



Auralisation of train pass-bys for virtual reality demonstration of combined noise mitigation measures

Reto Pieren^{a,*,id}, Fotis Georgiou^{a,id}, Giacomo Squicciarini^{b,id}, David J. Thompson^{b,id}

^a Empa, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland

^b ISVR, University of Southampton, Southampton, UK

ARTICLE INFO

Keywords:

Auralisation
Railway noise
Noise mitigation
Virtual reality
Sound synthesis

ABSTRACT

Railway noise can be reduced by various mitigation measures and combinations thereof. However, it is difficult to assess their combined effects and to communicate the options to stakeholders. For this, auralisation is a promising technique that can support communication and decision-making, and enable psychoacoustic evaluations. This paper presents a new physics-based auralisation model for train pass-bys that considers combined mitigation measures. The measures include acoustic rail grinding, avoidance of wheel flats, wheel and rail dampers, rail shields, mini barriers and classical noise barriers. Pass-by parameters such as train type, speed, track design and observer location can be selected. The proposed model includes contributions from rolling noise, impact noise, traction, auxiliary systems, and aerodynamic noise. The main novelty of this work lies in the improved time-domain synthesis of rolling noise. The sound radiated by each wheelset is modelled by multiple moving point sources. For the sound radiated by the track, a new hybrid model is proposed that consists of a combination of moving and fixed equivalent sources, reflecting the structural wave propagation in the rails. Separate source contributions for radial and axial wheel vibration, vertical and lateral rail vibration and sleeper vibration are considered using TWINS-based computations and an improved description of rolling damping. First comparisons of synthesised and recorded train pass-bys showed a very good agreement and a high degree of realism. The auralisations were coupled to an immersive virtual reality environment that allows for an interactive audio-visual experience of different train pass-by scenarios and to demonstrate noise mitigation options. The presented models were implemented in two software tools that are described in this paper and made available. The tools have already been successfully used in public demonstrations at international exhibitions and information events for residents.

1. Introduction

Noise pollution is an important environmental issue that adversely affects human health and well-being. Chronic exposure to elevated noise levels has been linked to a range of health problems, including sleep disturbances, cardiovascular diseases, and cognitive impairments [1]. The World Health Organization estimates that at least 20% of the European population lives in areas where noise levels are harmful to health [2].

Railway noise, as a significant contributor to environmental noise pollution, poses specific challenges. Noise from both mechanical and aerodynamic sources can disrupt communities situated near railway lines. Studies have shown that individuals residing close to railways experience increased annoyance, sleep disturbances, and a higher risk of cardiovascular issues [3,4].

Mitigating railway noise is crucial for reducing its negative impact on public health and enhancing the quality of life for affected populations. Implementing effective noise control measures, such as noise barriers, improved track maintenance, and the development of quieter train technologies, is essential. Additionally, comprehensive urban planning and policy interventions are necessary to manage and reduce noise exposure in communities adjacent to railway corridors.

Railway noise is conventionally assessed using standardised noise indicators such as A-weighted equivalent continuous sound level (L_{Aeq}), maximum sound level (L_{max}), and day-evening-night level (L_{den}) as well as the corresponding spectra. These metrics provide a quantitative evaluation of noise exposure and are widely used in environmental impact assessments and regulatory frameworks. However, such indicators often fail to fully capture the dynamic, transient, and perceptual

* Corresponding author.

E-mail address: reto.pieren@empa.ch (R. Pieren).

<https://doi.org/10.1016/j.apacoust.2025.111063>

Received 18 June 2025; Received in revised form 19 August 2025; Accepted 4 September 2025

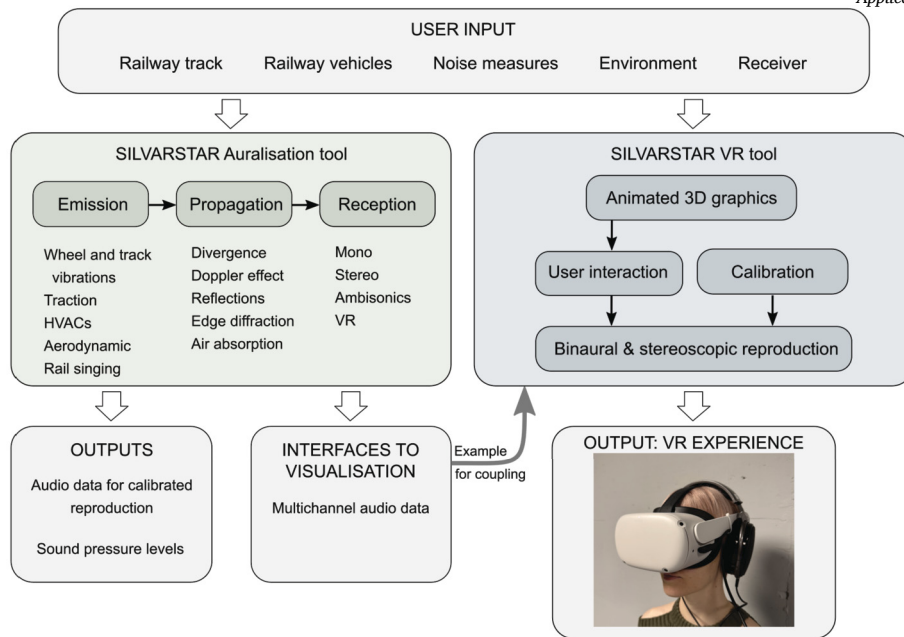


Fig. 1. Software architecture of SILVARSTAR VR demonstrator consisting of the auralisation tool that is coupled to the VR tool.

characteristics of railway noise, which significantly influence human annoyance and discomfort. A recent study explored the possibility of using psychoacoustic metrics to assess tram noise [5].

Auralisation, in contrast, offers a more immersive and perceptually relevant approach to railway noise assessment. By allowing listeners to experience simulated railway noise in realistic acoustic environments, auralisation provides insights that go beyond numerical noise levels. This enables researchers and engineers to assess the subjective effects of noise, such as tonal components, modulation, and impulsiveness—elements that conventional noise indicators may overlook.

Moreover, auralisation enhances public engagement and participatory decision-making in noise mitigation strategies. By enabling stakeholders to hear and compare different noise control measures, auralisation facilitates a more transparent and informed approach to railway noise management.

Several studies have explored the auralisation of railway noise. Most of the published work focusses on rolling noise, with only a few papers also addressing impact noise and other sources. In [6] a method that models the excitation of the wheel/rail system due to surface roughness in the time domain was proposed, aiming to achieve a realistic sound representation of rolling and impact noise. This approach was further extended by several research groups. Maillard et al. [7] proposed an auralisation model that represents the track by fixed distributed sources. In [8,9] Theyssen et al. developed a time-domain prediction model to estimate the sound pressure generated by a passing railway wheel, emphasising the sensitivity of modal contributions to the lateral contact position, which plays a crucial role in the accurate auralisation of rolling noise. And recently, Knuth et al. [10] used a numerical modelling approach to synthesise rolling noise.

Combining auralisation with virtual reality (VR) enhances railway noise assessment by creating fully immersive and interactive audio-visual experiences [11]. This integration allows users to hear as well as see how different changes impact their surroundings, improving the realism and accuracy of sound perception studies.

Using auralisation and VR to assess mitigation measures in railway noise was the main goal of this research conducted within the European research project SILVARSTAR [12]. To be able to auralise many different scenarios, we developed a new physics-based auralisation model for train pass-bys that can simulate combined mitigation measures. The auralisations were coupled to an immersive VR environment that allows

for an interactive audio-visual experience of different train pass-by scenarios and to demonstrate noise mitigation options. The models were implemented in software tools which are made available. The auralisation model as well as the software tools are described in this paper.

The work presented in [6,11] serves as a foundation for many aspects of this research, introducing a synthesis concept for rolling and impact noise based on a physical modelling approach in the time domain. The main novelties of this work are an improved modelling of wheel vibration damping, and a new efficient approach to model the sound radiated from the track using a combination of static and moving equivalent sources. These works, along with their technical details, are discussed later in this paper.

The remainder of this paper is structured as follows: Section 2 provides an overview of the new software tools, along with details of the various parameters that can be simulated, without going into technical details. Section 3 explains the technical aspects of the auralisation model, highlighting the key research contributions of this work. Section 4 explains the used sound reproduction strategies. Finally, Section 5 presents comparisons between pass-by measurements and synthesised results, as well as initial findings from a listening test.

2. The SILVARSTAR auralisation and VR tools

2.1. Overview

In the European project SILVARSTAR [12], new tools for the demonstration of railway noise mitigation measures were developed. The tools are mainly meant for demonstration purposes, comparisons between variants, in communication and decision-making. Fig. 1 presents the general software architecture of the VR demonstrator, which consists of two interconnected software tools: the auralisation tool and the VR tool. These software tools are designed to work in combination, where input parameters related to the acoustic environment and noise mitigation measures are processed and utilised by both systems.

The auralisation tool is responsible for generating auralised train pass-by sounds based on user-specified input parameters. The sounds are generated by synthesising virtual sound sources that move in a virtual 3D environment. The auralisation tool is written in MATLAB and the used auralisation models are described in Section 3. The output of the software includes sound pressure levels and calibrated audio data

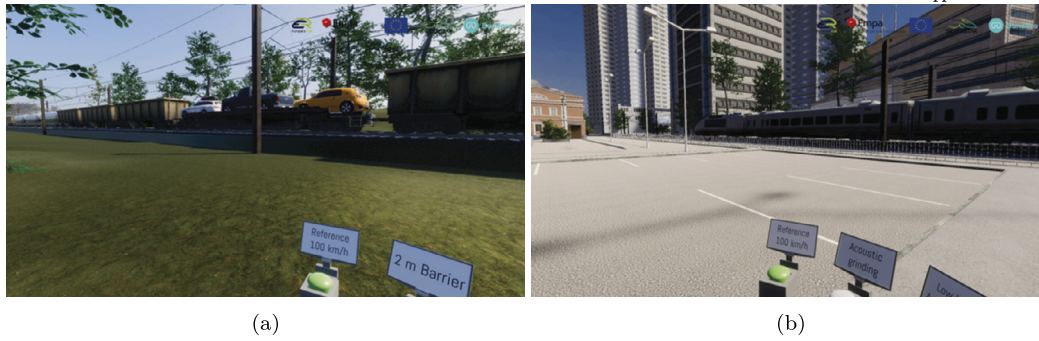


Fig. 2. Examples of the user view of the SILVARSTAR VR demonstrator with train passages in different environments and the virtual buttons (bottom right) for activating different mitigation measures in real time. (a) Freight train in rural setting. (b) Intercity train in urban setting.

Table 1

Input parameters related to vehicles and their operation.

Parameter	Values
Train type	Regional, Intercity, Freight
Train length	Short, Long
Speed	80–200 km/h for passenger 60–100 km/h for freight
Driving direction	RL, LR (LR = left to right)
Used track	Close, Far
Start position	50–200 m

Table 2

Input parameters related to track, environment and observer.

Parameter	Values
Track type	Ballasted, Slab
Sleeper type	Monoblock, Biblock
Rail pad type	Soft, Medium, Hard
Embankment height	0–4 m
Environment	Rural, Urban
Observer distance	7.5–50 m
Observer height	1.65–10 m

tailored for different reproduction systems including multichannel audio data, which serve as input for the VR tool. This audio data is then converted to binaural format within the VR tool for spatialised sound reproduction over headphones.

