

# CPX based synthesis for binaural auralization of vehicle rolling noise to an arbitrary positioned stander-by receiver

Frederico Pereira<sup>a,\*</sup>, Francisco Soares<sup>b</sup>, Carlos Silva<sup>a</sup>, Emanuel Sousa<sup>a</sup>, Elisabete Freitas<sup>b</sup>

<sup>a</sup>Centro de Computação Gráfica, Guimarães, Portugal

<sup>b</sup>Portugal, University of Minho, Portugal

## ARTICLE INFO

### Article history:

Received 21 January 2021

Received in revised form 24 May 2021

Accepted 25 May 2021

Available online 9 June 2021

### Keywords:

Binaural auralization

Traffic noise

CPX

Pedestrian

Virtual reality

Audiovisual scenarios

## ABSTRACT

Virtual reality is becoming an important tool for studying the interaction between pedestrians and road vehicles, by allowing the analysis of potentially hazard situations without placing subjects in real risk. However, most of the current simulators are unable to accurately recreate traffic sounds that are congruent with the visual scene. This has been recognized as a fault in the virtual audio-visual scenarios used in such contexts. This study proposes a method for delivering a binaural auralization of the noise generated by a moving vehicle to an arbitrarily located moving listener (pedestrian). Building on previously developed methods, the proposal presented here integrates in a novel way a dynamic auralization engine, thus enabling real-time update of the acoustic cues in the binaural signal delivered via headphones. Furthermore, the proposed auralization routine uses Close ProXimity (CPX) tyre-road noise signal as sound source input, facilitating the quick interchangeability of source signals, and easing the noise collection procedure. Two validation experiments were carried out, one to quantitatively compare field signals with CPX-derived virtual signal recordings, and another to assess these same signals through psychoacoustic models. The latter aims to assure that the reproduction of the synthesized signal is perceptually similar to one occurring on pedestrian/vehicle interactions during situations of street crossing. Discrepancies were detected, and emphasized when the vehicle is within close distance from the receiver (pedestrian). However, the analysis indicated that these pose no hindrance to the study of vehicle-pedestrian interaction. Improvements to the method are identified and further developments are proposed.

© 2021 Elsevier Ltd. All rights reserved.

## 1. Introduction

In urban centres, the safety of vulnerable road users and the annoyance of the population with noise levels are two of the most common problems directly associated to traffic [1,55]. Interestingly, these apparently unrelated problems are tied in such a way that one might be aggravated by policies designed to tackle the other. While quieter vehicles and road surfaces undoubtedly improve the quality of life of those living in busy areas, low-noise emitting vehicles can also impair pedestrian performance in road crossing scenarios [37], an issue that is gaining relevance as the number of electric vehicles in circulation tends to increase.

Exposing people directly to traffic noise to study the impacts and ways of mitigating these problems is methodologically difficult and can even be highly risky. Thus, experimental methodologies based on the presentation of stimuli in virtual reality

settings are gaining relevance as tools to carry out studies involving road vehicles-pedestrians interaction.

Virtual reality is defined as a computer- or video-generated environment that gives the user a sense of being present in a virtual world [46,49]. Among the commercialized virtual reality systems, projection-based (such as the CAVE system) and head-mounted displays truly isolate the user from the real world and they can immerse them into a traffic environment [13]. The feeling of being actually present in the virtual environment and not just perceiving it as a digital image is enhanced by multiple factors, such as: the quality of the graphical and acoustic representations; the possibility for user interaction with the virtual world; and the realism of the simulated environment [21].

However, there are still some limitations with some of the existing simulators used to study pedestrian safety, including: the inability to provide 'free pedestrian movement' opportunities; issues of presenting images of the virtual world in such a way that affect depth perception; and absence of appropriate sounds from the simulated traffic environment [13]. In fact, little relevance has been given to the audio component of virtual scenarios. A

\* Corresponding author.

E-mail address: [frederico.pereira@ccg.pt](mailto:frederico.pereira@ccg.pt) (F. Pereira).

considerable part of the developed simulators does not consider it (e.g. [21,36,41,50,57]), and, for those who do, little information is given about the process of editing and rendering the sounds emitted by the road traffic and the surrounding environment [7,10,12,13,40,49].

Because of its high cost, time-consuming procedures, and complex logistics (such as the frequent presence of uncontrolled and extraneous sources of noise), registering naturalistic sounds for implementation in virtual environments may be particularly difficult [51]. There is a need for new methodologies and tools for researchers studying the role of auditory perceptual variables in safety-relevant traffic situations. For perception evaluation of the acoustic cues, realistic implementation of the spatial-temporal characteristics of moving vehicles and moving listeners is paramount. Such methodologies greatly benefit from the application of dynamic binaural auralization (with real-time update to head position and orientation), enhancing virtual auditory scene realism, increasing immersion and plausibility [4].

With the purpose of simplifying and accelerating the development of traffic research experimental procedures, we propose in this paper the implementation of a new method that, based on nearfield tyre-road noise measurement methods, namely Close Proximity (CPX) [19] recordings, allows the synthesis of a vehicle rolling noise.

CPX recordings are less prone to extraneous noise contamination. Also, they allow estimating the noise generated by the contact between tyre and the road in a simple, quick and reliable way, without the need for complex and often incomplete generation prediction models [34]. The road characteristics are particularly relevant as they affect not only the generation but also the propagation mechanisms. Texture is one of the main pavement characteristics correlating well with tyre-road noise [15] and used in several noise generation prediction models.

Taking in consideration source characteristics and propagation transfer paths, the method presented here aims to render an auditory binaural signal plausibly corresponding to one emitted by a real vehicle and perceived by a listener at an arbitrary far-field position.

### 1.1. Background and related studies

Traffic noise emission has been a topic of study for over 30 years. During those years computational modeling evolved to a point where fast simulation of environmental variables and traffic noise propagation are made possible and are being constantly improved and refined. Pioneering research projects, such as *Harmonoise* [54], proposed approaches for predicting long-term average noise levels resulting from road and railway traffic. The *Harmonoise* model uses field measured statistical data to fine-tune mathematical-physical determination of noise parameters, such as in the estimation of source sound power levels. In this model, the separation of sound source characteristics from sound propagation effects enables a powerful generic propagation model that may be applicable to different scenarios. The *Harmonoise* engineering method resulted in a number of algorithms for outdoor noise prediction. Although its purpose did not preclude the development of an auralization engine for traffic noise, the method is of particular interest for the current study due to its advancements in the characterization of source sound power parameters and methods for replicating propagation effects. In the scope of *Harmonoise*, great improvement was made in comparison to the then available models, namely the separation of rolling noise and propulsion noise and attribution of directivity patterns [30]. In the study by Defrance et al. [14] a reference *Harmonoise* model was implemented and validated with experimental results, showing excellent agreement with flat ground surfaces cases.

Following the *Harmonoise* methods, the *LISTEN* project [42] developed a software-based auralization routine of pass-by road traffic at constant speed for a still-standing listener. The whole auralization process in the *LISTEN* project may be summarized as (1) starting with the creation of the noises and tones of the sources (2) followed by the modelling of the air attenuation, ground effects, directivity, Doppler effect, spherical spreading and (3) ending with the Head Related Transfer Functions (HRTF) convolution. Two approaches for source signals estimation were developed: a time-domain based and an additive synthesis based method. The first method relies on mono recordings of straight-line pass-by at fixed point. This signal is subjected to an inversion of the effects of propagation and directivity to obtain an estimation for the sound source. On the second, synthesis based method, sound source noise is estimated using the calculation of 1/3-octave sound power spectrum from pass-by data, to shape the noise-like character of the additive synthesis. Sources, disposition and ground effects on propagation are based on *Harmonoise* references. In the *LISTEN* project, perception of the pass-by sound was evaluated for realism, annoyance, velocity, and similarity between real and auralized sounds. Evaluation results for time-domain method suggested that listeners discriminate between real and auralized sounds and that the task is easier for lower speeds. Annoyance and velocity were approximately equivalent between real and auralized sounds. Similarity was judged mainly based on velocity. The higher the value, the more similar it was perceived. Greater levels of similarity between real and synthesized sounds, were found for vehicles passing at more than 65 km/h. In these experimental validations, the user was assumed to be stationary and as such, no dynamic cues from head movements were available. Several studies have demonstrated that, when using auralized sounds, dynamic cues from head movements are important so to minimize externalization problems, increase localization performance, and overall enhanced realism of auditory scene [4,44].

More recently, Hoffmann and Kropp [28] combined the tyre-noise model developed within *SPERON* [3] with the *LISTEN* auralization engine with the purpose of conducting perceptual evaluation of tyre-road noise. The approach enabled the simulation and auralization of pass-by (constant speed), with a single source assumed to be equivalent to four tyre-road sources. Perceived pleasantness, loudness, roughness and sharpness of the auralization was evaluated for 50 km/h vehicle speed, with listening tests showing a good correlation between auralized and recorded sounds for all perceptual evaluations but sharpness. As on the previous study, no binaural cues were available during presentation of the auralized sound.

