C.9}$$

where $k_{\mathrm{p,b}}$ and $c_{\mathrm{p,b}}$ are the stiffness and damping values for railpads and ballast.

Appendix D. Parameter space for the VTI model

The track parameters and value ranges used in the VTI model are listed in Table D.1. The value ranges of the six uncertain track parameters defined in Eq. (5) are determined based on the nominal/design values of the track components, as well as the parameter values that represent potential component degradations. In addition, some parameters were determined by calibrating the model to match the measurement results (see Section 3.1). Some considerations for defining the value ranges are as follows:

- The value ranges for the railpad and ballast stiffness were wider than the design values, representing potential component degradation. For example, the lower bound of the railpad stiffness was 10 MN/m, which was much lower than the design value. This accounts for the possibility of stiffness loss due to a lost railpad (see Fig. 1(e)).
- The value range for ballast damping was determined to be narrower than the design values for the real tracks by calibrating the measurements in the V-track. This is because the V-track used rubber pads instead of real ballast stones to simulate the ballast layer.
- Track stiffness evaluations should be robust against varying track irregularities. Therefore, the track irregularity levels A_{irr} defined in Eq. (B.6) are also considered uncertain parameters. The value range of A_{irr} was determined by calibrating the simulated PSD with the measured PSD, as described in Section 3.1. For the convenience of interpretation, A_{irr} was normalized between 1 and 5 to represent different track irregularity levels, with 5 being the worst.
- Some modelling errors can also be considered based on uncertain input parameters. First, in the fastening system, only the railpad was modelled while neglecting other components, such as clamps and bolts. The stiffness and damping of these neglected components were reflected in the stiffness and damping of the railpad. For example, a loose bolt reduces the clamping force, which is reflected in the reduced railpad stiffness. Furthermore, the sleeper bolts (Fig. 1(f)) were not modelled. To account for this modelling error, the upper bound of the sleeper bending stiffness was increased by approximately three times the design value. This is because the bending stiffness of the sleepers increased when the bolts connected the sleepers to the substructure. The increased values were determined by calibrating the measurements.

The Sobol sequence sampling technique [37] was used to obtain 10,000 sets of space-filling input parameters for the VTI model. Sobol sequences are quasi-random sequences that can generate input values that are more evenly distributed throughout the input space than random or other quasi-random sampling techniques, such as Latin hypercube sampling, particularly in high-dimensional space [52]. Therefore, using Sobol sequences can help reduce the variance in the training dataset and improve the accuracy of the machine learning model. In addition, compared with sampling each parameter at a time, this sampling technique leads to simultaneous variations in different track parameters between different samples.