The VR tool, developed with the game engine Unity, creates a realistic 3D virtual environment, which is displayed through a head-mounted display (HMD). It utilises user-defined input parameters to generate the virtual environment and animations, which are then synchronised with multichannel audio data. The sequence of virtual scenes is configurable via an input file (.json), allowing for flexible simulation scenarios.

The corresponding binaural audio is played through calibrated headphones, enhancing the immersive audio-visual experience. Additionally, the system features a graphical user interface (GUI), enabling an operator to guide users through various interactive virtual scenes. Users can instantly switch between different scenarios either by interacting with virtual buttons (Fig. 2) or by pressing the motion controller buttons. Two examples of the user's view in different environments and with different train types are shown in Fig. 2.

2.2. Creation of scenarios

With the presented system, a large variety of different scenarios can be created by the user, in contrast to the previous prototype demonstrator [11]. A scenario is characterised with a list of input parameters describing an outdoor train pass-by situation. Different types of rolling stock, operation, tracks, environments and observer locations can be chosen, as well as different noise mitigation measures. A total of six different train types, seven tracks and nine mitigation measures can be selected. The physics-based simulation approach described in Section 3 allows different mitigation measures to be combined. Some continuous parameters such as the travelling speed, the observer distance, or the barrier height can be selected within a given value range—offering a large number of possible scenarios. A total of 21 input parameters have to be selected to define and create a scenario. These parameters are described in the following.

Table 1 lists the input parameters related to the vehicles and their operation. Three different train types (Regional, Intercity, Freight) can be selected, each in two lengths. The trains operate at constant speed.

Table 3

Input parameters related to mitigation measures.

Parameter	Values
Rail roughness	Poor, Medium, Smooth
Freight wagon brake block CI to K rate	0, 50, 100%
Wheel flats	Yes, No
Wheel dampers	No, Yes
Rail treatment	No, Dampers, Shields
Attenuation of secondary sources	0–99 dB
Barrier type	No, Standard, Low
Standard barrier height	2–5 m

For a pass-by, the train speed can be continuously set within a given range, 80–200 km/h for the passenger, and 60–100 km/h for the freight trains, which reflects typical speeds and is within the range where the model implementations are valid. The track used, the driving direction and the vehicle starting position along the track have to be chosen to describe fully the train operation.

A straight double track is assumed with a geometry as illustrated in Fig. 3. Table 2 lists the input parameters related to the track, the environment and the observer. A total of seven tracks are implemented. The user can select a slab track or a ballasted track. For the latter, monoblock or biblock concrete sleepers and three classes of rail pads with different stiffnesses can be selected. The track can be situated at-level or on an embankment, and either in a rural or urban environment. The observer is at a fixed location and outdoors. The horizontal distance from the observer to the centre of the nearest track is limited to 7.5–50 m. The observer is either standing on the ground or at an elevated position, with the ear level being between 1.65 and 10 m above the terrain. The environment parameter determines the ground properties where for the urban environment asphalt and for the rural one grass is assumed.

A total of nine different noise mitigation measures are considered and implemented: 1) Rail grinding, 2) Composite brake blocks on freight wagons, 3) Avoidance of wheel flats, 4) Wheel dampers, 5) Rail dampers, 6) Rail shields, 7) Low height barriers, 8) Noise barriers, 9) Secondary source attenuation. Table 3 lists the user input parameters related to these noise mitigation solutions. Four of them are related to the vehicles, and five to the track or environment. Six measures only affect

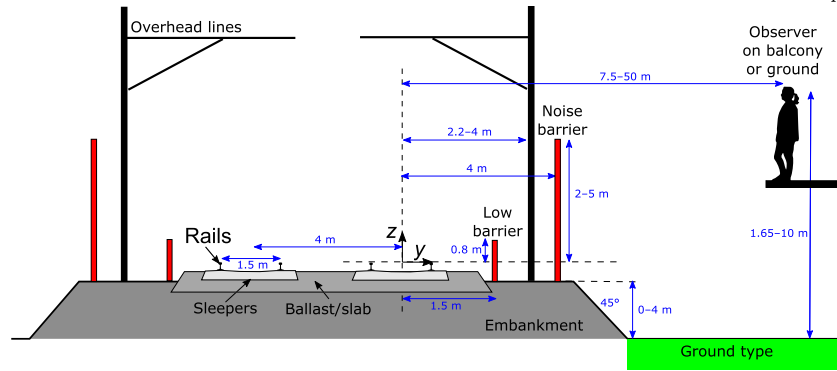


Fig. 3. Illustration of the simulation geometry as a cross-section of the double track on an embankment with possible noise barriers (red) and the virtual observer. Indications of dimensions and their allowed ranges are shown in blue.

Table 4

Descriptions of virtual test trains of the SILVARSTAR auralisation tool (loco = locomotive).

ID	Type	Size	Composition	Length	#Axles
1	Regional	Short	5 coaches	90 m	12
2	Regional	Long	10 coaches	180 m	24
3	Intercity	Short	2 loco + 6 coaches	200 m	32
4	Intercity	Long	4 loco + 12 coaches	400 m	64
5	Freight	Short	1 loco + 16 wagons	300 m	72
6	Freight	Long	1 loco + 21 wagons	550 m	110

rolling noise. The first three measures are related to the roughness excitation: 1) Acoustic rail grinding can be simulated by setting the rail roughness to ‘Smooth’. 2) The freight train wagons can be partially or fully retrofitted by composite K brake blocks. 3) Wheel flats can be omitted. Next, the wheels and the rails can separately or both be equipped with dampers. The rails can alternatively be equipped with rail shields. Secondary sources from equipment, traction and aerodynamic noise sources can be globally attenuated. Finally, either low height barriers with a fixed geometry, or conventional sound-absorbing noise barriers of height 2–5 m can be enabled.

2.3. Train type characteristics

This section describes the basic characteristics of the trains as they are implemented in the auralisation tool. All trains are propelled with electric locomotives. Table 4 lists the six trains implemented in the tool. There are three types of trains: Regional, Intercity and Freight. Each train type is available in two sizes, i.e. short and long. The regional trains are electric multiple unit (EMU) trains with articulated bogies and either five or ten coaches. They have two axles per bogie. The intermediate bogies of each EMU are Jacob’s bogies, i.e. a single bogie at the junction of two wagons. The short intercity train has a front and rear locomotive with six middle cars, and is duplicated for its long version. The intercity trains have two axles per bogie and two bogies per wagon, i.e. four axles per wagon. The freight trains have one front locomotive and either 16 or 21 wagons. Both trains have the same locomotive with four axles. The short freight train (ID 5) has wagons with four axles. The long freight train (ID 6) has 11 wagons with six axles and 10 wagons with four axles. The longest virtual train in Table 4 is the long freight train (ID 6) with 550 m length and 110 axles. The lengths, numbers of axles and the compositions of the intercity and freight trains were chosen based on statistics. For this, large datasets from the MineTrack project [13] of international train pass-bys operating on the Swiss network were analysed. For instance, freight trains were found to most often have a length of 550 m, one locomotive and 20 wagons.

The wheel roughness spectra used as an input to the auralisation are shown in Fig. 4. The spectra for cast iron, composite and disc brakes are derived from a measurement database from the ACOUTRAIN project

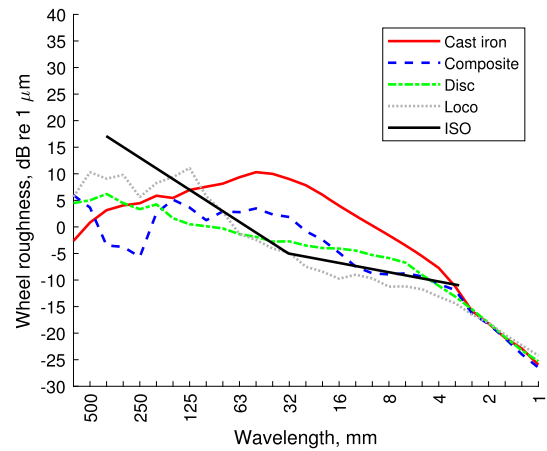


Fig. 4. Wheel roughness spectra in one-third octave bands used in the SILVARSTAR auralisation tool for freight wagons with optionally cast iron or composite brake blocks, for the passenger trains with disc brakes and for the freight train locomotive (‘loco’).

[14], which includes 52 cast iron block braked-wheels, 84 composite block braked-wheels and 138 disc-braked wheels. The representative roughness curves correspond to energetic averages. The limit curve for rail roughness from ISO 3095:2013 [15] is shown for reference (identified as ‘ISO’). The regional and intercity trains are equipped with disc brakes. The freight wagons are equipped with either cast iron or composite brake blocks depending on the setting of the input parameter describing the proportion of retrofitting. Potential wheel flats are characterised by equivalent roughness spectra for different speeds as displayed in Fig. 5. The spectra representing a worn flat are based on the model by Wu & Thompson [16]. For increasing speeds, the equivalent roughness levels become smaller due to the wheel briefly losing contact with the rail. The sonRAIL measurement campaign [17] has shown that about 2% of the wheels are affected by a flat. Therefore by default in the auralised trains, two wheels with flats are modelled for the freight trains, and one wheel flat for the passenger trains. However, the occurrence of wheel flats can be turned off by the user in the sense of a virtual noise mitigation measure.

2.4. Track characteristics

This section describes the basic characteristics of the tracks as implemented in the auralisation tool. As illustrated in Fig. 3, a double track is assumed, which may be on an embankment. The slope of the embankment is 45°. The total height of the rail and track bed is 0.35 m. Although the tool is intended to generate aural impressions for a virtual standing person, still a receiver at the standard measurement point at

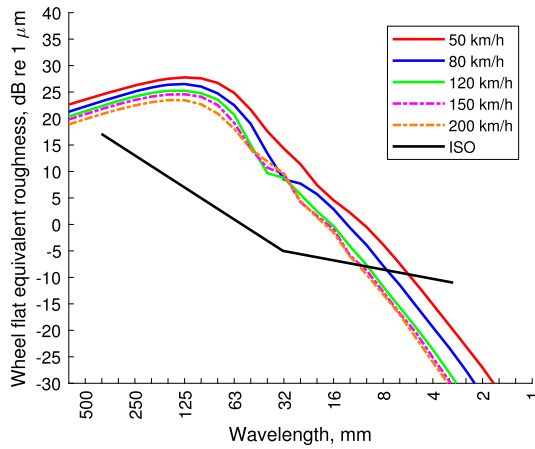


Fig. 5. Equivalent wheel roughness spectra in one-third octave bands used in the SILVARSTAR auralisation tool for wheel flats shown for different travelling speeds.

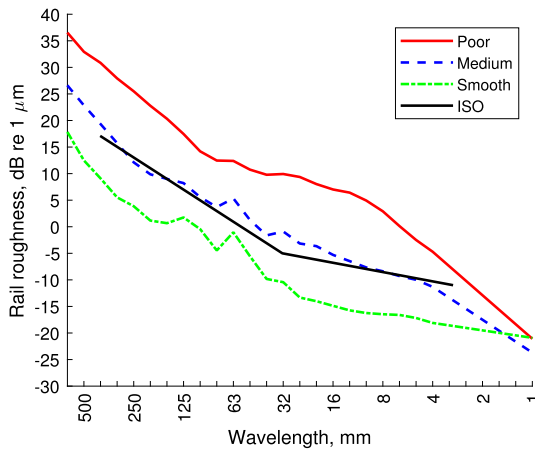


Fig. 6. Rail roughness spectra in one-third octave bands used in the SILVARSTAR auralisation tool for three selectable rail quality classes.