As the most common methods for assessing tyre-road noise are the Close Proximity – CPX and Controlled Pass-By – CPB methods. Anfosso-Lédée [2] investigated the relation between these, developing a “propagation filter” that relates the CPX signals to a point in the far-field, at the CPB location. The developed filter is analytically estimated in 1/3rd-octave bands for typical road surfaces configurations. Investigating the soundfield at the CPX near field measurement, a computational acoustic analysis employing the Boundary Element Method (BEM) was implemented for the estimation of the tyre-pavement source sound power. Validation experiments were carried with velocities ranging from 70 km/h to 100 km/h, with results indicating a fairly good agreement for reflecting pavements, across all 1/3rd-octave bands from 400 Hz to 4000 Hz. Anfosso-Lédée conclusions on the relation between CPX and sound power was corroborated on a later study by Campillo-Davo et al. [5]. Other studies also have addressed the relationship between CPX noise and roadside measurements [8,35].

In a work carried out by Freitas et al. [24] derived CPX-based virtual sounds through HRTF convolution, were used to generate binaural signals as stimuli for experiments with users rating

annoyance. The results showed the modelling proved its suitability for annoyance predictions.

## 1.2. Aims of this study

This study aimed at implementing a method for delivering a binaural auralization of a virtual moving vehicle to an arbitrarily located moving listener. The method is intended to provide an enough degree of realism to allow research on pedestrian/vehicle interactions during situations of street crossing within a controlled environment. Although the method builds on previous ones, it integrates a dynamic motion reacting auralization engine. It thus endows the sound signal delivered to headphones with dynamic binaural cues. Furthermore, the proposed auralization uses a CPX signal as sound source input, with no correspondent pass-by data required, facilitating the quick interchangeability of source signals. The method was developed to be employed on perception-action studies, using audiovisual virtual scenarios to explore pedestrians' crossing behaviour.

## 2. Method for auralization of rolling noise

The proposed auralization routine allows the headphones' rendering of a free-field, mono recorded sound source to a laboratory listener. The sound is perceived as it would naturally be perceived in a field at an arbitrary position relative to the sound source. The routine entails propagation effects from source to receiver, including the human anatomy effects of torso, head and pinna of an average listener. It is then a numerical description of a transfer function that has as input the CPX mono sound source and as output the signal at the listeners' left and right ear-canal entrance. In the current study, the rolling noise produced by a moving vehicle (the source) is auralized to a laboratory "stand-by pedestrian" listener.

### 2.1. Source modelling

On the basis of the auralization procedure is the characterization of the sound source, propagation path (direct and reflected) and sound receiver. As a starting point the rolling noise produced by a single tire is estimated.

#### 2.1.1. Rolling noise spectrum and levels

The rolling noise of a moving vehicle is generated by a combination of mechanisms such as tyre vibrations (mechanical excitation), air pumping and interactions with the car body [48], resulting in an extended source in height that is taken as centred at around 0.01 m from the ground. Considering the size of the sound source and typical distances of vehicle-pedestrian listener, a simplified model of a point source centred at tyre width, at 0.01 m from the ground with spherical sound energy radiation is adopted. As an additional simplification, the total rolling noise from a vehicle is assumed as an equivalent point source located at the geometrical centre of the four tyres (similar to the strategy for auralization adopted by Anfosso-Lédée [2], Freitas et al. [24] and Hoffmann and Kropp [28]). This model assumes that the pedestrian listener is in the far-field, where the sound energy decay with distance obeys to the inverse square law, as described in the second term of Eq. (1) [31]. Analytical calculation of the time-varying sound pressure levels at a generic position 'p' in the far-field resultant from a moving point source, can be described by:

$$L^p(f, t) = Lw(f, t)^{source} + 10 \log_{10} \left( \frac{1}{4\pi \cdot (d(t)^p)^2} \right) - Att(f, t)^{free} \quad (1)$$

with:

$$Att(f, t)^{free} = Att(f, t)^{atm} + Att(f, t)^{excess} + Att(f, t)^{refl} \quad (2)$$

where:

$L^p(f, t)$  = the sound pressure levels at a generic far-field position 'p', for frequency 'f' and time instant 't' (dB);

$d(t)^p$  = the distance from the sound source at time instant 't' (m);

$Lw(f, t)^{source}$  = the sound power spectrum of the noise source at time instant 't' (dB);

$Att(f, t)^{free}$  = the frequency dependent attenuation relative to free-field that accounts for propagation effects at time instant 't' (dB);

$Att(f, t)^{atm}$  = frequency dependent attenuation due to air absorption at time instant 't' (dB);

$Att(f, t)^{excess}$  = frequency dependent attenuation due to ground reflections and temperature gradients at time instant 't' (dB);

$Att(f, t)^{refl}$  = frequency dependent attenuation due to sound energy loss to other reflection surfaces at time instant 't' (dB);

$f$  = 25 Hz to 10 kHz frequency bin array;

$t$  = time array with duration of CPX time signal (44.1 kHz sample rate);

For the estimation of the rolling noise, a Bruel and Kjaer Pulse type 3560-C system and two Bruel and Kjaer type 4189 microphones were mounted on a light passenger electric propulsion vehicle (Renault Zoe with Acoustic Vehicle alerting system turned off), with an arrangement in accordance to that described in EN/ISO 11819-2:2007 for CPX tests (Fig. 1). The vehicle was equipped with Michelin Green 195/55 R16 Tyres. The microphones were fitted with a wind screen and the mounting device is designed to minimize wind turbulence effects. It is assumed that the signal captured by a CPX microphone is predominantly noise emission from the closest tyre-road generation mechanism (minimal cross talk from other noise sources). The leading edge microphone (45° with trajectory direction) signal is used for analysis, as suggested by previous studies Donovan and Lodico [18], so that the best correlation with pass-by signals is obtained. The recorded signals correspond to a driven trajectories/conditions as presented in Table 1 and Fig. 2.

A Bruel and Kjaer type 4128-C head and torso simulator (HATS) is positioned at a "pedestrian listener" reference position, nominated at a perpendicular distance from 1.55 m of the trajectory centreline, aligned with the centreline of the crosswalk, as represented in Fig. 2. HATS is set with a head orientation of 35°(left), and its pinna at a height of 1.65 m. HATS captured signals are recorded with a Bruel and Kjaer sound level meter (SLM) Type 2270, time synchronized to the CPX measurements. Time synchrony for recordings from the two systems is guaranteed by the use of a digital clapperboard.

The field measurements were taken with an electric vehicle, Renault Zoe equipped with Michelin Primacy 4 205/45 R17 tyres, driving over an asphalt concrete pavement (AC14). The ambient temperature was of 8 °C and the surface was dry. Measured background noise levels were of about LeqA = 46 dB. The vehicle was driven by a professional driver and all trajectories followed a straight line. Measurements were carried through night-time in an industrial area in the outskirts of the city of Guimarães (Portugal) to minimize extraneous sources of noise.

The CPX microphone is in the vicinity of the tyre/road interaction patch, deemed to be measuring at a point where the size of the source is not small if compared to the measurement distance, where the far-field assumption is not valid for frequencies below about 1000 Hz. At this position, a reactive sound field that does not radiate to the far-field is captured. Hence, the inverse square

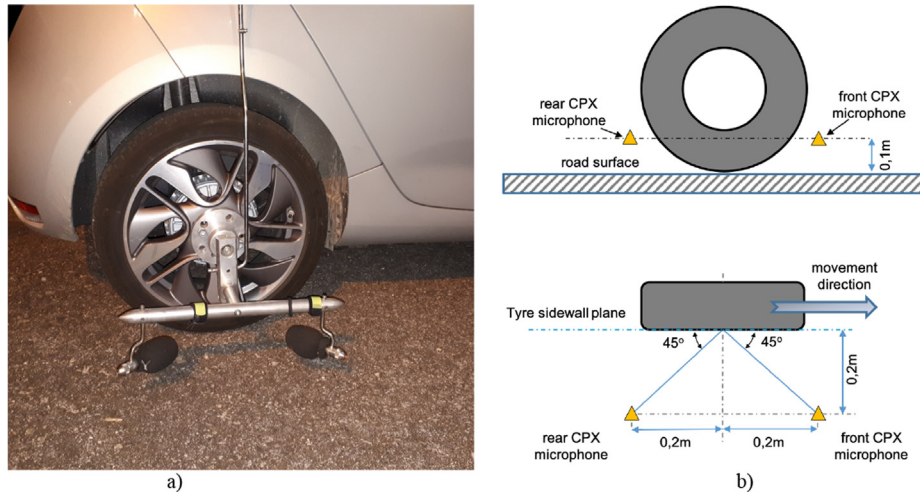


Fig. 1. a) CPX measurement device assemble in the wheel. Two mandatory microphones are mounted and connected to the recording device placed inside the vehicle. b) Scheme representing microphone distances relative to the wheel and ground surface.