References

[1] J.C.O. Nielsen, E.G. Berggren, A. Hammar, F. Jansson, R. Bolmsvik, Degradation of railway track geometry – Correlation between track stiffness gradient and differential settlement, Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. 234 (2020) 108–119, https://doi.org/10.1177/0954409718819581.
[2] I. Grossoni, A.R. Andrade, Y. Bezin, S. Neves, The role of track stiffness and its spatial variability on long-term track quality deterioration, Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. 233 (1) (2019) 16–32.
[3] L. Le Pen, D. Milne, D. Thompson, W. Powrie, Evaluating railway track support stiffness from trackside measurements in the absence of wheel load data, Can. Geotech. J. 53 (2016) 1156–1166, https://doi.org/10.1139/cgj-2015-0268.
[4] H.F. Lam, S.A. Alabi, J.H. Yang, Identification of rail-sleeper-ballast system through time-domain Markov chain Monte Carlo-based Bayesian approach, Eng. Struct. 140 (2017) 421–436, https://doi.org/10.1016/j.engstruct.2017.03.001.
[5] A. Remennikov, S. Kaewunruen, Experimental investigation on dynamic railway sleeper/ballast interaction, Exp. Mech. 46 (2006) 57–66, https://doi.org/10.1007/s11340-006-5868-z.
[6] E. Arlaud, S. Costa D'Aguiar, E. Balmes, Receptance of railway tracks at low frequency: Numerical and experimental approaches, Transp. Geotech. 9 (2016) 1–16, https://doi.org/10.1016/J.TRGEO.2016.06.003.
[7] C. Shen, R. Dollevoet, Z. Li, Fast and robust identification of railway track stiffness from simple field measurement, Mech. Syst. Signal Process. 152 (2021), 107431, https://doi.org/10.1016/j.ymssp.2020.107431.
[8] J.S. Theyssen, E. Aggestam, S. Zhu, J.C.O. Nielsen, A. Pieringer, W. Kropp, W. Zhai, Calibration and validation of the dynamic response of two slab track models using data from a full-scale test rig, Eng. Struct. 234 (2021), 111980, https://doi.org/10.1016/j.engstruct.2021.111980.
[9] M.P.N. Burrow, A.H.C. Chan, A. Shein, Deflectometer-based analysis of ballasted railway tracks, Proc. Inst. Civ. Eng. Geotech. Eng. 160 (3) (2007) 169–177.
[10] C. Shen, Z. Li, R. Dollovoet, A Novel Method for Railway Crossing Monitoring Based on Ambient Vibration Caused by Train-Track Interaction, in: Lect. Notes Mech. Eng., 2020: pp. 133–141. https://doi.org/10.1007/978-3-030-38077-9_16.
[11] J.A. Priest, W. Powrie, Determination of Dynamic Track Modulus from Measurement of Track Velocity during Train Passage, J. Geotech. Geoenvironmental Eng. 135 (2009) 1732–1740, https://doi.org/10.1061/(ASCE)GT.1943-5606.0000130.
[12] D. Milne, J. Harkness, L. Le Pen, W. Powrie, The influence of variation in track level and support system stiffness over longer lengths of track for track performance and vehicle track interaction, Veh. Syst. Dyn. 59 (2021) 245–268, https://doi.org/10.1080/00423114.2019.1677920.
[13] E.G. Berggren, A. Nissen, B.S. Paulsson, Track deflection and stiffness measurements from a track recording car, Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. 228 (2014) 570–580, https://doi.org/10.1177/0954409714529267.
[14] E.G. Berggren, A.M. Kaynia, B. Dehlbom, Identification of substructure properties of railway tracks by dynamic stiffness measurements and simulations, J. Sound Vib. 329 (2010) 3999–4016, https://doi.org/10.1016/J.JSV.2010.04.015.
[15] M. Fermér, J.C.O. Nielsen, Vertical interaction between train and track with soft and stiff railpads—full-scale experiments and theory, Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. 209 (1995) 39–47, https://doi.org/10.1243/PIME_PROC_1995_209_253_02.
[16] P. Quirke, D. Cantero, E.J. OBrien, C. Bowe, Drive-by detection of railway track stiffness variation using in-service vehicles, Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit. 231 (4) (2017) 498–514.
[17] A. Malekjafarian, E.J. OBrien, P. Quirke, D. Cantero, F. Golpayegani, Railway Track Loss-of-Stiffness Detection Using Bogie Filtered Displacement Data Measured on a Passing Train, Infrastructures. 6 (6) (2021) 93.

[18] Y.B. Yang, Z.L. Wang, K. Shi, H. Xu, X.Q. Mo, Y.T. Wu, Two-axle test vehicle for damage detection for railway tracks modeled as simply supported beams with elastic foundation, Eng. Struct. 219 (2020), 110908, https://doi.org/10.1016/j.engstruct.2020.110908.

[19] L. Wang, Y. Zhang, S.T. Lie, Detection of damaged supports under railway track based on frequency shift, J. Sound Vib. 392 (2017) 142–153, https://doi.org/10.1016/j.jsv.2016.11.018.

[20] X.Q. Zhu, S.S. Law, L. Huang, Identification of Railway Ballasted Track Systems from Dynamic Responses of In-Service Trains, J. Aerosp. Eng. 31 (2018) 04018060, https://doi.org/10.1061/(asce)as.1943-5525.0000898.