1.2 m above rail head can, for example, be simulated by choosing the lowest allowed observer height of 1.65 m above ground and setting the embankment height to 0.1 m.

The rail roughness is assumed to be identical for the four rails and in a statistical sense independent on the location along the rail. As an input to the auralisation, rail roughness is given in one-third octave bands in the wavelength range 1000 to 1 mm. No corrugation, rail joints or switches are considered here. The roughness spectra used are shown in Fig. 6; the limit curve from ISO 3095:2013 [15] is again shown for reference. The spectral values of the rail roughness at 1000 mm and 800 mm follow the same rising trend as shown in Fig. 6 for wavelengths 250–630 mm. The roughness spectra are specified for the three rail quality classes: Poor, Medium and Smooth. The curves are based on an analysis of approximately 30 rail roughness measurements collected across the UK. They are introduced into the tool using their energetic average (medium), as well as the 10th and 90th percentiles.

The monoblock sleepers have a mass of 280 kg and the biblock sleepers 120 kg. The vertical pad stiffness for the soft, medium and hard pads are 100, 250 and 1000 MN/m. The effect of the ballast on sound propagation (e.g. sound transmission through the ballast or reflection at the ballast bed) is not considered in the model.

One representative slab track is modelled consisting of a concrete slab below the rails. In the model the slab is set to be rigid and does not contribute to the sound radiation. The rails are modelled as resting on a double layer foundation with a baseplate set between a lower and an

upper rubber pad. The parameters describing the slab track are taken from Zhang et al. [18].

As a possible mitigation measure, each of these seven tracks can optionally be equipped with rail dampers. These dampers are not intended to represent a specific product but are intended to show the effect of adding additional decay to the propagation of vibration along the track. The methodology developed in the STARDAMP project is followed here [19], which introduces an additional decay rate representing the dampers. Alternatively, the rails of each of the seven tracks can be encased with rail shields. Spectral insertion losses obtained from the recent study by Zhang et al. [20] are used to acoustically characterise and model the rail shields.

3. Physics-based sound synthesis model

3.1. Model overview

In the presented auralisation model, a train pass-by is generated by a collection of synthesised sound sources moving in a virtual environment. All the major sound sources are considered, and a separate emission signal is synthesised for each source. Typically, several hundred sources are used to auralise a full train pass-by.

The auralisation model consists of the three steps: Source signals are first synthesised using physics-based or signal-based sound synthesis. Propagation from the source to the receiver is then introduced by applying time-varying filters to model different propagation effects. Finally, different reproduction renderers are used to generate signals for different sound reproduction systems, such as stereo loudspeakers, Ambisonics or headphones (see Section 4.1).

Expressed formally, the sound pressure p at time t at the observer location $\mathbf{r}$ is modelled by

$$p(\mathbf{r}, t) = \sum_m \sum_n \int_{-\infty}^t s_m(\tau) h_{\text{prop},m,n}^t(\mathbf{r}, t - \tau) d\tau \quad (1)$$

with the source time signals s_m for source index m , the time-varying propagation function h_{prop}^t and index n for the propagation paths, such as reflected and diffracted paths (see Section 3.5). The propagation function h is time-dependent because of the source motion. The integral in Eq. (1) represents a time-varying convolution which is limited to t because of causality, i.e. $h_{\text{prop}}^t(\mathbf{r}, t') = 0$ for $t' < 0$. To compute audio signals for spatial sound reproduction, a time-varying reproduction rendering function g^t is introduced between the two sums in Eq. (1) to obtain the audio signal u_l per audio channel l :

$$u_l(\mathbf{r}, t) = \sum_m g_{l,m}^t(t) * \sum_n \int_{-\infty}^t s_m(\tau) h_{\text{prop},m,n}^t(\mathbf{r}, t - \tau) d\tau \quad (2)$$

where $*$ denotes the linear time-varying convolution. In Eq. (2) the reproduction rendering is applied per source but irrespective of the different propagation paths n . This computational simplification is based on the assumption that the first wavefront will dominate the perceptual localisation according to the “precedence effect” or the “law of the first wavefront” [21,22]. This is valid for the considered geometries, which have short time delays between the propagation paths.

The following sections explain how the source signals s , the propagation filters h and the reproduction filters g in Eq. (2) are created.

3.2. Synthesis of rolling and impact noise

In many railway noise situations, rolling noise is the dominant source in terms of the sound pressure level. The wheel/rail contacts may also lead to transient noise, for instance in the case of wheel flats, rail joints or switches. Rolling and impact noise are both very characteristic in the perception of railway noise.

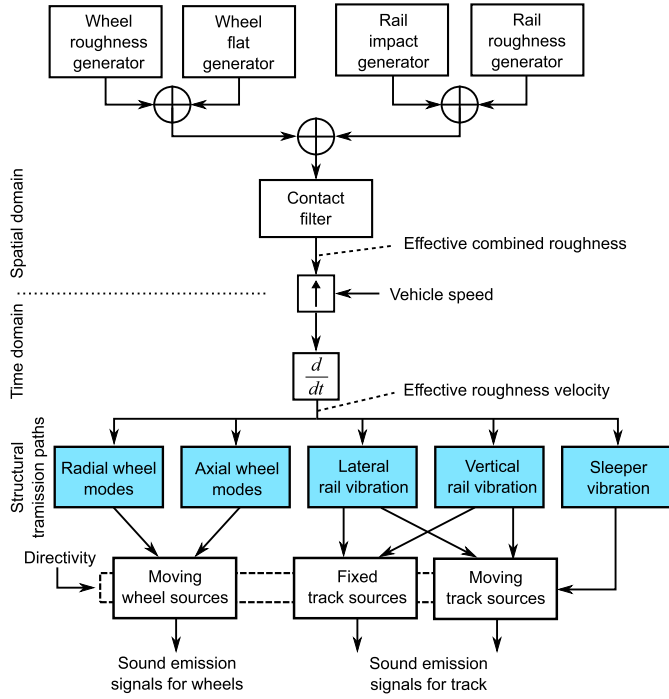


Fig. 7. Signal flow diagram of the rolling and impact noise synthesis model. Blue boxes represent the structural transmission paths. The synthesis delivers emission signals for moving wheel sources, fixed track sources, and for moving track sources.

The emission of rolling noise is simulated by a collection of distributed point sources within the virtual environment. Separate sources are introduced for the sound radiated from each wheel and for the sound radiated from the track. The corresponding source signals are created by a new physics-based synthesis method presented in the following.

A signal flow diagram of the sound synthesis of rolling and impact noise is shown in Fig. 7. For each wheelset i of the train, an effective roughness velocity time signal $v_{r,i}(t)$ at the wheel/rail contact on the near rail is computed as a time-domain signal, transformed from an initial one-third octave spectral resolution, based on an earlier approach [6]. This takes the wheel and rail roughness spectra shown in Figs. 4 and 6 as an input. The roughness spectra are converted to a narrow-band form using a piecewise cubic interpolation and transformed into spatial signals assuming random phase relations (i.e. assuming no periodic patterns due to e.g. rail grinding). For wheels with a wheel flat, an impulse train is additionally generated reproducing the equivalent roughness spectra in Fig. 5 using the method from [6], i.e. with an inverse procedure to the one used in [16]. The roughness signals are generated with a spatial sampling interval of 0.45 mm.

To synthesise realistic train pass-bys, some random variation in source strength between wheelsets of the same type is needed. This is particularly important for the simulation of freight trains consisting of different wagons. Statistical analyses of freight train pass-bys showed substantial differences between wheels (of the same brake block type). The dataset consisted of pass-by of 281 international freight trains measured in Switzerland with a total of 18'400 wheelsets. Based on these analyses, a hierarchical random variation between freight wagon wheelsets is introduced, as established in the DESTINATE project [11]. For the wheels within one wagon, a standard deviation of $\sigma = 2$ dB is used. For the wheels within similar wagons $\sigma = 2.8$ dB and across all freight wagons $\sigma = 3.5$ dB.

To obtain an effective roughness excitation signal, a contact filter is introduced, representing the finite size of the contact patch, which is based on the DPRS model [23]. In contrast to [6], here the velocity signals are explicitly calculated from the displacement signals using a

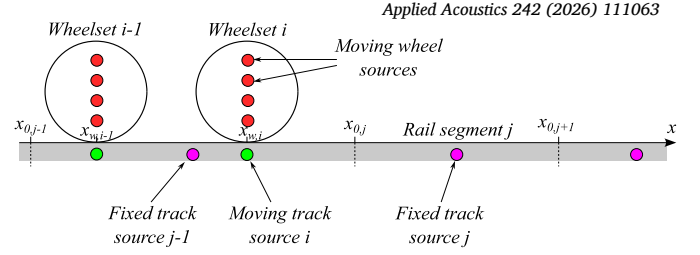


Fig. 8. Equivalent sources model for the synthesis of rolling noise. The sound radiated from the wheels is represented by four moving sources per wheelset (red). The track contribution is modelled by a combination of moving and fixed sources. A moving track source (green) is introduced below each wheel/rail contact. The fixed track sources (magenta) represent track segments of different length.

5th-order FIR differentiator. The resulting spatial signal per wheelset i is then transformed to the time domain. Using the constant train travelling speed V as an input, each spatial signal is resampled to obtain the effective roughness velocity time signal $v_{r,i}(t)$.

To account for the dynamic behaviour of the wheel/track system, the excitation signal is fed to multiple filters representing different structural transmission path filters; these are arranged in parallel as illustrated in Fig. 7. These filters are applied in the time domain, resulting in rolling noise source signals $s(t)$. The following sections describe how these source signals are computed for sources related to the wheels and the track, also considering source directivities.

3.2.1. Wheel sources

For the wheel contributions, moving sources are associated with each wheelset i of the train. These virtual sources move with the train speed V in the driving direction. Whereas earlier auralisation models use a single equivalent point source to represent a wheel [6,8], here we propose an extension to account for partial shielding from a low height barrier and to obtain a smooth spectral ground effect pattern. For each wheelset, multiple vertically stacked sources above the wheel/rail contact are used to represent the spatial extent of the wheel. As a reasonable compromise between accuracy and computational effort, a total of four equally spaced equivalent sources are introduced. The wheel sources are illustrated in Fig. 8.

The computational effort depends linearly on the number of wheelsets of the train. According to Table 4, the number of wheelsets varies strongly between the trains, by almost a factor of 10, and amounts to more than 100 wheelsets for the long freight train. To prevent excessive computation times, a simplification regarding the wheel sources was made: all wheel sources are located above the near side rail. Contributions from the far side wheel are considered in a simplified way with an artificial amplification by 3 dB.