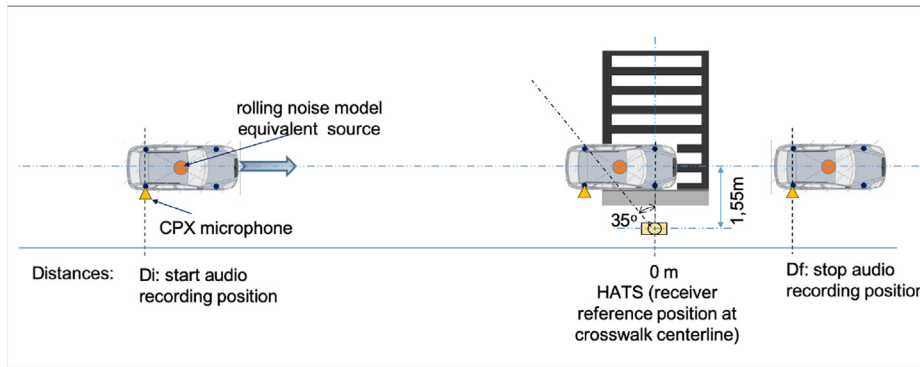


Fig. 2. Scheme of field measurements. The vehicle is driven along a straight-line trajectory (forward direction). CPX microphone is mounted at rear right tyre. CPX recordings and binaural recordings with HATS are time synchronized. Recordings start at the moment the vehicle back axle reaches the Di distance to crosswalk centreline mark, and stop after Df from it.

Table 1

Field measurements trajectories conditions. Vi is initial velocity at the segment of interest; Vf is the final velocity at the segment of interest. Di and Df positions for start and stop recording respectively.

Velocity (nominal)	Description	Vi (km/h)	Vf (km/h)	Di (m)	Df (m)
30 km/h	Electric vehicle at constant speed	30	30	35	11

law for sound energy decay to a far-field receiver is not applicable. To numerically estimate sound pressure levels for a listener in the far-field it is necessary to estimate the tyre-road source sound power  $Lw^{tyre}$ .

The derivation of  $Lw(t)^{tyre}$  from the nearfield CPX measurement requires an analysis of the near-field radiation pattern (such analysis is beyond the scope of this study). In a previous study carried out by Anfosso-Lédée [2], which adopted a similar simplified source-receiver model to investigate the propagation from a point source at the tyre/road contact to the CPX microphone, a detailed acoustical analysis was performed. The results indicated that for a reflective pavement surface as asphalt concrete, the tyre/road noise sound power  $Lw(t)^{tyre}$  may be estimated applying an attenuation of 6 dB across the frequency range to the CPX measured levels (Fig. 3).

As mentioned before, the source model considers an equivalent source at the geometric centre of the four tyres. Each tyre sound

power contribution to the level at this position amounts to an equivalent source of sound power spectrum  $Lw(f, t)^{eq}$ :

$$Lw(f, t)^{eq} = 10 \log_{10} (4 \times 10^0 \cdot 0.1 \times (Lw(f, t)^{tyre} - 10 \log_{10} (\frac{1}{4\pi \cdot (d(t)^g)^2}))) \quad (3)$$

where:

$Lw(f, t)^{tyre}$  = sound power spectrum of tyre-road source at time instant 't' (dB);

$d(t)^g$  = the distance from the tyre to the geometric centre at time instant 't' (m);

It is noteworthy the dependence of electric vehicles' measured sound power and velocity, which is numerically described with regression equations from velocities as low as 20 km/h, indicating

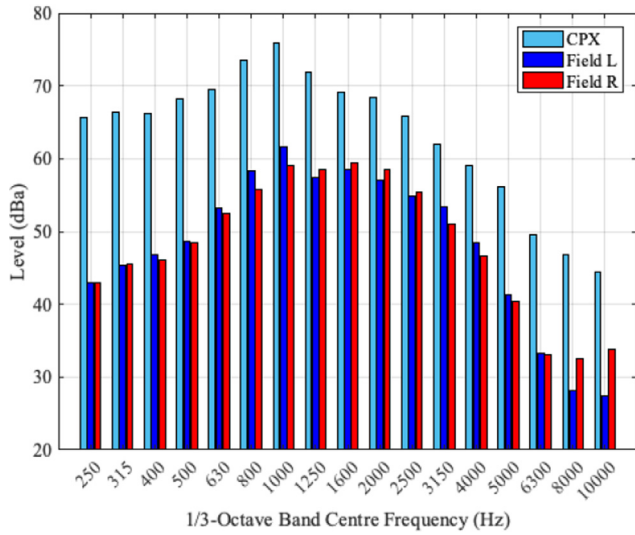


Fig. 3. A-weighted 1/3rd octave band pressure levels of the signal captured by the CPX microphone vs Hats left and right ears captured signals.

a negligible contribution of the propulsion mechanism even at lower speeds [56].

2.1.2. Source directivity and ground effects

A directivity pattern is assigned to the rolling noise source power spectrum. The horizontal directivity accounts for noise generation mechanisms (such as the horn effect, simplified as symmetrical for front and backwards), while the vertical directivity accounts for shielding effects of the car body. The directivity is characterized by a 1/3-octave band equalization, as proposed in the Harmonoise model [31].

Ground effects are of particular importance for sources close to the ground. For non-porous pavements, the ground effect results in a significant increase in sound levels at the receiver. At the same time, it may also, at certain frequencies, result in notches due to destructive interference, causing a distinctive coloration of the sound. The ground effects' calculation for the Harmonoise engineering model is taken as reference and applied to the auralization routine with the single case of flat ground. Straight ray-based (image source) and linear acoustics are assumed, where the effects of

the reflection can be separated from the source signal. The ground reflected path is calculated from the relative positions of the source-listener. The finite acoustic impedance of the pavement is described by the Delaney and Bazley method, using a single parameter (flow-resistivity) model, here a dense asphalt with  $2 \times 10^8 \text{ Nsm}^{-4}$  flow resistivity is assumed [14,16], Fig. 4 plots the calculated attenuation in function of car distance for a receiver at the reference position (the data is processed in 1/3rd oct. bands from 25 Hz to 10 kHz, for better readability only oct. bands from 315 Hz to 5 kHz are plotted).

2.1.3. Auralization routine

The auralization routine (Fig. 5) is implemented in *Cycling74 Max/Msp* environment coupled with the freeware *BlenderVR* add-on [32]. The *Max* patch receives coordinates of source and receiver via the "Open Sound Control (OSC)" protocol. Receiver coordinates are, in the case of this study, tracked by a *Vicon* motion tracking system and sent through the *BlenderVr* add-on to a *Max spatial* audio processor. Source coordinates are sent from a *Blender* "virtual world" model, incorporating an animated vehicle trajectory. The auralization routine reacts dynamically to receiver motion and head orientation.

Taking as input the coordinates of virtual sound sources and receiver, the patch executes real-time attenuation of a monaural signal, according to Eq. (2). The *Max* patch uses the Ircam's *Spat* [6] patch for audio spatialization of the source signal. The binaural signals' directional auditory cues are obtained from real-time convolution of the monaural CPX recording with the corresponding direction of arrival HRTFs.

2.2. Auralization validation

Analysis of the auralization output was conducted in a laboratory at the University of Minho. A standard personal computer was used to reproduce the field captured CPX files running through the auralization *Max* patch and rendered to HATS binaurally with headphones (Sennheiser HD-600). The reproduction system was calibrated with pink noise, using a Bruel&Kjaer HATS and a sound level meter Bruel&Kjaer Type 2270. Unnecessary laboratory equipment was turned off to minimize background noise, but this was not possible for HVAC systems, that were kept running during the validation measurements. Two validation procedures were carried out to evaluate the auralization binaural signal.

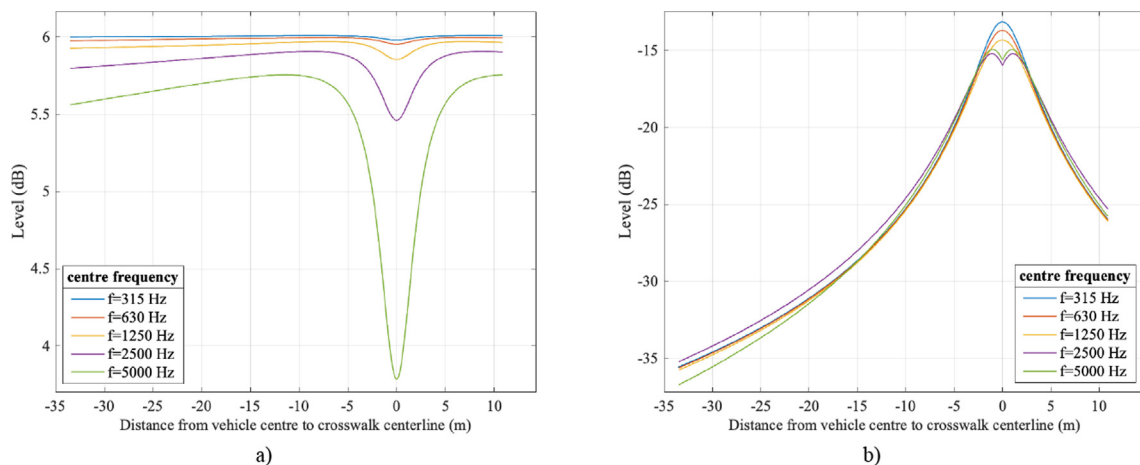


Fig.4. a) Harmonoise analytically estimated attenuation levels due to ground reflections for oct. bands from 315 Hz to 5 kHz. b) Harmonoise estimated SRTF including ground reflection directivity, energy spreading and air absorption to apply to a rolling tyre noise omnidirectional source power as propagated to a receiver at reference position (oct. bands from 315 Hz to 5 kHz). Closer to the reference position the attenuations tend to become significant for the higher frequency range. (due to reflections, car body shielding and horn effect).