[21] Y.B. Yang, Z.L. Wang, B.Q. Wang, H. Xu, Track modulus detection by vehicle scanning method, Acta Mech. 231 (2020) 2955–2978, https://doi.org/10.1007/s00707-020-02684-w.

[22] M. Molodova, Z. Li, A. Nunez, R. Dollevoet, Automatic Detection of Squats in Railway Infrastructure, Ieee Trans. Intell. Transp. Syst. 15 (5) (2014) 1980–1990.

[23] M. Bocciolone, A. Caprioli, A. Cigada, A. Collina, A measurement system for quick rail inspection and effective track maintenance strategy, Mech. Syst. Signal Process. 21 (3) (2007) 1242–1254.

[24] D. Cantero, B. Basu, Railway infrastructure damage detection using wavelet transformed acceleration response of traversing vehicle, Struct. Control Heal. Monit. 22 (2015) 62–70, https://doi.org/10.1002/stc.1660.

[25] D.R.M. Milne, L.M. Le Pen, D.J. Thompson, W. Powrie, Properties of train load frequencies and their applications, J. Sound Vib. 397 (2017) 123–140, https://doi.org/10.1016/j.jsv.2017.03.006.

[26] Y. Zhang, L. Wang, H. Zhao, S.T. Lie, Detection of Damaged Supports under Railway Track Using Dynamic Response of a Passing Vehicle, Int. J. Struct. Stab. Dyn. 19 (10) (2019) 1950117.

[27] T.G. Ritto, F.A. Rochinha, Digital twin, physics-based model, and machine learning applied to damage detection in structures, Mech. Syst. Signal Process. 155 (2021), 107614, https://doi.org/10.1016/J.YMSSP.2021.107614.

[28] S. Chakraborty, S. Adhikari, Machine learning based digital twin for dynamical systems with multiple time-scales, Comput. Struct. 243 (2021), 106410, https://doi.org/10.1016/j.compstruc.2020.106410.

[29] M.G. Kapteyn, Mathematical and Computational Foundations to Enable Predictive Digital Twins at Scale, Massachusetts Institute of Technology, 2021.

[30] P. Zhang, J. Moraal, Z. Li, Design, calibration and validation of a wheel-rail contact force measurement system in V-Track, Measurement. 175 (2021), 109105, https://doi.org/10.1016/j.measurement.2021.109105.

[31] I. Daubechies, J. Lu, H.-T. Wu, Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool, Appl. Comput. Harmon. Anal. 30 (2) (2011) 243–261.

[32] G. Thakur, E. Brevdo, N.S. Fučkar, H.-T. Wu, The Synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications, Signal Processing. 93 (5) (2013) 1079–1094.

[33] C. Shen, X. Deng, Z. Wei, R. Dollevoet, A. Zoeteman, Z. Li, Comparisons between beam and continuum models for modelling wheel-rail impact at a singular rail surface defect, Int. J. Mech. Sci. 198 (2021), 106400, https://doi.org/10.1016/j.ijmecsci.2021.106400.

[34] Z. Friedman, J.B. Kosmatka, An improved two-node timoshenko beam finite element, Comput. Struct. 47 (1993) 473–481, https://doi.org/10.1016/0045-7949(93)90243-7.

[35] P. Zhang, Z. Li, Experimental study on the development mechanism of short pitch corrugation using a downscale V-Track test rig, Tribol. Int. 180 (2023), 108293, https://doi.org/10.1016/j.triboint.2023.108293.

[36] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning (2005), https://doi.org/10.7551/mitpress/3206.001.0001.

[37] A. Marrel, B. Iooss, B. Laurent, O. Roustant, Calculations of Sobol indices for the Gaussian process metamodel, Reliab. Eng. Syst. Saf. 94 (2009) 742–751, https://doi.org/10.1016/J.RESS.2008.07.008.