The sound radiated from each source wheelset i is described using an equivalent source model with point sources b . The wheel source signal is modelled by the combined contributions of radial ('rad') and axial ('ax') wheel vibration

$$s_{w,i,b}(t, \varphi_i, \theta_i) = \sqrt{2} \rho_0 (\dot{Q}_{\text{rad},i,b}(t) + \dot{Q}_{\text{ax},i,b}(t) \Psi_{\text{di},i}(t)) \Psi_{\text{ts},i}(t) \quad (3)$$

with the air density ρ_0 , the volume acceleration $\dot{Q} = dQ/dt$ being the time derivative of volume velocity Q (in m^3/s) for one wheel subsource, and the directivity functions Ψ depending on the azimuthal angle φ and the elevation angle θ . The general factor of $\sqrt{2}$ approximates the sound power increase due to the far side wheel by about 3 dB. A dipole directivity Ψ_{di} is applied to the axial vibration contribution [24] and none to the radial vibration. Additionally, the directivity function Ψ_{ts} accounts for the shielding by the train body and the bogie cavity at grazing radiation angles ("installation effect"). Both directivities are modelled as independent of frequency. Furthermore, the dependence on the elevation angle θ is neglected, i.e. only horizontal directivity is explicitly considered. This simplification, together with how the far side wheel is

considered, reduce the reliability of the predictions for large absolute elevation angles such as in the case of a high noise barrier or for direct ground reflections. For a typical railway line geometry even with a 5 m noise barrier the elevation angle remains below 60° . Up to this elevation the radiation of the lateral and vertical rail vibration was found to be quite uniform [24], which justifies this simplification. The two directivities for the wheel contributions are modelled by:

$$\Psi_{di} = \sqrt{3} \cos \varphi \cos \theta \approx \sqrt{3} \cos \varphi \quad (4)$$

$$\Psi_{tbs} \approx \kappa + (1 - \kappa) |\cos \varphi| \quad (5)$$

The azimuthal angle φ is defined as 0 in the direction perpendicular to the track and limited to the range $-\pi/2$ to $\pi/2$. The dipole scaling factor $\sqrt{3}$ assumes a 3D dipole and ensures energy is conserved. A shielding coefficient of $\kappa = 0.1$ was estimated experimentally from regional train pass-by measurements. This setting limits the shielding effect to -20 dB. Note that for the computation of a source signal for pass-by auralisation, the azimuthal radiation angle φ is a function of time, i.e. $\varphi(t)$.

The volume accelerations $\dot{Q}$ in Eq. (3) are derived by linear convolution of the roughness velocity excitation signal $v_{r,i}$ with impulse responses of respective structural transmission paths:

$$\dot{Q}_{rad,i,b}(t) = (v_{r,i} * h_{rad,i,b})(t) \quad (6)$$

$$\dot{Q}_{ax,i,b}(t) = (v_{r,i} * h_{ax,i,b})(t) \quad (7)$$

For each wheelset, Eq. (3) requires eight transmission path filters h , i.e. for the radial and axial vibration for each of the four source heights. These time-domain filters are designed based on a superposition of second order oscillators as in [6] expressed in the time domain by exponentially decaying sinusoids:

$$h_{osc}(t) = \sum_{n=1}^N A_n e^{-\alpha_n(t-t_n)} \sin(2\pi f_n(t-t_n)) \Theta(t-t_n) \quad (8)$$

with amplitudes A_n , decay rates α_n , frequencies f_n , time delays t_n , and the Heaviside function Θ . The decay rate is related to the modal damping ratio ζ by

$$\alpha_n \approx 2\pi f_n \zeta_n. \quad (9)$$

Resonance frequencies of wheel modes were computed from wheel cross-sectional geometries of different wheels using the finite element method (FEM). These modal frequencies are attributed to the respective response frequencies f_n in Eq. (8) with a slight random variation of 2% between wheels of the same type within the train. A typical number of modal frequencies in Eq. (8) is $N = 50$. To correctly reproduce the narrowband behaviour in the volume velocity signals, appropriate estimates for the total damping of the wheels are needed. For a wheel rolling over a railway track, the total effective damping ratio, ζ_{tot} , can be written as the sum of the structural damping ratio, ζ_s , and the apparent rolling damping ratio [24], ζ_{roll} , the latter being due to the coupling between the wheel and the rail:

$$\zeta_{tot} = \zeta_s + \zeta_{roll} \quad (10)$$

Both ζ_s and ζ_{roll} are frequency dependent, thus giving a frequency dependent curve for ζ_{tot} . In the prior auralisation model by Pieren et al. [6] a damping ratio of $\zeta_{tot} \approx \frac{10}{f}$ was used.

As the total damping is affected by both the structural damping and by the apparent rolling damping, different wheelset designs can result in different levels of total damping. For example, it is known that for a lightly damped wheelset the total damping is mostly determined by the rolling damping [24]. On the other hand, if the structural damping increases, such as in the presence of wheel-mounted disc brakes or wheel dampers, this will contribute more to ζ_{tot} .

Three different wheelset set-ups are considered in this paper and each is given a different total damping curve. These are (i) a standard

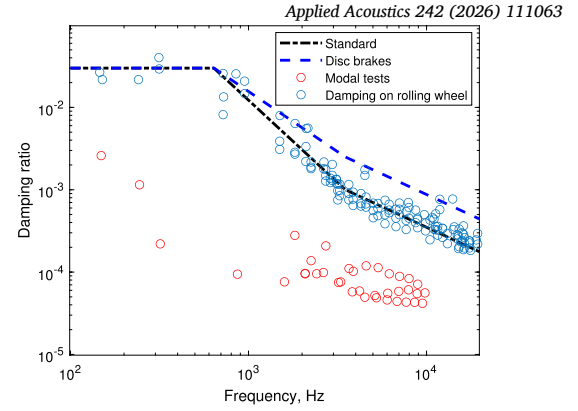


Fig. 9. Total damping curves for a standard wheel and for a wheel with wheel-mounted disc brakes, and measurement data on a standard wheel shown as circles.

wheelset; (ii) a wheelset with wheel-mounted brake discs and (iii) a wheelset fitted with wheel dampers. The total damping curve for a standard wheelset was first calculated based on measurements on a running wheel made available from the European Roll2Rail project [25], and those of the other two types were then inferred empirically. More specifically, wheel vibration was recorded by means of wheel-mounted accelerometers. Spectral peaks were analysed by means of the half power bandwidth method to evaluate the total effective damping ratio during rolling. The total effective damping was found to depend on frequency but not to be associated with the mode type. Static impact tests were also performed in axial and radial directions for the same wheel when lifted from the track. These have been analysed to extract structural damping ratios at the natural frequencies. The measured total damping ratios varied from 0.03 at low frequency to about 2×10^{-4} at the highest frequency recorded (20 kHz). The structural damping extracted from impact test measurements are lower and mostly between 0.001 and 5×10^{-5} . Above 1 kHz, the structural damping ratio was predominantly around 10^{-4} . These results demonstrate that, for the case considered in these tests, the apparent rolling damping can be significant in the whole frequency range. Different wheels may give slightly different results, but it is expected that the main trends and levels would be similar.

Curve fitting was used on the measured data in Fig. 9 to express the total effective damping as a function of frequency. The standard damping settings shown in Fig. 9 are used for the wheels of the freight wagons.

Measurements have shown that wheel-mounted brake discs can affect the structural damping of the wheel and therefore modify the total damping curves. The structural damping introduced by wheel-mounted disc brakes was extracted from modal impact tests performed on a wheelset fitted with disc brakes, and the additional damping, compared to a standard wheelset, was used to empirically modify the damping curve as shown in Fig. 9. These settings are used for the wheels of passenger trains and locomotives.

To determine the amplitudes A_n in Eq. (8), transfer functions for radial and axial wheel motions are predicted using a TWINS-based model [26]. These are obtained as transfer functions in one-third octave bands from 50 Hz to 5 kHz and used as a target curve in an iterative optimisation process.

The four vertically distributed point sources per wheelset are assumed to be incoherent. Their incoherency is achieved in the filter design by randomising the time delays in Eq. (8).

3.3. Track sources

This section describes a new efficient approach to model the track contributions in a train pass-by auralisation. This approach makes use of previous work by combining two approaches. Whereas, previous work

modelled the track contribution either by moving sources (e.g. [6]) or by fixed coherent sources (e.g. [7]), here we propose a combination of both moving and fixed sources for the track. The moving sources represent strongly decaying structural waves that are localised around the excitation point and thus travel with it, whereas fixed sources represent propagating structural waves in the rail that make the rail appear as an extended source. The different types of equivalent sources for the track are illustrated in Fig. 8.

The sound radiated from the track is modelled by equivalent point sources that are a combination of moving sources and spatially distributed sources fixed to the track. All track sources are located 0.1 m below the near side rail head. Contributions from the far side rail are considered in a simplified way with an artificial amplification.

The track contribution considers and distinguishes the following components: i) Vertical rail motion of propagating waves, ii) Lateral rail motion of propagating waves, iii) Decaying structural waves in the rail (both vertical and lateral), iv) Sleeper vibration. For these four track contributions, transfer functions from unit roughness to sound power in one-third octave bands are computed using a TWINS-based model. This model also provides corresponding vertical and lateral track decay rates (TDRs).

3.3.1. Moving track sources

The moving track sources are associated to each wheelset i of the train below the wheel/rail contact and thus move with the train speed V in the driving direction. The moving sources represent the radiation of: i) All waves in the rail with a TDR greater than a certain value (i.e. 4 dB/m), ii) Sleeper vibration. The moving track source signal per wheelset i is modelled by

$$s_{\text{tr, mov}, i}(t, \varphi_i, \theta_i) = \sqrt{2}\rho_0 (\dot{Q}_{\text{rail, mov, lat}, i}(t)\Psi_{\text{lat}, i}(t) + \dot{Q}_{\text{rail, mov, vert}, i}(t)\Psi_{\text{vert}, i}(t) + \dot{Q}_{\text{sleeper}, i}(t)) \quad (11)$$

with separate volume velocities for the lateral and vertical rail and for the sleeper contributions, and directivity functions Ψ for the two rail contributions. The factor $\sqrt{2}$ in Eq. (11) accounts for a sound power increase by the far side rail by 3 dB. The two rail directivities are modelled with maximum radiation perpendicular to the rail web and no radiation in the track direction:

$$\Psi_{\text{lat}} = \sqrt{3} \cos \varphi \cos \theta \approx \sqrt{3} \cos \varphi \quad (12)$$

$$\Psi_{\text{vert}} = \sqrt{2} \cos \varphi \quad (13)$$

For both rail contributions, a dipole characteristic is assumed in the horizontal plane ($\theta = 0$). In the vertical plane, for the lateral rail contribution a dipole directivity is assumed whereas the vertical rail contribution is omnidirectional. Since the considered receiver geometries have small elevation angles, the vertical directivity of the lateral contribution is omitted to simplify the subsequent propagation simulation. This approximation leads to an overestimation of the lateral contribution for large absolute elevation angles (e.g. 6 dB for $|\theta| = 60^\circ$), such as in the case of a high noise barrier or for ground reflections when there is a high embankment. Also the contribution from the far side rail will be overestimated in these situations because it is just modelled by a 3 dB increase and thus the larger distance from the receiver and shielding by the train body are not considered.

The volume accelerations in Eq. (11) are calculated by linear convolution of the effective roughness excitation signal $v_{r,i}$ with the impulse response h of a structural transmission path filter:

$$\dot{Q}_{\text{rail, mov, v}, i}(t) = (v_{r,i} * h_{\text{rail, mov, v}, i} * g_{\text{rs}, v})(t) \quad (14)$$

$$\dot{Q}_{\text{sleeper}, i}(t) = (v_{r,i} * h_{\text{sleeper}, i})(t) \quad (15)$$

with the directions $v \in \{\text{lat}, \text{vert}\}$. The functions g_{rs} introduce the frequency-dependent effects of the rail shields on the respective rail con-

tributions. They are computed from the spectral insertion loss IL by the inverse Fourier transform

$$g_{\text{rs}}(t) = \mathcal{F}^{-1} \{10^{-IL(f)/20}\}. \quad (16)$$

Without rail shields, $g_{\text{rs}, v} = \delta(t)$. The three time-domain filters h for the structural transmission paths are created based on transfer functions and TDRs in one-third octave resolution. For the sleeper contribution in Eq. (15), the whole sound power of the sleeper vibration is attributed to the moving source.