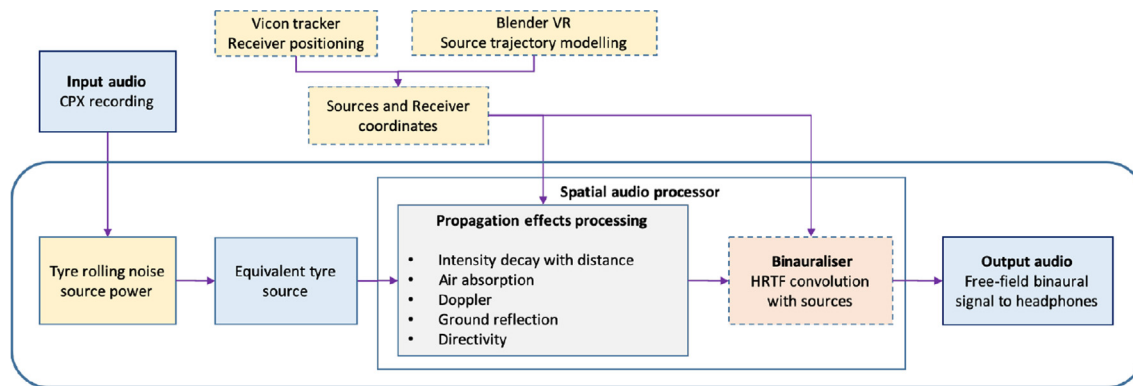


Fig. 5. Schematic representation of the auralization routine. The integrated processing of CPX data for estimation of the resultant free-field binaural signal is shown inside the rounded corners box. Dashed outline boxes correspond to processes not in the scope of this document, thus a thorough description is not considered.

2.2.1. Validation procedure 1: Objective comparisons

The ultimate goal of the auralization engine is to reproduce a synthesized signal that is perceptually similar as to one occurring in the field. Physically identical signals would warrant the perceptual similarity, but due to technical limitations, such feature is generally unattainable. Nevertheless, an objective comparison between spectra of recordings from the field HATS and equivalent condition CPX-derived signals is worth carrying as to enable repeatability and for quantitatively gauge differences and respective perceptual impact. The CPX-derived signals are delivered to headphones and recorded through the HATS-SLM system, positioned at the reference position in the virtual world (tracked by Vicon). This CPX-derived virtual signal recording is gain adjusted to the field signal, for the point of interaural level differences (ILD) equal to zero. At this point, the source is located in the listener median plane (theoretically, free field condition). For this first task, an objective approach is taken to quantify differences between the A-weighted SPL in 1/3rd-octave bands of the two signals.

2.2.2. Validation procedure 2: Psychoacoustic analysis

Relying on purely objective acoustic parameters has proven insufficient for description and quality evaluation of soundscapes and perception related processes of human hearing were considered [25]. Arguably, although two signals may not be physically identical, if their perception is similar, a similar behaviour response is evoked. Thereby, maintaining the same setup as in the previous procedure, in this second task, psychoacoustic models are employed for sound perception analysis. The models' output is an estimative of human participants subjective responses. Estimated loudness, sharpness and roughness from the synthesized signal are set against those estimated from the binaural field recordings as to evaluate perception related similarities between the two recorded signals.

Loudness is a fundamental dimension of auditory perception. It is attributable to the sensation of sound intensity (related to A-weighted SPL), it is then a psychological quantity. Loudness models were developed with the aim of explaining and predicting sound perception by humans, and are generally regarded as the most important psychoacoustic model, with sharpness roughness and fluctuation strength models partially based on it [22,47]. Loudness contains temporal and spectral components, it may be interpreted dynamically in time (momentary loudness), or globally, at the end of the auditory event (global loudness, represented by a scalar) [43]. Both ears contribute to loudness, and the model proposed by Moore and Glasberg [39] for binaural loudness (time-varying) was employed for the presented analysis.

In the perception of an approaching vehicle from a pedestrian point of view, loudness and its rate of growth have a major role on judgments of time-to-contact (TTC), which are of particular importance on decision making for street crossing [17,26,27,38]. See Fig. 6 for an illustration of TTC.

Sharpness impression has been identified to correlate well with perceived annoyance ratings, and, conjointly with roughness, are important attributes on the description of machinery and traffic rolling noise [33,28]. Sharpness sensation is sensitive to high-frequency content, increasing with the growth of energy proportion in the high end of the spectrum. It may be regarded as the perceptual equivalent to the spectral centroid [20]. On the perception of an approaching sound emitting source, one cue that has revealed important to judgements of proximity is the high frequency content variation. The reason for this is that high frequencies are easily attenuated over distance, hence, these frequencies get progressively audible as a source approaches a listener [26]. Roughness sensation is produced by rapid temporal fluctuations in loudness, consequence of modulations in frequency and/or amplitude. Increasing roughness is correlated to an increase in annoyance, and in general terms, to a more noticeable sound [52]. In the current analysis sharpness has been evaluated implementing the Chalupper and Fastl [9] model, and for roughness, the Daniel and Weber [11] model.

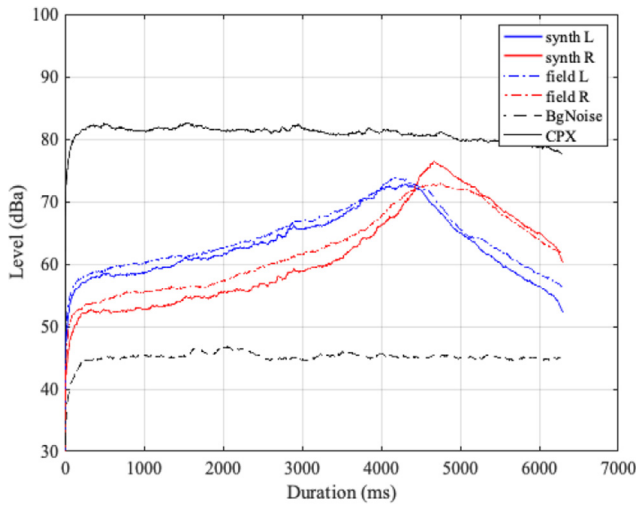
3. Results and discussion

3.1. Objective comparison

In this section, plots from the analysis of the CPX based synthesis and field signals are presented. In Fig. 7, the SPL along time from the recordings of the virtual signal and the field signal is shown. The crossing point indicates that the right and left ear levels are



Fig. 6. Time between a moving object on a collision path to an observer may be defined by a relation between its distance and velocity. Frequently, the terms time-to-contact (TTC), time-to-arrival (TTA) or time-to-collision (ITC), are employed for this measure. Schematic representation in the context of the current study.



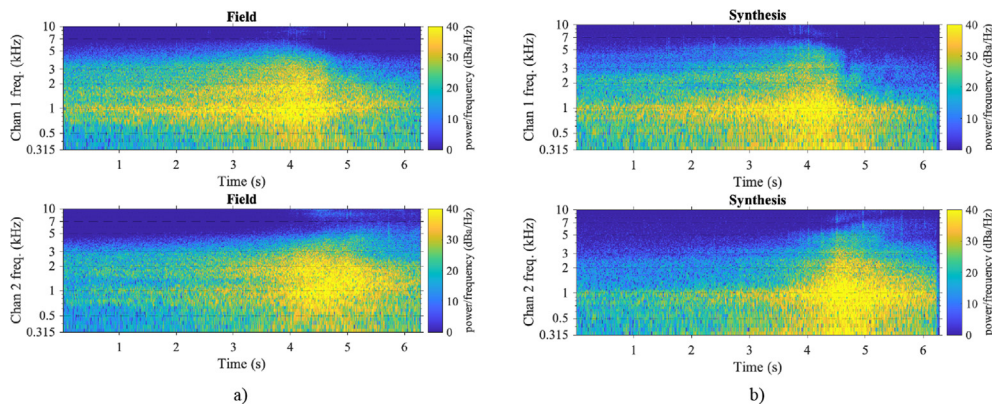
**Fig. 7.** A-weighted SPL synthesis, field and CPX: The plot shows the SPL over time during a crossing trial for audio signals correspondent to the synthesized (full line), field CPX (black line), field HATS recordings (dashed line), and field measured background noise (black dashed). The intersection between left (blue) and right (red) channels levels occur at the position/time of  $ILD = 0$  dBa. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the same, and it is approximately the moment where the source(s) were directly in front of the listener's rotated head, with interaural level difference (ILD) of zero (approximately, as reflections from vertical planes are not accounted for). The auralised signal is level-equalized to the field signal at this point. This equalization

