Thesis for the degree of Licentiate of Engineering

Background road traffic noise synthesis

Georgios Zachos

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers University of Technology
Göteborg, Sweden, 2017

Background road traffic noise synthesis

Georgios Zachos

Lic / Department of Civil and Environmental Engineering
Chalmers University of Technology
Technical report: lic 2017:5
ISSN 1652-9146

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Telephone +46 (0) 31-772 2269

Printed by
Chalmers Reproservice
Göteborg, Sweden, 2017

Background road traffic noise synthesis

Georgios Zachos

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers University of Technology

Abstract

As urban planning evolves along with societal life, new requirements that should be met are rising. One of those, is the inclusion of acoustic quality in urban environments, before, during, or after a planning process. To address this for a yet to be built area, metrics are needed that try to quantify acoustic quality. Currently, due to limited models of auditory perception, it is essential that sound samples for the said environments are produced, in order to be used either for further perception based research, or to provide audio examples that can be judged on a case by case basis. The products from this procedure, auralisations, are usually realised with physically valid models, which simulate physical processes to create realistic sounds. While there are several methods aiming for this, fast and computationally expensive simulations of extended in area urban environments, is still a challenge. This report suggests a method for auralising background traffic noise, produced by cars travelling on roads distant from a listener, usually obscured by buildings or other barriers. To achieve computationally efficient auralisation of these scenarios, some parts of the technique are modelled using simulations of physical processes, while others are more simplified methods. For the former, ground reflection and air turbulence are modelled in detail. Individual pass-by events and the Doppler effect, are not modelled explicitly for each vehicle, but instead an approach that considers the traffic as cumulated noise is used. Based on the long distance limitation, shifts in the frequency domain, combined with modulation transfer functions and controlled coherence between the two channels of a listener, attempt to simulate the Doppler effect, fluctuations due to traffic flow inhomogeneities, and the spatial audio image. The modelling of the source output power of the vehicles, uses recorded data as a basis. For validation, auralisations are tested against mixed output from a previously developed and validated demonstrator.

Keywords: *urban sound planning, auralisation, background traffic noise synthesis*

Acknowledgements

The work included in this report is part of the SONORUS project, which was officially funded by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7 2007-2013, under the Research Executive Agency (REA) grant agreement n. 290110.

Of great importance was the smooth and constructive cooperation with my supervisor and co-coordinator of the project Jens Forssén, who helped me figure out how is research actually conducted. During our constructive discussions I got the chance to swiftly overcome obstacles and shift to more efficient paths. Together with Jens and the coordinator of the project, Wolfgang Kropp, I was provided with input which showed to be invaluable.

Due to the nature of the project, the collaboration with the rest of the participants, and especially the thirteen plus one early stage researchers (ESRs) and colleagues, proved necessary to understand the concept of urban sound planning as a whole, without strict engineering blinkers. Special shout-out to the ESR of the project and colleague Laura Estévez Mauriz, for her multimodal and stoical support.

Research visits and secondments, at the technical university of Eindhoven and the noise and sound division of the environment department (miljöförvaltningen) of Gothenburg's municipality, allowed a pleasant experience of different research and working environments. Regards for this to Maarten Hornikx, and the ESRs Fotis Georgiou and Raúl Pagán Muñoz for the former, and Maria Holmes, Henrik Nystedt and Martin Knape for the latter.

Finally, a salute to all of my colleagues from any post at the division of Applied Acoustics for the hospitable working environment, as well to Lars Hansson for his enthusiastic technical help and involvement. And of course, a gratuitous respect to Börje Wijk and Gunilla Skog, not only for their technical and administrative support respectively, but foremost for their patience and ability to form a good atmosphere for the division.

Contents

List of Figures

Part I

Report

1 Introduction

Evolution through cumulative cultural traditions has enabled humans to develop rich cultures rarely evidenced in other species [1]. It can be said that an outcome of this process is the creation of societies, and furthermore cities when population has sufficiently grown. These, in turn, have to be carefully planned for harmonised functionality, where high mobility is essential. Currently, historic information can be drawn through the development of urban planning dating from the end of the nineteenth century [2], but recently further enhancement of the planning process is shown to be possible with the involvement of acousticians [3], of which the current work is part.