The link to the volume acceleration is achieved through the relationship between the sound power W and the root-mean-square volume acceleration, $\dot{Q}_{\text{rms}}$, of a monopole point source given by

$$\dot{Q}_{\text{rms}}^2 = \frac{4\pi c_0}{\rho_0} W \quad (17)$$

with speed of sound c_0 . Since the source allocations are based on power, it has to be ensured that the signals in Eq. (11) add up incoherently. This is achieved by introducing a partial phase randomisation between the three filter functions h used in Eqs. (14) and (15). As for the wheel sources, the computational effort for the moving track sources depends linearly on the number of wheelsets of the train.

3.3.2. Fixed track sources

Highly propagating waves in the rail are characterised by a low TDR. The contribution of these waves is modelled using an equivalent source model. It represents the radiated sound field by an array of fixed, spatially distributed sources. In [27], a line array of equally spaced coherent monopole sources was used. Requiring at least four sources per acoustic wavelength, this would result in more than 1000 sources for a full train pass-by auralisation. As a consequence, the high computation times make this approach unacceptable from a user perspective. Therefore, a different approach is proposed here assuming an array of incoherent sources.

In the following, the coordinate in track direction is denoted as x . Due to geometrical spreading and source directivity, track sections close to the observer are more relevant for the sound at the observer compared with distant sections. This is used in the model to reduce further the number of sources and accordingly the computational load. The equivalent sources are therefore non-equally spaced. The source density is highest at the shortest distance to the observer. To create the equivalent sources, the track is longitudinally segmented into sections of different length. The segments are chosen to cover similar aspect angles from the receiver. A source j fixed to the track is introduced at the centre of each segment. The source signal of the fixed track source j is modelled by the sum of the lateral and vertical rail vibration contributions from all wheelsets i using

$$s_{\text{tr, fix}, j}(t, \varphi_j, \theta_j) = \sqrt{2}\rho_0 \sum_i (\Psi_{\text{lat}, j} \dot{Q}_{\text{rail, fix, lat}, i, j}(t) + \Psi_{\text{vert}, j} \dot{Q}_{\text{rail, fix, vert}, i, j}(t)) \quad (18)$$

where Ψ represent directivities for lateral and vertical rail motion. The volume velocity for the fixed track sources consists of the superposition of two volume accelerations, i.e. for vertical and lateral rail motion. A proportion of the energy of the propagating waves is allocated to the fixed track sources, separately for vertical and lateral direction. The rail contributions are computed with a TWINS-based model and attributed to the fixed sources if the TDR is lower than 4 dB/m. The movement of the rail excitation point and the frequency dependence of the spatial decay of the corresponding propagating waves is considered. For the latter, the TDRs in one-third octave bands are used. From the TDR Δ (expressed in dB/m) the imaginary part of the complex wavenumber k in the rail is calculated by [24]

$$\delta = \Im\{k\} = -\Delta/8.686 \quad (19)$$

Assuming a simple beam model and neglecting the evanescent wave component, the sound power of a vibrating rail is approximately given by [24]

$$W \approx \frac{1}{4} \rho_0 c_0 \sigma P |v(0)|^2 \int_{-\infty}^{\infty} |e^{-ikx}|^2 dx, \quad (20)$$

with the air density ρ_0 , speed of sound c_0 , the frequency dependent radiation ratio σ , the rail perimeter length P and the velocity $v(0)$ at the excitation point. After changing the integration limits and the excitation point in Eq. (20), the integral was solved analytically to obtain an expression for the partial sound power $W_{j,i}$ per track segment j excited by wheelset i :

$$W_{v,j,i}(t, f) = \begin{cases} \frac{W_{v,i}}{2} \left(2 - e^{2\delta_v |x_{w,i}(t) - x_{0,j+1}|} - e^{2\delta_v |x_{w,i}(t) - x_{0,j}|} \right), & \text{if } x_{w,i} \in [x_{0,j}, x_{0,j+1}] \\ \frac{W_{v,i}}{2} \left| e^{2\delta_v |x_{w,i}(t) - x_{0,j+1}|} - e^{2\delta_v |x_{w,i}(t) - x_{0,j}|} \right|, & \text{else} \end{cases} \quad (21)$$

for the lateral and vertical directions $v \in \{\text{lat}, \text{vert}\}$ and the total sound power W_i associated with wheelset i of the highly propagating waves in the rail

$$W_{v,i} \approx \sum_j W_{v,j,i}. \quad (22)$$

$x_{w,i}$ denotes the time-varying position of the wheel/rail contact at wheelset i . $x_{0,j}$ and $x_{0,j+1}$ are the bounds of track segment j . The geometry is depicted in Fig. 8. The first condition in Eq. (21) describes the situation when the wheel is running on the track segment j ; the second condition accordingly when the wheel is outside the respective track segment. Eq. (21) is evaluated separately for the vertical and lateral rail motion using the respective values for δ . The frequency dependence is introduced by both the spatial decay δ and the sound power W_i . The time dependence is due to the variable position of the wheel/rail contact. Time histories of the spectral partial sound powers $W_{v,j,i}$ are used to design time-varying linear filters in order to create the two volume velocity signals, i.e. lateral and vertical direction, per track segment in Eq. (18).

The ratio between the total sound power $W_{v,i}$ associated with one wheel per vibration direction and a unit roughness excitation is computed with a TWINS-based model and expressed as a transfer function in one-third octave bands. Based on these transfer functions and the derived expression for the allocation of the sound power per segment j (Eq. (21)), the volume velocities in Eq. (18) are derived by linear convolution of the excitation signal $v_{r,i}$ with the impulse response of a time-varying structural transmission path filter h^t :

$$\dot{Q}_{\text{rail,fix},v,i,j}(t) = \left(v_{r,i} * h_{\text{rail,fix},v,i,j}^t * g_{\text{rs},v} \right)(t) \quad (23)$$

Similarly to the moving track source contributions, the TWINS-based model results for the lateral and vertical rail motions are assumed to add up incoherently. Before summing up these contributions using Eq. (18), the corresponding signals have to be decorrelated. To achieve this, random phase differences are added within the design of the two required filters h in Eq. (23) for $v = \text{lat}$ and $v = \text{vert}$.

The track discretisation is critical for the performance and the fidelity of the model. The number of track segments affects the overall computational load and the memory demands of the auralisation. A small number of track segments is thus beneficial. Furthermore, too closely spaced sources would violate the assumption of incoherence between the track segment contributions, because the assumed incoherence is in fact approximated by the different propagation time delays from the fixed track sources to the observer. At the same time, the chosen discretisation should result in an adequate spectral reproduction and be perceptually acceptable, i.e. result in a smooth sound perception in time, frequency and source localisation. For the tested propagation scenarios,

using around 20 fixed track sources with a minimum separation of 2 m seems to be a reasonable compromise between fidelity and computation time. This small number of fixed track sources makes this modelling approach applicable in an auralisation tool for users. The computational effort for the fixed track sources depends linearly on the number of fixed track sources and on the number of wheelsets in the train.

3.4. Synthesis of other sources

Sound sources related to the traction, auxiliary systems or aerodynamics can also be perceptually important and, in some cases, even dominate the overall sound pressure level of a rail vehicle pass-by. This applies especially when considering specific frequencies, particularly at low frequencies. Even if rolling noise dominates the spectral band levels, it is found that these additional sources affect the aural impression and thus need to be included for a convincing train pass-by auralisation. The following sections briefly describe the synthesis of these sources.

3.4.1. Rail singing

Rail singing was found to be relevant for auralisation in the DESTINATE project [11]. Rail singing is a narrowband sound that is radiated from the rail and can be perceived when the observer distance to the train is much larger than the distance to the track. Due to its large horizontal source extension, it is modelled as a coherent time-varying line source, as in previous auralisation work [11].

A single source signal is synthesised by a subtractive synthesis technique, i.e. by filtering a white noise signal. The band-limited sound is created with a centre frequency of 1.2 kHz and a bandwidth of 100 Hz, as derived from pass-by measurements. Unlike the other rolling noise sources, the rail singing effect is not evaluated through physics-based modelling. Instead, it is empirically tuned to produce a plausible contribution in the 1.2 kHz band, consistent with observations from measurements. As its generation mechanism is still based on roughness excitation, the source strength is empirically linked to the 1.25 kHz one-third octave band of the rolling noise track contribution.

3.4.2. Traction, auxiliary equipment and aeroacoustic sources

With the exception of specific tonal components, source signals for traction, auxiliary equipment and aerodynamic noise are generated by a subtractive synthesis technique, i.e. by filtered random noise signals. The acoustical inputs are expressed as sound powers in one-third octave bands from 50 Hz to 10 kHz and corresponding source directivities.

To allow for variation in the train speed V , the sound power L_W of the sources are interpolated between the known sound powers at neighbouring speeds, typically 80 and 160 km/h as used during certification, by

$$L_{W,m}(f, V) = L_{W,m,V_0}(f) + 10 \log_{10} \left(\left(\frac{V}{V_0} \right)^{b_m(f)} \right) \quad (24)$$

with the sound power L_{W,m,V_0} of source m at reference speed V_0 , and the exponent b_m for the speed dependence. Sound power spectra for the different components such as heating, ventilation, and air conditioning (HVAC) and motors were obtained from vehicle manufacturers or from microphone array pass-by measurements [17]. Pass-bys of international freight trains from the long-term railway noise monitoring data from Switzerland were analysed to adjust the sound power of the freight locomotive at low frequencies. Aeroacoustic sources were taken from the Swiss sonRAIL model [17] where a constant speed exponent of $b_m = 6$ is used for all frequencies. Due to lack of detailed source directivity data, general frequency-independent dipole characteristics are used.

3.5. Propagation simulation

The propagation effects from source to observer are applied to each source signal individually, as expressed by the equations in Section 3.1.

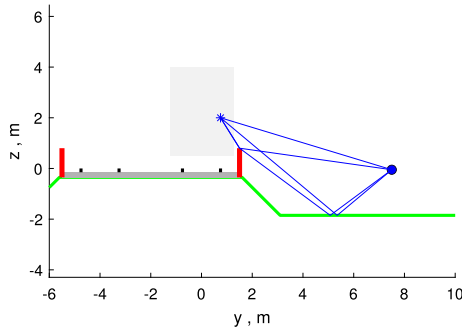


Fig. 10. Propagation geometry in a cross section showing the four propagation paths (blue), the embankment and terrain (green), and low-height barriers (red), for a source 2 m above rail and a receiver at 7.5 m horizontal distance to the centre of the close track.

This is realised by processing the source signals with a set of digital linear time domain filters. Due to source motion, all propagation effects change over time which makes them time-varying filters. The propagation model considers Doppler effects, propagation delay, geometrical spreading, ground reflections, edge diffraction and air absorption. Meteorological effects are ignored but could conceptionally be added (e.g. [28–30] for turbulence effects). Sound propagation simulation is performed in vertical cross-sections through source and receiver. Conventional noise barriers or low height barriers are simulated by modelling the sound diffracted by the barrier top edge, and the inherent changes to the ground effect. The ground effect is also affected by whether the track is on an embankment, the observer location and the ground properties. In the absence of a barrier, the embankment edge may also shield low height sources such as the rail.