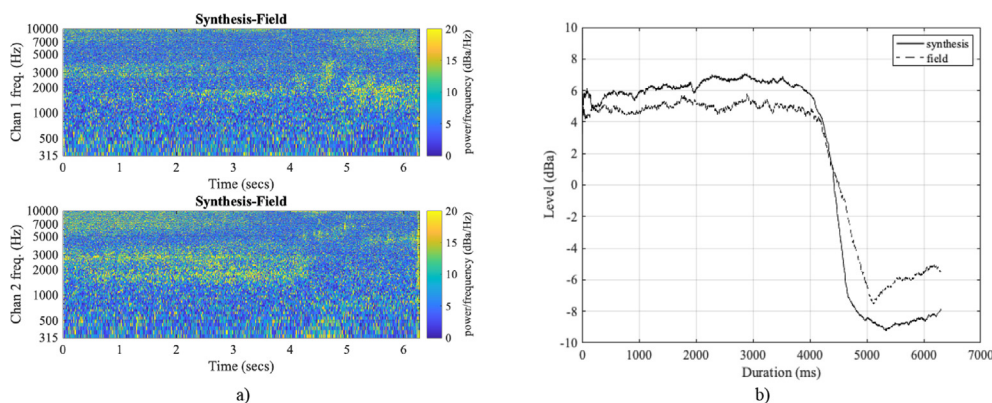
may be partially justified by model simplifications, including that in the synthesis sound, shadowing by the car body is not comprehensively taken into account (particularly from the two tyres on the farthest side). Also, a single reflective ground surface is modelled, and a point source assumed. Fig. 7 shows that the rate of level changes, increasing to about 4500 ms (time the vehicle front axle reaches the crosswalk midline) is similar between the synthesized and field recordings.

In the vicinity of the  $ILD = 0$ , the differences resulting from the choice of considering a single equivalent source for the propagation modelling are apparent on the greater rate of increase on the sound level of the synthesized right ear signal, and its higher peak. It's worth noting that with a head rotation of  $35^\circ$ , the right ear is the one that is closer to the passing vehicle, with the left ear slightly shadowed by the head. The reduced level decrease after the  $ILD$  point in the field measurement evidences the physical distance between vehicle tyres. Unexpectedly, for the field recording, the head rotation is not evident, with the right ear having about the same SPL as the left. This is not explained by a simple binaural theoretical model, and the reflections from adjacent buildings, as well as background noise, have to be taken into consideration to better understand this apparent contradiction.

The aforementioned remarks are confirmed by observing the spectrograms from Fig. 8. The field signals are smoother and the effect of head shadowing less prominent, with the virtual signal resembling a theoretical point source binaural model behaviour, showing a discernible head shadowing. The vertical (broadband) spikes that appear on the spectrograms result from dust particles hitting the CPX measuring equipment, indicating that the quality of the virtual signal is directly dependent on the CPX field recording quality. Fig. 9(a) illustrates the spectral content differences



**Fig. 8.** a) Spectrogram (power spectral density dB) of a-weighted audio signal correspondent to HATS field recordings and b) correspondent to virtual listener recordings.



**Fig. 9.** a) Power spectral density absolute difference between synthesized signal and field signal. The plot illustrates where in frequency and when in time major differences occur. b)  $ILD$  of both signals (a-weighted).

between the two recordings, showing that somewhat larger differences occur in the head shadowed signal, a fact explained by background noise and the existence of reflections. (HATS is oriented on an angle, as a pedestrian looking towards oncoming traffic; thus the spectrogram is not symmetric about the moment of crossing, around 4 s). The plot shows that the right “virtual ear” lacks energy below 2000 Hz (approximately), during vehicle approaching, explained by the head shadowing from a point source and absence of lateral reflections on the theoretical modelling. Fig. 9(b) plots the ILD for the compared signals, with both agreeing on the general curve shape. ILDs are greater for the synthesized signal, which apart from the ground reflections, assumes free field condition; hence an expected greater ILD than in the field measurement.

In investigating acoustic time to arrival (TTA) estimations), previous studies identified that judgments are best if based on cues from the central region of the spectrum (from 350 Hz to 5000 Hz) [26]. In Fig. 10(a), synthesized and field signals energy content plot shows that differences in the 400–6300 Hz range are generally less than 5 dB. Outside this range, the discrepancy between the two signals increases, with the auralization process overestimating energy on the low and higher ends of the spectrum. It is worth noting that specific recording location characteristics may significantly influence the spectral pattern, namely field reflections on concrete/steel walls from nearby buildings would affect the higher end, while background noise and wind turbulence on the CPX microphone affect the lower end. Car body reflections’ effects on the CPX may have an influence of about 1–2 dB on the 315–4 kHz range [18]. As reported before, dust particles resulting in broadband peaks would also have an influence (see spectrograms). Fig. 10(b) graphs spectra of the binaural field and synthesized signals in comparison to the acquired CPX signal. From visual inspection, a correspondence between the three is apparent, that is more evident for the CPX and synthesis signals, reinforcing the suggestion that the synthesized signal is directly dependent on the CPX capture.

The objective analysis shows the differences between the temporal and spectral structures of the analysed recordings. This result is not unexpected and the question of how this measurable difference relate to human perception is investigated in in the psychoacoustical analysis discussion presented next.

### 3.2. Psychoacoustic analysis

The perception of distance and velocity of a moving sound source in free-field involves complex processes. Few studies have investigated this matter. However, a consensual conclusion is that temporal variation on intensity, ILD, and ITD, as well as Doppler

shift, and spectral changes due to air absorption are all relevant cues for distance and velocity estimates, with some studies finding loudness (perception of intensity) as the main cue [44,53].

Loudness and sharpness were estimated for CPX-derived signal and field HATS recording. Loudness model estimates are shown for each signal on Fig. 11. The plots show specific loudness, representing loudness of different auditory filters (spectral components), as well as ‘short-term’ (momentary), and ‘long-term’ time-varying loudness (temporal components). The short-term loudness (STL) is representative of the perceived loudness at any moment. On the other hand, long-term loudness (LTL) involves memory, remaining briefly available after sound termination and it relates to global loudness [39,43].

Comparing time-varying instantaneous specific loudness patterns as a function of time, the field recording (Fig. 11) yields a smoother representation of loudness, with an agreeing representation of its pattern in function of auditory filters. On the other hand, synthesized signal loudness representation (Fig. 11) reveals a distinctively more pronounced peak centred at around 7erb (low frequency energy). Previous investigations have identified that in a decision making task for road crossing, generally, a pedestrian makes the decision before the approaching vehicle gets within 3 m from the crossing line [17]. This line mark is represented in

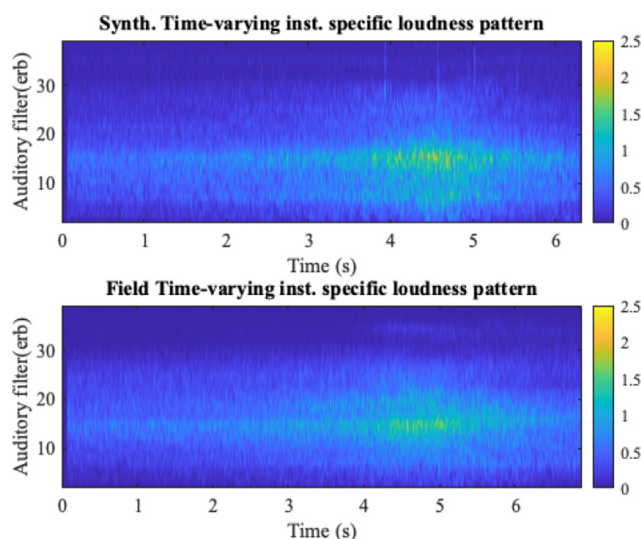


Fig. 11. Synthesis and Field signals’ binaural time-varying instantaneous specific loudness patterns plots.

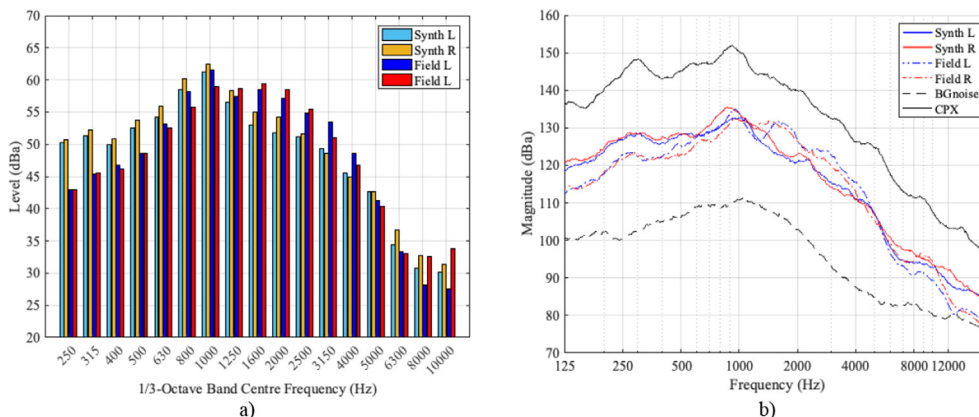


Fig. 10. a) A-weighted 1/3-octave frequency bands level differences between synthesized and field binaural recordings. b) Spectral content superposition plots for background noise, synthesized, field and CPX signals.