[38] P.H. Franses, A note on the Mean Absolute Scaled Error, Int. J. Forecast. 32 (2016) 20–22, https://doi.org/10.1016/J.IJFORECAST.2015.03.008.

[39] F. Peng, D. Yu, J. Luo, Sparse signal decomposition method based on multi-scale chirplet and its application to the fault diagnosis of gearboxes, Mech. Syst. Signal Process. 25 (2011) 549–557, https://doi.org/10.1016/J.YMSSP.2010.06.004.

[40] W. Fan, G. Cai, Z.K. Zhu, C. Shen, W. Huang, L. Shang, Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction, Mech. Syst. Signal Process. 56 (2015) 230–245, https://doi.org/10.1016/j.ymssp.2014.10.016.

[41] I.W. Selesnick, Resonance-based signal decomposition: A new sparsity-enabled signal analysis method, Signal Processing. 91 (2011) 2793–2809, https://doi.org/10.1016/j.sigpro.2010.10.018.

[42] T.X. Wu, D.J. Thompson, VIBRATION ANALYSIS OF RAILWAY TRACK WITH MULTIPLE WHEELS ON THE RAIL, J. Sound Vib. 239 (2001) 69–97, https://doi.org/10.1006/jsvi.2000.3157.

[43] T.P. Le, P. Argoul, Continuous wavelet transform for modal identification using free decay response, J. Sound Vib. 277 (2004) 73–100, https://doi.org/10.1016/j.jsv.2003.08.049.

[44] W.J. Staszewski, Identification of non-linear systems using multi-scale ridges and skeletons of the wavelet transform, J. Sound Vib. 214 (1998) 639–658, https://doi.org/10.1006/jsvi.1998.1616.

[45] M. Ruzzene, A. Fasana, L. Garibaldi, B. Piombo, Natural Frequencies and Dampings Transform: Application To Real Data, Mech. Syst. Signal Process. 11 (1997) 207–218, https://doi.org/10.1006/mssp.1996.0078.

[46] K. Dziedziech, W.J. Staszewski, T. Uhl, Wavelet-based modal analysis for time-variant systems, Mech. Syst. Signal Process. 50–51 (2015) 323–337, https://doi.org/10.1016/J.YMSSP.2014.05.003.

[47] Y. Guo, A. Kareem, Non-stationary frequency domain system identification using time–frequency representations, Mech. Syst. Signal Process. 72–73 (2016) 712–726, https://doi.org/10.1016/J.YMSSP.2015.10.031.

[48] G. Yue, Z. Xu, L. Wang, C. Liu, T. Ren, WSN-Based Vibration Characteristic Research for Various Railway Track Structures for Pattern Classification, Int. J. Pattern Recognit. Artif. Intell. 30 (2016) 1650020, https://doi.org/10.1142/S0218001416500208.

[49] P. Salvador, V. Naranjo, R. Insa, P. Teixeira, Axlebox accelerations: Their acquisition and time–frequency characterisation for railway track monitoring purposes, Measurement. 82 (2016) 301–312, https://doi.org/10.1016/j.measurement.2016.01.012.

[50] Z. Wei, X. Sun, F. Yang, Z. Ke, T. Lu, P. Zhang, C. Shen, Carriage interior noise-based inspection for rail corrugation on high-speed railway track, Appl. Acoust. 196 (2022), 108881, https://doi.org/10.1016/j.apacoust.2022.108881.

[51] X. Lei, N.-A. Noda, ANALYSES OF DYNAMIC RESPONSE OF VEHICLE AND TRACK COUPLING SYSTEM WITH RANDOM IRREGULARITY OF TRACK VERTICAL PROFILE, J. Sound Vib. 258 (2002) 147–165, https://doi.org/10.1006/jsvi.2002.5107.

[52] S. Kucherenko, D. Albrecht, A. Saltelli, Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling, Techniques (2015) 1–30. http://arxiv.org/abs/1505.02350.