Fig. 10 illustrates the four propagation paths considered. These paths are used to model the sound signal produced by source m at the observer location. Ground reflected paths are created by mirroring the observer in the terrain below the observer. Reflections near the track or propagation through the ballast bed (as investigated in [31]) are neglected—implicitly they are assumed to be contained in the source signals. Fully sound absorbing barriers are assumed, i.e. no reflections from barriers or buildings are considered. The propagation path index n introduced in Section 3.1 thus takes values from 1 to 4. The total sound pressure is modelled as the sum of the four corresponding partial sound pressures for direct, ground-reflected, diffracted and diffracted-reflected sound:

$$p_m(\mathbf{r}, t) = p_{m,\text{dir}}(t) + p_{m,\text{refl}}(t) + p_{m,\text{diff}}(t) + p_{m,\text{diff/refl}}(t) \quad (25)$$

Each path term is calculated considering its own time-varying propagation path geometry and by applying a cascade of time-varying filters to the source signal:

$$p_{m,n}(\mathbf{r}, t) = \left(\frac{s_m}{4\pi r_{m,n}} * g_{\text{Doppler},m,n}^t * g_{\text{refl},m,n}^t * g_{\text{diff},m,n}^t * g_{\text{air},m}^t \right)(t) \quad (26)$$

where s_m is a source signal of source m as described in Sections 3.2 and 3.4, r denotes the time-varying length of the respective propagation path from source to receiver. The time-varying functions g^t denote filter impulse responses for different propagation effects. Each path has its own time-dependent Doppler frequency shift, Doppler amplification, propagation delay (summarised in g_{Doppler}^t) and geometrical spreading. For the ground-reflected sound paths, g_{refl}^t represents the reflection factors, which for the other two paths is omitted (formally, $g_{\text{refl}}^t = \delta(t)$). A spherical wave reflection at a finite impedance plane is assumed [32]. The complex frequency-dependent ground impedance is calculated as a function of the ground type using the Miki model [33]. The static air flow resistivity is set to 10,000 kPa s/m² for asphalt or 100 kPa s/m² for grassland. For the diffracted sound paths, g_{diff}^t represents the diffraction,

which for the other two paths is omitted (formally, $g_{\text{diff}}^t = \delta(t)$). Single-edge diffraction at an infinitely extended edge is modelled based on Pierce's analytical complex-valued frequency domain expression [34]. Accordingly, if there is no line of sight between source and observer, $p_{m,\text{dir}} = 0$, and analogously for the ground reflected-sound $p_{m,\text{refl}}$. For shielded sources, the propagation simulation thus only requires the computation of two propagation paths, resulting in a halving of the computational load.

To simulate air attenuation, spectral air absorption coefficients are calculated from ISO 9613-1 [35]. A homogeneous atmosphere is assumed. Air absorption is simulated by a time-varying minimum phase FIR filter and, unlike other filters, identically applied to all paths. This simplification increases the computational efficiency without impairing the results since the small relative differences in path length will not substantially affect air absorption.

4. Reproduction

4.1. Spatialisation

To support different reproduction systems, different rendering strategies and output formats are used. The spatial rendering requires as input the time-varying azimuthal angle ϕ of sound incidence at the observer for every source m .

4.1.1. Monophonic

A single-channel mono signal is created by ignoring the angles of incidence at the observer and simply storing the sound pressure signal in Pascal according to Eq. (1). This type of signal has to be used for calculations of sound pressure levels or psychoacoustic parameters.

4.1.2. Two-channel stereo

For two-channel loudspeaker playback, a virtual microphone array is emulated. The mixed-stereo technique by the Office de Radiodiffusion Télévision Française (ORTF) is chosen. ORTF consists of two tilted cardioid microphones with a separation of 17 cm. The spatial sensor separation leads to time-of-arrival differences depending on the angle of sound incidence. The cardioid patterns lead to different sensitivities depending on the angle of sound incidence. This ORTF rendering was implemented according to [36].

4.1.3. Ambisonics

A scene-based audio format is chosen as an interface of spatial audio with various commercial systems. Ambisonics is based on a sound field decomposition using spherical harmonic functions. It is increasingly used in academia and commercial Virtual Reality (VR) environments such as Unity or YouTube. A First Order Ambisonic (FOA) B-format is encoded and stored in the AmbiX format with the four channels:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \sum_m \begin{bmatrix} 1 \\ \sin \phi_m(t) \cos \beta_m(t) \\ \sin \beta_m(t) \\ \cos \phi_m(t) \cos \beta_m(t) \end{bmatrix} \cdot p_m(t) \quad (27)$$

with the azimuthal angle ϕ and the elevation angle β of incidence. For simplicity, the elevation angle β was set to 0 which effectively makes it a 2D-FOA format. If needed, a representative elevation angle can be inferred from the propagation geometry. For increased reproduction accuracy, the first-order ambisonic format could be extended to higher order ambisonics (HOA).

4.1.4. Binaural rendering for headphones

For interfacing with the VR tool, a tailored intermediate audio format was developed. For compactness and simplicity, a channel-based format is used where all observer sounds are panned to a 2D array of virtual loudspeakers. Seven virtual speakers are used that are located on a tilted semi-circle facing the track. The azimuthal angles of the virtual speakers

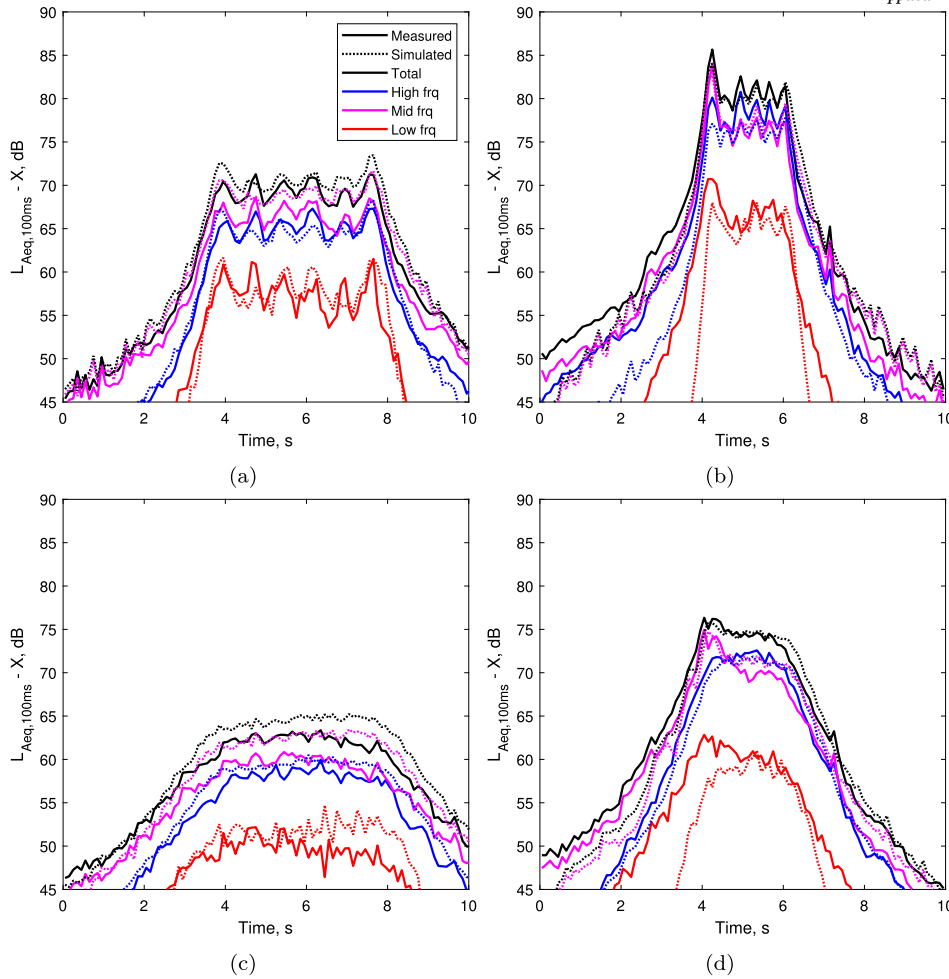


Fig. 11. Comparison of measured and simulated pass-by sound pressure levels for an electric multiple unit train running at 80 km/h (left) and 160 km/h (right) on a ballasted track at distances 7.5 m (top) and 25 m (bottom). The total A-weighted 100 ms short-term L_{eq} are shown in black and three frequency bands in colours (with an arbitrary offset X for confidentiality).

are $\pm 90^\circ$, $\pm 60^\circ$, $\pm 30^\circ$, and 0° , where 0° means facing the track. This arrangement of fixed virtual sources is recreated in the VR environment within the VR tool where real-time binaural rendering for headphones is performed through dynamic head-related transfer function (HRTF) filtering. For this, the physical head orientation of the user is tracked in real time using sensors in the head-mounted display that the user is wearing.

For the panning to the virtual speakers, a type of vector-base amplitude panning (VBAP) introduced by Pulkki [37] is used. The current dual-band VBAP works in two separate frequency bands with a crossover frequency of 400 Hz [38]. This band-specific handling of the signals is due to field distortions introduced by the presence of the listeners head. In the high frequency band, the presence of the head leads to a rather incoherent superposition of the speaker signals at the ears. In classical VBAP this leads to an uneven frequency response with an attenuation of high frequencies, which is here accounted for. This intermediate 7-channel format already includes the pre-equalisation of the headphones required by the VR tool, and the equalisation of the HRTFs used inside the VR tool.

4.2. Calibration

Equalisation and calibration of the binaural headphone reproduction was performed in the lab based on measurements taken using a head-and-torso simulator (45BB KEMAR, GRAS), a free-field measurement microphone, and a sound level meter as a reference, along with a

broadband loudspeaker in an anechoic environment. Once calibrated, the audio interface output voltage was measured using an AC voltmeter while playing a reference signal. This reference voltage serves as a benchmark for future recalibrations, reducing reliance on a HATS. Additionally, in the calibration panel of the software, users are provided with instructions on how to perform this calibration process effectively.

4.3. Required hardware

The SILVARSTAR demonstrator was designed such that independent persons can set up their own system and use it. Apart from the SILVARSTAR VR software, only a few commercially available hardware components are needed. The primary requirements are a PC, a specific VR headset and specific headphones. To ensure high-quality video and audio reproduction, the following hardware is required:

- PC with NVIDIA GeForce RTX 2080 graphics card or higher
- Oculus Meta Quest 2 VR headset and motion controllers
- Meta Quest Link cable and Meta Quest 2 Elite Strap (recommended for better fit and comfort)
- Sennheiser HD650 headphones
- 2-channel audio interface with gain control
- AC voltage meter (for sound level calibration)

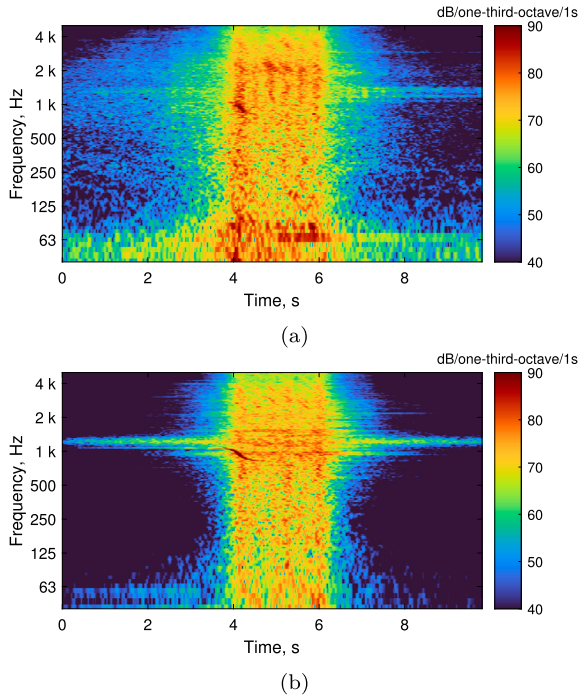


Fig. 12. Spectrograms of measured (a) and synthesised (b) pass-by sounds for an electric multiple unit train running at 160 km/h at 7.5 m distance.