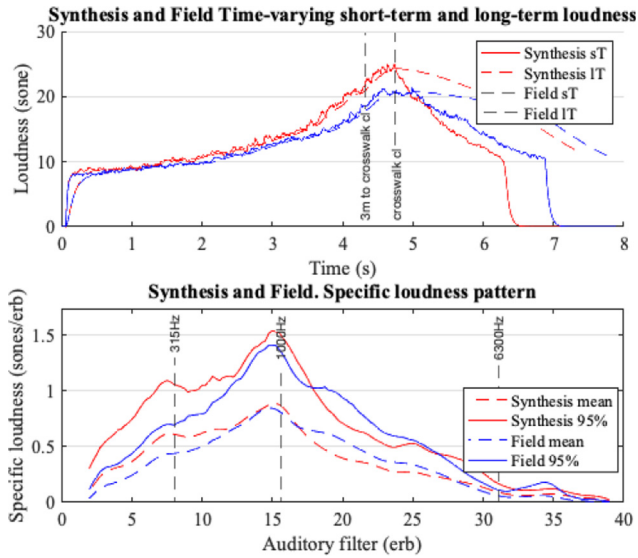


Fig. 12. Synthesized and field binaural time-varying loudness comparison plots (1 sone is defined as the loudness produced by a 1 kHz tone at 40 dB).

Fig. 12, where it is shown that, both signals' short-term loudness grows quite similarly until just before the 3 m mark (4.32 s). The two STL lines maintain a closely matched progression, an important feature for the agreement of distance and velocity estimation [26]. The long-term plot, shows the same agreement, indicating a similar overall loudness perception.

From around the 3 m mark onwards, the synthesis signal results in an overestimation of sound pressure levels (higher values of loudness). This divergence is partially explained by the geometric

Table 2

Global loudness and relative differences obtained from analysis of recordings from 35 m until 3 m (0–4.32 s) vehicle distance to crossing line. (The global loudness value is the maximum from the time-history of short-term values).

Field global loudness 0 s–4.32 s (sone)	Synth global loudness 0 s–4.32 s (sone)	$\Delta$ (%)
21.6	22.8	5.26

inaccuracy from the modelling assumption of an equivalent rolling noise point source. As source/receiver distance decreases, the (angular) direction of arrival from the equivalent point source gets increasingly different from each of the spatially distributed tyre-pavement sources. Consequently, at close distances, the focused sound power of the model results in an overestimation of sound pressure levels relative to the spatially distributed sound power of the field scenario. In Table 2, a single loudness level value is obtained from the listener recorded signal of a vehicle from initial distance of 35 m to 3 m of reaching the crossing centreline (corresponding to analysis of recordings from 0 s to 4.32 s).

Similarly, in Fig. 13(a) time-varying sharpness is plotted, and agreeing conclusions may be drawn. From the start of the trajectory to until just about reaching the crosswalk line, values of sharpness progression are well matched for synthesized and field signals. A single sharpness value, estimated from the recorded signals time history of a vehicle at initial distance of 35 m to 3 m (4.32 s) of reaching the crosswalk centreline is shown in Table 3. This value is calculated taking the mean of the time-history, suggested to be the most representative value for traffic noise perception.

As with sharpness, calculated single values for roughness estimates are shown in Table 4. The resulting roughness estimates yield low values, (generally <0.1 asper), with sporadic peaks in the synthesized signal (Fig. 13b). These results suggest that the model is not able to capture a temporal modulation that may be perceived as roughness. These low values are not unexpected, as the tyre-pavement sound emission is largely unmodulated broadband noise, which does not impart roughness sensation [52]. Despite the low values from the plots, no temporal pattern correspondence is identifiable between the synthesized and field estimate. This fact is clarified if taking into account the synthesis process, as the effect of wind turbulence in the CPX apparatus is inherently different from those in the vicinity of HATS field microphones, contributing to dissimilar temporal structures of the roughness plot. Furthermore, the synthesis roughness plot reveals sporadic peaks, which, when set against the spectrograms of Fig. 8b expose a temporal correspondence with the bursts of broadband energy. As mentioned before, these bursts are attributable to small stones/dust particles hitting the CPX measuring equipment, once again indicating the direct correspondence of the synthesized signal with the CPX recording.

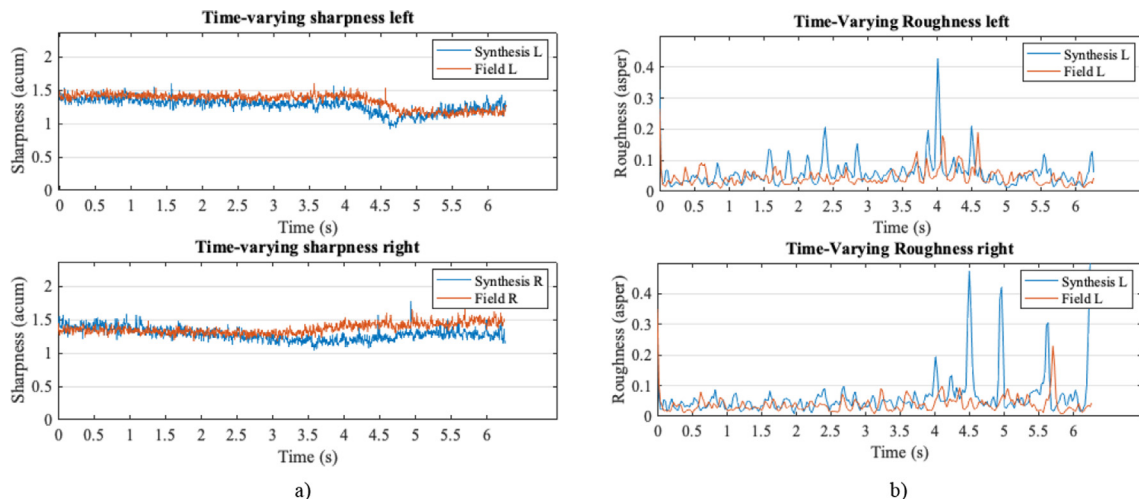


Fig. 13. a) Synthesis and field signals time-varying sharpness (1 acum refers to the sharpness sensation produced by a narrow band noise, one critical band-wide at 1 kHz centre frequency with 60 dB level). b) Synthesis and field signals time-varying roughness. (1 asper is defined as the roughness of a 60 dB 1 kHz sine tone, amplitude modulated at a rate of 70 Hz and modulation depth of 1).

**Table 3**

Summary of sharpness mean values and relative differences obtained from analysis of recordings from 35 m until 3 m vehicle distance to crossing line (0–4.32 s).

Field Left sharpness 0 s–4.32 s mean (acum)	Field Right sharpness 0 s–4.32 s mean (acum)	Synth Left sharpness 0 s–4.32 s mean (acum)	Synth Right sharpness 0 s–4.32 s mean (acum)	$\Delta$ Left (%)	$\Delta$ Right (%)
1.41	1.33	1.32	1.29	6.38	3.01

**Table 4**

Summary of roughness mean values and relative differences obtained from analysis of recordings from 35 m until 3 m vehicle distance to crossing line (0–4.32 s).

Field Left roughness 0 s–4.32 s mean (acum)	Field Right roughness 0 s–4.32 s mean (acum)	Synth Left roughness 0 s–4.32 s mean (acum)	Synth Right roughness 0 s–4.32 s mean (acum)	$\Delta$ Left (%)	$\Delta$ Right (%)
0.05	0.04	0.06	0.05	20	20

The synthesized and field recordings have been further assessed for perceptual similarity using PEMO-Q psychoacoustic validated auditory model [29]. This method simulates the peripheral auditory system signal processing extracting the information for the input of higher cognitive stages. Taking a reference signal and a test signal, the method outputs the PSM\_inst and PSMt measures (instantaneous Perceptual Similarity Measure, and its sequence 5th percentile). PEMO-Q results may be interpreted as an objective audio degradation quality measure from reference to test signal, and have shown to correlate well with subjective ratings.

From the PEMO-Q outputs, the Subjective Degradation Rate (SDG) is estimated [45], which scales as:

- 0: Imperceptible;
- -1: Perceptible but not annoying;
- -2: Slightly annoying;
- -3: Annoying;
- -4: Very annoying.

The instantaneous SDG plot of Fig. 14 shows that there is a tendency for greater perceived degradation from the field to synthesized signal as the virtual source approaches the listener, a consequence from the single equivalent source simplification. Generally, the rate during vehicle approach is above -3, indicating “annoying” judgements at worst. Overall rates are of -2.81 for left and -3.08 for right, but it should be noted that, for the overall rating, the algorithm emphasizes regions of lower rates. It should also

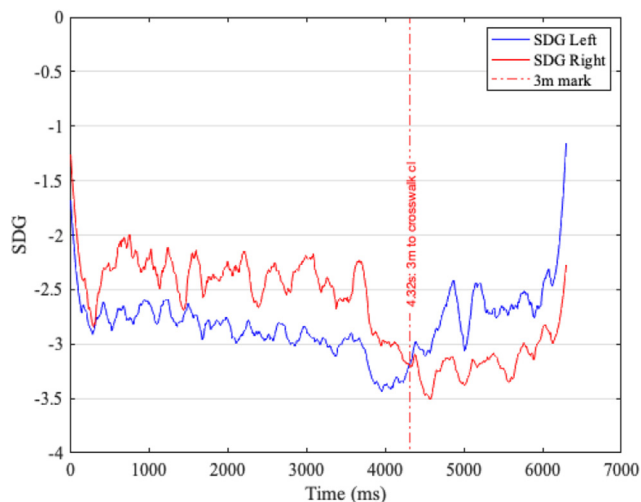
be noted that the 3 m mark, resulting from setup geometry, is in proximity of the position where the vehicle is directly in front from the listener (azimuth = 0, ILD = 0). The crossing of the left and right SDG lines at this moment suggests the weight of the model head shadow discrepancies in the perceived differences between field and synthesized signals.

#### 4. Conclusions

The present paper discusses a method for dynamic binaural auralization of an approaching vehicle, updated in real-time to a motion-tracked listener. The auralization is based on a monaural recording from tyre-road noise. The method’s aim is the capability of synthesizing an approaching vehicle audio signal that is perceptually equivalent to one obtained from field recordings. An objective analysis comparing synthesized and virtual signals revealed that although both signals’ overall spectral and temporal structure is similar, the synthesized signal, due to its simplified theoretical formulation is limited in some details. The apparent enhanced head shadowing from the absence of lateral reflections and a point source formulation result in discrepancies in the lower and higher ends of the frequency range, with the range between 400 and 6300 Hz having lesser differences. To evaluate these differences from a psychoacoustic point of view, loudness and sharpness were assessed. Loudness estimates, as for sharpness, show that the general progression of values until the vicinity of the crosswalk centreline line are well matched. Closer to the crosswalk centreline, and consequently to the receiver, larger discrepancies are obtained. Here, the auralization method is overestimating loudness levels, which is mainly explained by the adoption of a point source radiating the power of four tyres. The signals have been further assessed for a perceptual motivated similarity measure. Results show that listeners would be able to detect differences between signals, which is acceptable.

The method shows promising results, and from the analysis of the validation results, it is deemed capable of allowing laboratory research in pedestrian crossing behaviour resourcing to virtual auditory environments with a freely moving listener. The flexibility that it affords enables for controlled manipulation of conditions, such as the variation of background noise, environmental sounds, source signals, and geometry of scenarios with higher repeatability than field experiments.

On the current implementation, limitations are identified and taken as a simplification that poses no hindrances for the aim of the studied scenario. Limitations include the absence of consideration for barriers or reflections from nearby buildings. Scattering and diffraction from surrounding objects are not calculated, as these are not precluded for the simplified open space scenario of the current application. Vertical temperature gradient and wind



**Fig. 14.** PEMO-Q analysis results of estimated Subjective Degradation Rate (SDG) in function of recording time history. The analysis is based on 10 ms segments.

atmospheric refraction are neglected due to the short propagation distances involved (these effects would be relevant for larger distances) [23]. The most impacting simplification is seen as the assumption of a point source for the vehicle rolling noise equivalent source. Nevertheless, the model proved to be able to recreate the spectral and temporal patterns characteristic from the sound of an approaching vehicle. Analysis results attest to the dependence of the synthesized signal patterns from the acquired CPX signal.

Drawing from the conclusions of the present study, a number of implementations for the method's next iteration are planned:

- Upgrade of the source description to four-point sources, spatially distributed accordingly to vehicle dimensions.
- Enhance propagation mechanisms by integrating the effect of road surface characteristics, namely sound absorption.
- Attending that tyre-road noise sound power is dependent on velocity and that it is directly reflected on the CPX recording sound pressure levels, the method may be applied to trajectories with varying velocities.
- During measuring, protection of the CPX measuring device from stones/dust particles, as to improve the quality of source signals.

### Role of the funding source

This work was supported by the “Fundação para a Ciência e a Tecnologia” [PTDC/ECM-TRA/3568/2014, SFRH/BD/131638/2017, UIDB/04029/2020]

This work is part of the activities of the research project AnPeB – “ANalysis of PEdestrians Behaviour based on simulated urban environments and its incorporation in risk modelling” (PTDC/ECM-TRA/3568/2014), funded by the “Promover a Produção Científica e Desenvolvimento Tecnológico e a Constituição de Redes Temáticas” (3599-PPCDT) project and supported by the “European Community Fund FEDER” and the doctoral scholarship SFRH/BD/131638/2017, funded by “Fundação para a Ciência e a Tecnologia (FCT)”.

### CRediT authorship contribution statement

**Frederico Pereira:** Conceptualization, Software, Validation, Writing - original draft. **Francisco Soares:** Data curation, Investigation, Writing - review & editing. **Carlos Silva:** Funding acquisition, Investigation, Writing - review & editing. **Emanuel Sousa:** Data curation, Investigation, Writing - review & editing. **Elisabete Freitas:** Funding acquisition, Project administration, Conceptualization, Writing - review & editing.