5. Verification and evaluation

5.1. Comparison with pass-by measurements

This section presents some comparisons between pass-by measurements and syntheses. It is generally very difficult to obtain complete measurement datasets which allow for a direct comparison due to lack of sufficient detailed data that are needed as input for a simulation, such as rail and wheel roughness spectra, wheel geometries and TDRs.

A first validation case consists of pass-bys of an electric multiple unit (EMU) train recorded at two speeds and for two receiver distances. The train consists of five articulated coaches and six bogies, of which three are driven bogies. The considered speeds are 80 and 160 km/h. Microphone positions according to ISO 3095 [15] are used, with horizontal receiver distances of 7.5 and 25 m, and heights of 1.2 and 3.5 m, respectively. The track was a ballasted track with monoblock sleepers, hard rail pads and a smooth rail. Sound level time histories calculated from the recordings and from the syntheses are shown in Fig. 11 for both speeds and receivers. For these four pass-bys, the differences in the A-weighted sound exposure level, L_{AE} , between synthesis and recording lie between -0.5 and 2.2 dB. A-weighted 100 ms short-term L_{eq} are shown for the total and for three selected frequency bands, namely the octave bands 63–250 Hz (low), 500 and 1000 Hz (mid), 2 and 4 kHz (high). For all frequency bands, the level-time histories calculated from the synthesised sounds agree very well with the pass-by recordings. Also, the sound level ripples caused by the bogie passages for the close observer, are very well reproduced by the synthesis.

Fig. 12 shows spectrograms for the pass-by of the EMU running at 160 km/h, from the recording and from the synthesis. The continuous component slightly above 1 kHz is due to rail singing, present both in the synthesis and the recording, however somewhat masked in the latter during the approach of the train. The spectro-temporal patterns at 1–4 kHz are well reproduced by the synthesis. They are due to the bogie passages and the Doppler shift of the wheel contribution due to the high train speed. Some deficiencies in the auralisation can be identified below 100 Hz. Here, sources related to the traction seem to be underpredicted. Note that a strong tonal noise component at 900 Hz associated with

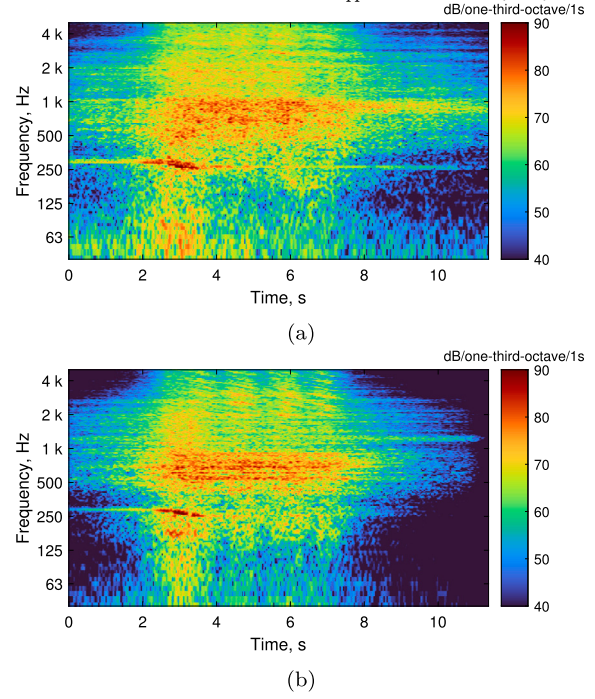


Fig. 13. Spectrograms of measured (a) and synthesised (b) pass-by sounds for a locomotive with three passenger coaches running at 80 km/h at 8.3 m distance.

the first propelled carriage was introduced in the model empirically to match the measurement.

Different types of vehicle are compared in the spectrograms in Fig. 13. Here a recording and synthesis of the test train from the Roll2Rail project [25] travelling at 80 km/h are shown. The train consisted of an electric locomotive with four wheelsets (type Siemens ES64U2) pulling three passenger coaches of 26 m length. The microphone is at 8.3 m from the track centre and 3.5 m above rail head. Also in this example, spectro-temporal patterns are well reproduced by the model. Above 1 kHz, horizontal patterns are visible in Fig. 13 which are due to the modal contributions from the wheels. The rolling noise contributions in the 250 Hz to 1 kHz range are stronger and more broadband. At around 300 Hz, there are noticeable tonal components from the locomotive which are pitched down in frequency during the passage due to the Doppler shift. These tones were added to the model to match the measurements. For the whole train pass-by, the difference between synthesis and recording in L_{AE} is -1.2 dB for 80 km/h, and -1.4 dB for 160 km/h.

Across all six train pass-by cases, the root mean square deviation (RMSD) of the L_{AE} is 1.4 dB. In a single 1/3 octave band between 50 Hz and 10 kHz, the average RMSD is 2.9 dB. The largest RMSDs occur at the low frequencies with more than 5 dB in the 1/3 octave bands below 100 Hz.

5.2. Analysis of partial sources

In the physics-based auralisation model presented in Section 3, the different source contributions are separately synthesised and propagated to the receiver. This allows an analysis of the individual components which can be listened to separately.

As an example, Fig. 14 shows such an analysis of the EMU train pass-by at 80 km/h at 7.5 m distance from Fig. 11a. One-third octave spectra and A-weighted level time histories are shown for six different source contributions. Three of these contributions are radiated from the track and the other three are related to the vehicles. The sound radiated from the track is shown in terms of the contributions from the rails and the sleepers, as well as rail singing. For the vehicles, rolling

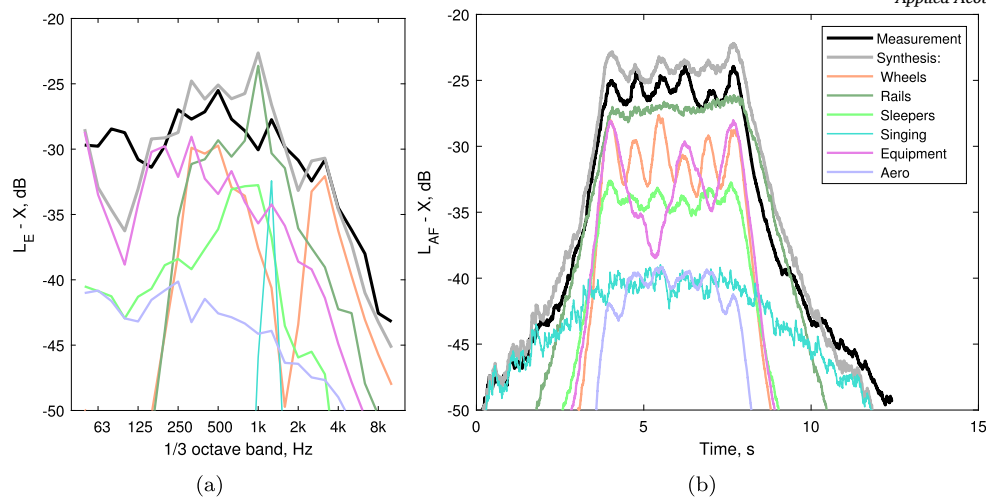


Fig. 14. Source-resolved analysis of a synthesised train pass-by with (a) sound exposure levels in 1/3 octave bands and (b) A-weighted FAST-time-weighted sound pressure levels 14b, showing separately the generated source components (in colours) related to the wheels, the rails, the sleepers, rail singing, equipment and traction, and aeroacoustic sources. The measurement (black) and the complete synthesis (grey) are shown for reference. The electric multiple unit train pass-by with 80 km/h at 7.5 m distance from Fig. 11a is shown.

noise radiated from the wheels, equipment and traction, and aeroacoustic contributions are displayed. This example shows that depending on frequency and time, different contributions dominate the total sound pressure level.

Below 250 Hz, contributions from auxiliary equipment and the traction dominate the sound exposure level. In the frequency range 250–800 Hz, three components provide a similar energy contribution, i.e. rolling noise from the wheels and the rails, as well as equipment and traction. At 1 kHz, there is a strong peak from the rail contribution which is over-predicted. Above 2 kHz, the wheel contribution dominates the energy spectrum. Looking at the temporal structure, the level time histories show that rail singing is dominant about 1 s before and after the pass-by. A different temporal behaviour can be observed for the two rolling noise contributions, i.e. from the rails and from the wheels. The ripples caused by the six bogie passages are only present for the wheels, but not for the rails. This is a consequence of the strong contribution from propagating waves in the track for this particular track. The three peaks in the level time history for the equipment and traction contribution are related to the traction-related source from the three driven bogies. From the shown results, one may conclude that the sleeper contributions and aeroacoustic sources are of minor importance. However, this is not always the case—e.g. for higher speeds aeroacoustic sources will increase in importance.

This example and other cases showed that a variety of sources needs to be considered for a full train pass-by auralisation. It was found that an auralisation goes beyond traditional noise level predictions because also frequencies outside the 100–5000 Hz range (e.g. traction and aeroacoustic sources) and sounds before and after the pass-by matter and thus need to be accounted for (e.g. rail singing and source directivities).

5.3. Perceptual evaluations

The authors and other experts in railway acoustics critically compared the syntheses produced with the described model with recorded train pass-bys. The syntheses were attested to have high quality and realism. The recorded situations however only consisted of situations without noise mitigation measures in place.

Calibrated audio data, along with the necessary software for auralisation, is freely available, allowing readers and researchers to evaluate the quality of the auralisations themselves [39]. While the authors have not conducted formal perceptual listening tests, perceptual tests were performed using the results of the model presented in this paper in the European FINE-2 project. Below, a concise summary of their findings is

given, focusing specifically on the plausibility of the auralisations. More detailed information is available in [40].

In [40], an AB comparison test was conducted, evaluating eight train pass-by stimuli—four auralisations and four measurements. Additionally, two validation tests were included, comparing an auralisation against another auralisation and a measurement against another measurement, resulting in a total of six auralisation stimuli and six measurement stimuli. All stimuli were presented monaurally, and participants were exposed only to the auditory stimuli, without any accompanying virtual environment.

The study involved 19 experts in railway acoustics, who evaluated each of the 12 stimuli by answering two key questions:

1. Does the recording sound like a real train?
2. Is the stimulus an auralisation or a measurement?

For the auralisation stimuli, 74% of participants responded “Yes” to the first question, 12% were “Unsure”, and 14% answered “No”. In contrast, for the measurement stimuli, 58% responded “Yes”, 18% were “Unsure”, and 24% answered “No”.