### 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.

### References

- [1] Adminaité-Fodor, Jost. How safe is walking and cycling in Europe? PIN Flash Report 38 (pp. 73). Brussels, Belgium: European Transport Safety Council; 2020.
- [2] Anfosso-Lédée. Modeling the local propagation effects of tyre-road noise: propagation filter between CPX and CPB measurements. In: Paper presented at the Inter-Noise, Prague, Czech Republic; 2004, 22-25 August.
- [3] Beckenbauer, Kuijpers. Prediction of pass-by levels depending on road surface parameters by means of a hybrid model. In: Paper presented at the Inter-Noise, The Hague, Netherlands; 2001, 27-30 August.
- [4] Begault, Wenzel, Anderson. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *J Audio Eng Soc*, 2001; 49(10): 904–916.
- [5] Campillo-Davo, Peral-Orts, Campello-Vicente, Velasco-Sanchez. An alternative close-proximity test to evaluate sound power level emitted by a rolling tyre. *Appl Acoust* 2019;143:7–18. <https://doi.org/10.1016/j.apacoust.2018.08.017>.
- [6] Carpentier. A new implementation of Spat in Max. In: Paper presented at the 15th Sound and Music Computing Conference (SMC2018), Limassol, Cyprus; 2018, 2018-07-04.
- [7] Cavallo, Dommès, Dang, Vienne. A street-crossing simulator for studying and training pedestrians. *Transp Res Part F: Traff Psychol Behav* 2019;61:217–28. <https://doi.org/10.1016/j.trf.2017.04.012>.
- [8] Cesbron, Klein. Correlation between tyre/road noise levels measured by the Coast-By and the Close-Proximity methods. *Appl Acoust* 2017;126:36–46.
- [9] Chalupper J, Fastl H. Dynamic loudness model (DLM) for normal and hearing-impaired listeners. *Acta Acust United Acust* 2002;88(3):378–86.
- [10] Charron, Festoc, Guéguen. Do child pedestrians deliberately take risks when they are in a hurry? An experimental study on a simulator. *Transp Res Part F: Traff Psychol Behav* 2012;15(6):635–43. <https://doi.org/10.1016/j.trf.2012.07.001>.
- [11] Daniel P, Weber R. Psychoacoustical roughness: Implementation of an optimized model. *Acta Acust United Acust* 1997;83(1):113–23.
- [12] de Clercq, Dietrich, Núñez Velasco, de Winter, Happee. External human-machine interfaces on automated vehicles: effects on pedestrian crossing decisions. *Hum Factors* 2019;61(8):1353–70. <https://doi.org/10.1177/0018720819836343>.
- [13] Deb, Carruth, Sween, Strawderman, Garrison. Efficacy of virtual reality in pedestrian safety research. *Appl Ergon* 2017;65:449–60. <https://doi.org/10.1016/j.apergo.2017.03.007>.
- [14] Defrance, Noordhoek, Plovsing, Jonasson, Premat, Aballea, et al. Outdoor sound propagation reference model developed in the European Harmonoise project. *Acta Acustica united with Acustica* 2007;93(2):213–27.
- [15] Del Pizzo A, Teti L, Moro A, Bianco F, Fredianelli L, Licitra G. Influence of texture on tyre road noise spectra in rubberized pavements. *Appl Acoust* 2020;159:107080.
- [16] Delany ME, Bazley EN. Acoustical properties of fibrous absorbent materials. *Appl Acoust* 1970;3(2):105–16. [https://doi.org/10.1016/0003-682X\(70\)90031-9](https://doi.org/10.1016/0003-682X(70)90031-9).
- [17] DeLucia, Preddy, Oberfeld. Audiovisual integration of time-to-contact information for approaching objects. *Multisens Res*, 2016; 29(4-5): 365-395. doi: <https://doi.org/10.1163/22134808-00002520>.
- [18] Donavan, Lodico. Measuring tire-pavement noise at the source: precision and bias statement. NCHRP Transportation Research Board of The National Academies; 2011.
- [19] EN ISO 11819-2:2017. Acoustics - Measurement of the influence of road surfaces on traffic noise - Part 2: The close-proximity method. Brussels, Belgium: European Committee for Standardization.
- [20] Fastl, Zwicker. Sharpness and Sensory Pleasantness. In: Fastl, Zwicker (eds.), *Psychoacoustics: Facts and Models*. Berlin, Germany: Springer Berlin Heidelberg; 2007. p. 239–46.
- [21] Feldstein I, Dietrich A, Milinkovic S, Bengler K. A pedestrian simulator for urban crossing scenarios. *IFAC-PapersOnLine* 2016;49(19):239–44. <https://doi.org/10.1016/j.ifacol.2016.10.531>.
- [22] Florentine. Loudness. In Florentine, Popper, Fay (eds.), *Loudness* New York, USA: Springer New York; 2011. p. 1–15.
- [23] Forssén, Kaczmarek, Lundén, Nilsson & Alvarsson. (2009). Auralization of traffic noise within the LISTEN project : Preliminary results for passenger car pass-by. Paper presented at the Euronoise, Edinburgh, UK. <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-34998>
- [24] Freitas, Lamas, Silva, Soares, Mouta, Santos. Tyre-road noise annoyance assessment through virtual sounds. In *InTech* (ed.), *Noise Reduction and Control*. Rijeka, Croatia; 2016.
- [25] Genuit, Fiebig. *Psychoacoustics and its Benefit For The Soundscape Approach*. *Acta Acustica united with Acustica* 2006;92(6):952–8.
- [26] Gordon, Russo, MacDonald. Spectral information for detection of acoustic time to arrival. *Atten Percept Psychophys* 2013;75(4):738–50. <https://doi.org/10.3758/s13414-013-0424-2>.
- [27] Hecht, Savelsbergh. (2004). Theories of time-to-contact judgment. In: Hecht, Savelsbergh (eds.), *Advances in Psychology* (Vol. 135): Elsevier; 2011. p. 1–11
- [28] Hoffmann A, Kropp W. Auralization of simulated tyre noise: Psychoacoustic validation of a combined model. *Appl Acoust* 2019;145:220–7. <https://doi.org/10.1016/j.apacoust.2018.08.026>.
- [29] Huber R, Kollmeier B. PEMO-Q—A new method for objective audio quality assessment using a model of auditory perception. *IEEE Trans Audio Speech Lang Process* 2006;14(6):1902–11. <https://doi.org/10.1109/TASL.2006.883259>.
- [30] Jonasson. Acoustical source modelling of road vehicles. *Acta Acust United with Acust*. 2007; 93(2): 173-184.
- [31] Jonasson, Sandberg, Van Blokland, Ejsmont, Watts, Luminari. Source modeling of road vehicles, Deliverable 9 of the Harmonoise project: Swedish National Testing and Research Institute, Borås; 2004.
- [32] Katz, Felinto, Touraine, Poirier-Quinot, Bourdot. BlenderVR: Open-source framework for interactive and immersive VR. In: Paper presented at the 2015 IEEE Virtual Reality (VR), Arles, France; 2015, Mar 23–27.
- [33] Kuwano. Sharpness evaluation of temporally varying sounds. In: Paper presented at the Inter-Noise, Hamburg, Germany; 2016.
- [34] Li T, Burdisso R, Sandu C. Literature review of models on tire-pavement interaction noise. *J Sound Vib* 2018;420:357–445.

- [35] Licitra G, Teti L, Bianco F, Chetoni M, Ascari E. Relationship between Pass By results, CPX ones and roadside long-term measures: some considerations. In: INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 253, No. 6). Institute of Noise Control Engineering; (2016, August). p. 2508–16.
- [36] Meir A, Oron-Gilad T, Parmet Y. Can child-pedestrians' hazard perception skills be enhanced?. *Accid Anal Prev* 2015;83:101–10. <https://doi.org/10.1016/j.aap.2015.07.006>.
- [37] Mendonça C, Freitas E, Ferreira JP, Raimundo ID, Santos JA. Noise abatement and traffic safety: The trade-off of quieter engines and pavements on vehicle detection. *Accid Anal Prev* 2013;51:11–7. <https://doi.org/10.1016/j.aap.2012.10.018>.
- [38] Middlebrooks JC, Green DM. Sound Localization by Human Listeners. *Annu Rev Psychol* 1991;42(1):135–59. <https://doi.org/10.1146/annurev.ps.42.020191.001031>.
- [39] Moore BCJ, Glasberg BR. Modeling binaural loudness. *J Acoust Soc Am* 2007;121(3):1604–12. <https://doi.org/10.1121/1.2431331>.
- [40] Morrongiello BA, Corbett M, Milanovic M, Pyne S, Vierich R. Innovations in using virtual reality to study how children cross streets in traffic: evidence for evasive action skills. *Injury Prev* 2015;21(4):266–70. <https://doi.org/10.1136/injuryprev-2014-041357>.
- [41] Nagamatsu LS, Voss M, Neider MB, Gaspar JG, Handy TC, Kramer AF, et al. Increased cognitive load leads to impaired mobility decisions in seniors at risk for falls. *Psychol Aging* 2011;26(2):253–9. <https://doi.org/10.1037/a0022929>.
- [42] Peplow, Forssén, Lundén, Nilsson. Exterior Auralization of Traffic Noise within the LISTEN project. In: Paper presented at the Forum Acusticum 2011, Aalborg, Denmark, June 26–July 1, 2011. <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-70246> (2011).
- [43] Ponsot. Global loudness processing of time-varying sounds. Université Pierre et Marie Curie. Retrieved from <https://hal.archives-ouvertes.fr/tel-01266653>. (2015).
- [44] Pörschmann C, Störig C. Investigations Into the Velocity and Distance Perception of Moving Sound Sources. *Acta Acustica united with Acustica* 2009;95(4):696–706. <https://doi.org/10.3813/AAA.918198>.
- [45] Recommendation ITU-R BS.1116-3 02/2015, Methods for the subjective assessment of small impairments in audio systems.
- [46] Reid. Virtual reality and the person–environment experience. *CyberPsychol Behav.* 2002; 5(6); 559–564. doi: <https://doi.org/10.1089/109493102321018204>.
- [47] Sahai AK, Snellen M, Simons DG. Objective quantification of perceived differences between measured and synthesized aircraft sounds. *Aerosp Sci Technol* 2018;72:25–35. <https://doi.org/10.1016/j.ast.2017.10.035>.
- [48] Sandberg, Ejsmont. Tyre/road noise: reference book. Kisa, Sweden Informex. 2002.
- [49] Schwebel DC, Gaines J, Severson J. Validation of virtual reality as a tool to understand and prevent child pedestrian injury. *Accid Anal Prev* 2008;40(4):1394–400. <https://doi.org/10.1016/j.aap.2008.03.005>.
- [50] Simpson G, Johnston L, Richardson M. An investigation of road crossing in a virtual environment. *Accid Anal Prev* 2003;35(5):787–96. [https://doi.org/10.1016/S0001-4575\(02\)00081-7](https://doi.org/10.1016/S0001-4575(02)00081-7).
- [51] Soares F, Freitas E, Cunha C, Silva C, Lamas J, Mouta S, et al. Traffic noise: Annoyance assessment of real and virtual sounds based on close proximity measurements. *Transp Res Part D: Transp Environ* 2017;52:399–407. <https://doi.org/10.1016/j.trd.2017.03.019>.
- [52] Sottek R. Modeling engine roughness. *SAE Int J Engines* 2009;2(1):1891–9. <https://doi.org/10.4271/2009-01-2153>.
- [53] Störig, Pörschmann. Investigations into velocity and distance perception based on different types of moving sound sources with respect to auditory virtual environments. *J Virtu Real Broadcast*, 2013; 10(4). doi: <https://doi.org/10.20385/1860-2037/10.2013.4>.
- [54] Watts. Harmonoise prediction model for road traffic noise, Report No. PPR034. Wokingham, UK: Transport Research Laboratory; 2005.
- [55] WHO. Environmental noise guidelines for the European region. Copenhagen, Denmark: World Health Organization Regional Office for Europe; 2018.
- [56] Yamauchi, Yabuno. Measurement of sound power level of electric vehicles in steady low speed travelling. Paper presented at the Inter-noise, Madrid, Spain; 2019.
- [57] Zito GA, Cazzoli D, Scheffler L, Jäger M, Müri RM, Mosimann UP, et al. Street crossing behavior in younger and older pedestrians: an eye- and head-tracking study. *BMC Geriatr* 2015;15(1). <https://doi.org/10.1186/s12877-015-0175-0>.