Regarding stimulus classification, 54% of the auralisation stimuli were correctly identified as auralisations, while 46% were mistaken for real measurements. Similarly, 48% of the measurement stimuli were correctly identified as measurements, while 52% were incorrectly classified as auralisations.

These results indicate that participants found it challenging to distinguish between auralised and real measurement stimuli, underscoring the high plausibility and realism of the auralisation methodology. It should be noted that the method is currently being tested and applied in ongoing research projects by various research groups.

6. Conclusions

This paper presented a new auralisation model for railway noise. In this model, train pass-by sounds are synthesised using physics-based synthesis methods. This approach allows the integration of different noise mitigation measures and their combination. The presented methods are computationally efficient which is a fundamental requirement for them to be used in VR applications.

For validation purposes, synthesised train pass-bys were compared to recordings. The comparisons showed that the models are capable of reproducing typical spectral and temporal patterns of train pass-bys. In addition, expert listeners rated the syntheses to be plausible. The

models were thus implemented in two software tools. These tools allow for an intuitive and interactive VR demonstration of railway noise. As some first applications, the tools were successfully used for railway noise demonstrations at public exhibitions and information events. Both software tools, along with pre-configured demonstration scenarios, audio files, and supporting materials, are freely available for non-commercial use [39], facilitating further research and application in railway noise assessment and mitigation strategies.

Further work could focus on a thorough perceptual validation of the synthesis model and on extending its applicability towards other train types or mitigation measures, complex propagation environments and special situations and sources, like curves, switches, bridges and tunnel openings.

CRediT authorship contribution statement

Reto Pieren: Writing – original draft, Software, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Fotis Georgiou:** Writing – review & editing, Software, Methodology, Investigation. **Giacomo Squicciarini:** Writing – review & editing, Methodology, Investigation. **David J. Thompson:** Writing – review & editing, Methodology, Investigation, Funding acquisition.

Declaration of competing interest

There are no competing interests of any author. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Dario Sala and Daniel Gremlé from the company Bandara VR for their development work on the VR application. The work presented in this paper has received funding from the Europe's Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101015442). The contents of this publication only reflect the authors' views; the Joint Undertaking is not responsible for any use that may be made of the information contained in the paper.

Data availability

The authors do not have permission to share data.

References

- [1] World Health Organization. Burden of disease from environmental noise: quantification of healthy life years lost in Europe. World Health Organization Regional Office for Europe; 2011. <https://www.who.int/publications-detail-redirect/burden-of-disease-from-environmental-noise-quantification-of-healthy-life-years-lost-in-europe>.
- [2] European Environment Agency. How does environmental noise pollution impact my health? 2023.
- [3] International Union of Railways (UIC). Nuisance and health impacts of railway noise; 2016.
- [4] Seidler A, Schubert M, Mehrjerian Y, Krapf K, Popp C, van Kamp I, et al. Health effects of railway-induced vibration combined with railway noise - a systematic review with exposure-effect curves. *Environ Res* 2023;233:116480. <https://doi.org/10.1016/j.envres.2023.116480>. <https://pubmed.ncbi.nlm.nih.gov/37352957/>.
- [5] Nowakowski T, Komorski P. Tram noise annoyance: the role of different psychoacoustic measures in the assessment of noise. *Appl Acoust* 2024;219:109946. <https://doi.org/10.1016/j.apacoust.2024.109946>. <https://www.sciencedirect.com/science/article/pii/S0003682X24000975>.
- [6] Pieren R, Heutschi K, Wunderli JM, Snellen M, Simons DG. Auralization of railway noise: emission synthesis of rolling and impact noise. *Appl Acoust* 2017;127:34–45.
- [7] Maillard J, Kacem A, Martin N, Faure B. Physically-based auralization of railway rolling noise. In: Proc. of 23rd international congress on acoustics (ICA) 2019; 2019. p. 1667–74.
- [8] Theyssen J, Deppisch T, Pieringer A, Kropp W. On the efficient simulation of pass-by noise signals from railway wheels. *J Sound Vib* 2023;564:117889. <https://doi.org/10.1016/j.jsv.2023.117889>. <https://www.sciencedirect.com/science/article/pii/S0022460X23003383>.
- [9] Theyssen Jannik. The radiation from railway wheel modes and their effect on loudness, sharpness, and equivalent pressure level. *Acta Acust* 2024;8:20. <https://doi.org/10.1051/aacus/2024012>.
- [10] Knuth C, Squicciarini G, Thompson DJ. Railway rolling noise synthesis using a numerical modelling approach. In: Proc. of annual German conference on acoustics (DAS/DAGA); 2025.
- [11] Pieren R, Lauper D, Heutschi K. Demonstrator for rail vehicle pass-by events. In: Proc. of euronoise 2018; 2018. p. 1409–14.
- [12] Bouvet P, Thompson D, Pieren R, Degrande G, Nuber A, Garcia M. Soil vibration and auralisation software tools for application in railways. *Transp Res Proc* 2023;72:973–80. <https://doi.org/10.1016/j.trpro.2023.11.525>. <https://www.sciencedirect.com/science/article/pii/S2352146523008232>.
- [13] Jäggi J, Pieren R, Heusser A, Heutschi K, Wunderli JM. Data mining on railway track noise measurements (minetrack). Final Report Nr. 5211.01856, Tech. Rep. Empa, Swiss Federal Laboratories for Materials Science and Technology; 2023.
- [14] Squicciarini G, Toward M, Thomson D, Jones C. Statistical description of wheel roughness. In: Proceedings of the 11th international workshop on railway noise. Noise and vibration mitigation for rail transportation systems. Uddevalla, Sweden: Springer; 2015. p. 651–8.
- [15] International Organization of Standardization. ISO 3095: acoustics - railway applications - measurement of noise emitted by railbound vehicles; 2013.
- [16] Wu TX, Thompson DJ. A hybrid model for the noise generation due to railway wheel flats. *J Sound Vib* 2002;251:115–39.
- [17] Thron T, Hecht M. The sonrail emission model for railway noise in Switzerland. *Acta Acust* 2010;96:873–83.
- [18] Zhang X, Jeong H, Thompson D, Squicciarini G. The noise radiated by ballasted and slab tracks. *Appl Acoust* 2019;151:193–205. <https://doi.org/10.1016/j.apacoust.2019.03.012>. <https://www.sciencedirect.com/science/article/pii/S0003682X18310491>.
- [19] Squicciarini G, Toward M, Thompson D. Experimental procedures for testing the performance of rail dampers. *J Sound Vib* 2015;359:21–39. <https://doi.org/10.1016/j.jsv.2015.07.007>. <https://www.sciencedirect.com/science/article/pii/S0022460X15005805>.
- [20] Zhang X, Thompson D, Ryue J, Jeong H, Squicciarini G. The effect of rail shields on railway rolling noise. *Int J Rail Transp* 2023;11(4):552–72. <https://doi.org/10.1080/23248378.2022.2106599>.
- [21] Blauert J. Spatial hearing: the psychophysics of human sound localization. The MIT Press; 1996.
- [22] Litovsky RY, Colburn HS, Yost WA, Guzman SJ. The precedence effect. *J Acoust Soc Am* 1999;106(4):1633–54. <https://doi.org/10.1121/1.427914>. https://pubs.aip.org/asa/article-pdf/106/4/1633/7519853/1633_1_online.pdf.
- [23] Thompson DJ, Remington PJ. The effects of transverse profile on the excitation of wheel/rail noise. *J Sound Vib* 2000;231(3):537–48.
- [24] Thompson D. Railway noise and vibration - mechanisms, modelling and means of control. 2nd edition. Oxford: Elsevier; 2024.
- [25] Thompson D, Squicciarini G, Zhang J, Lopez Arteaga I, Zea E, Dittrich M, et al. Assessment of measurement-based methods for separating wheel and track contributions to railway rolling noise. *Appl Acoust* 2018;140:48–62. <https://doi.org/10.1016/j.apacoust.2018.05.012>. <https://www.sciencedirect.com/science/article/pii/S0003682X1731071X>.
- [26] Thompson DJ, Hemsworth B, Vincent N. Experimental validation of the twins prediction program, part 1: description of the model and method. *J Sound Vib* 1996;193(1):123–35.
- [27] Kitagawa T, Thompson D. The horizontal directivity of noise radiated by a rail and implications for the use of microphone arrays. *J Sound Vib* 2010;329(2):202–20. <https://doi.org/10.1016/j.jsv.2009.09.002>. <https://www.sciencedirect.com/science/article/pii/S0022460X09007020>.
- [28] Pieren R, Georgiou F, Squicciarini G, Thompson DJ. Auralisation of combined mitigation measures in railway pass-by noise. In: Proc. of InterNoise 2022; 2022. p. 518ff.
- [29] Forsén J. Scintillating and decorrelating signals for different propagation paths in a random medium. *Appl Acoust* 2024;221:110038. <https://doi.org/10.1016/j.apacoust.2024.110038>. <https://www.sciencedirect.com/science/article/pii/S0003682X24001890>.
- [30] Lincke Dorothea, Pieren Reto. Auralization of atmospheric turbulence-induced amplitude fluctuations in aircraft flyover sound based on a semi-empirical model. *Acta Acust* 2024;8:47. <https://doi.org/10.1051/aacus/2024036>.
- [31] Heutschi K. Sound propagation over ballast surfaces. *Acta Acust* 2009;95:1006–12.
- [32] Chessel CI. Propagation of noise along a finite impedance boundary. *J Acoust Soc Am* 1977;62(4):825–34. <https://doi.org/10.1121/1.381603>. https://pubs.aip.org/asa/jasa/article-pdf/62/4/825/12036803/825_1_online.pdf.
- [33] Miki Y. Acoustical properties of porous materials. Modifications of Delany-Bazley models. *J Acoust Soc Jpn* 1990;11(1):19–24. <https://doi.org/10.1250/ast.11.19>.

- [34] Pierce AD. Diffraction of sound around corners and over wide barriers. *J Acoust Soc Am* 1974;55(5):941–55. <https://doi.org/10.1121/1.1914668>.
- [35] International Organization of Standardization. ISO 9613-1: acoustics – attenuation of sound during propagation outdoors - Part 1: calculation of the absorption of sound by the atmosphere; 1993.
- [36] Pieren R, Büttler T, Heutschi K. Auralization of accelerating passenger cars using spectral modeling synthesis. *Appl Sci* 2016;6(1). <https://doi.org/10.3390/app6010005>. <https://www.mdpi.com/2076-3417/6/1/5>.
- [37] Pulkki V. Virtual sound source positioning using vector base amplitude panning. *J Audio Eng Soc* 1997;45(6).
- [38] Taghipour A, Pieren R, Schäffer B. Short-term annoyance reactions to civil helicopter and propeller-driven aircraft noise: a laboratory experiment. *J Acoust Soc Am* 2019;145(2):956–67. <https://doi.org/10.1121/1.5090500>. https://pubs.aip.org/asa/jasa/article-pdf/145/2/956/13531756/956_1_online.pdf.
- [39] Empa - Swiss Federal Laboratories for Materials Science and Technology. Silvarstar project - virtual railway noise and vibration simulation tools. <https://www.empa.ch/web/silvarstar>, 2025.
- [40] Nemitz T. Fine-2 deliverable 9.4: final evaluation and analysis of the a&v software tool. https://projects.shift2rail.org/s2r_ipcc_n.aspx?p=fine-2, 2021. [Accessed 10 February 2025].