The need for prediction of acoustic qualities in urban environments, has led to the creation of noise maps by which calculation of sound pressure levels over an area using various methods and standards is possible. These maps combined with urban planning metrics and acoustic indicators are frequently used to quantify the effects of noise within a population. For example, the World Health Organisation (WHO) presents a research in which lost living years are calculated using noise maps, self-reported health data, and questionnaires for assessing noise-induced annoyance and sleep disturbance levels [4]. The term "living years" does not only include time not lived, but is a value calculated also considering the quality of life of these years. The noise map data used for that report is based on the indicators L_{day} and L_{night} and L_{den}, which describe average levels over a period of time during the day, so the link between different distributions of noise in time is lost. It should be noted though, that the L_{den} calculation includes penalties for non-daytime hours, and that guidelines for night noise exposure exist as well [5]. But various researches have come to different conclusions on how health effects of extended exposure to noise should be viewed. For example, [6] shows that noise and air pollution should be addressed in conjunction, or according to [7, 8], socio-economic factors should also be considered. Annoyance is a term with high popularity in acoustics, but due to yet uncharted lands of neural and emotional responses triggered by audio stimuli [9], a standardised definition of annoyance which can be quantified does not exist. In works like [10, 11, 12] it can be seen that quantifying annoyance is a process that needs careful design in order to be aware of the bias level of the results. Furthermore, WHO's role to introduce proposals for policy makers, should consider results that cannot be misinterpreted, avoiding a future of inaccurately based policies.

To address these issues, the right tools predicting the impacts of an urban sound environment are needed. Many computational tools are constantly being developed, and each one is fitted best for a specific aim, in order to produce either sound levels over time, or auralisations. Auralisations, products of a process which outputs audible stimuli representing the required sound emitting elements, can be used in urban sound planning as a communication tool, presenting to interested parties future sonic scenarios, but can also be a method that potentially provides insights to the human perception of these stimuli. Bringing together the fields of acoustics, psychology and neuroscience, these insights could be used to create more suitable metrics and indicators that represent physical and emotional responses. But a step is firstly needed, and that is to create models that can produce convincing auralisations.

Several methods have been developed, auralising the dominant element of urban environments, i.e. cars, and their passing through surrounding environments [13], but most of them focus in the micro-scale [14], and become easily computationally demanding when applied to more general situations. This report is focused on a simplified but perceptually valid background traffic noise auralisation, which targets on computationally low demands. What is considered as background here, is traffic that occurs at roads distanced sufficiently away from the listener such that it is difficult for one to distinguish individual vehicle pass-bys.

1.1 Scope

From a both computational and psychoacoustics point of view, it is interesting to discover a minimum level of complexity required from an auralisation model, in order to be perceived as realistic when experienced sonically. This report proposes a simplified model for background traffic noise, where not all physical qualities are explicitly calculated, but are simulated with other ways. Validating the model by using recordings from identical scenarios would be ideal. However, the auralisations simulate flat city scenarios, so acquiring such recordings is nearly impossible. It has been instead subjectively tested against auralisation mixes of car pass-bys from a previously developed demonstrator, a product of the LISTEN project [15]. Considering the vehicle velocity range that the background traffic noise auralisation model aims to auralise, LISTEN has been succesfully validated, using subjective tests, on the perceived similarity between its output and recorded events of the same scenarios, as well on the perceived speed and distance of individual pass-by events. Conducting listening tests against the LISTEN output, the suitability of the model for usage in combination with auralisations of local pass-bys, is investigated. Although the model is partially based on simplified techniques, which do not represent exact simulations of physical processes, results provide insights on its suitability for usage in combination with pass-by events close to a listener.

1.2 Outline

Firstly, an introduction to traffic auralisation is presented in Chapter 2, covering the developed technique for background traffic noise auralisation. There, the important parts that should be included in a model are noted, as well as a brief reference on existing and developed methods. The validation procedure is explained in Chapter 3 presenting results from listening tests. Lastly, Chapter 4 summarises conclusions derived from the validation's results, and suggests possible ways improving such a model.

At the conference Inter-noise 2016, hold in Hamburg, Germany, Paper I was presented, which includes a summary of the developed technique, and preliminary results. Those led to the conduction of more involved listening tests, presented in Paper II, and will be presented at the DAGA 2017 conference, in Kiel, Germany.

2 Introduction to traffic auralisation

A recent definition of auralisation can be bound in [16] where it is described as the methodology of producing audio stimuli using simulated, measured or synthesised data. These data will carry information on mostly physical qualities that are needed for the situation to be auralised. In the case of traffic sounds, the information needed can be broken down in parts and be considered separately for individual vehicles for creating a model.

2.1 Parts involved

When auralising a specific acoustic situation, an important part that usually should be considered is the source from which the audible event initiated. For traffic it is the vehicle, but one can go in further detail. When a passenger car travels with a speed between 30 and 100 km/h, the dominant sound source is the interaction between the tires and the road surface, considering the A-weighted level of the average vehicle [17]. The characteristics of the emitted sound depends on the speed of the vehicle as well as the type of the surface. Above and below that speed range, vibrations originating from the engine of the vehicle, and air intake and exhaust will be the most prominent, thus important to explicitly include in an auralisation process. As the vehicles are not ideal radiating omni-directional sources, their directivity should also be considered.

The first important event a wavefront encounters in its propagation from the source, is the ground effect. Strong interferences between direct and reflected waves, will result in attenuation of the signal in certain frequency ranges [18]. For long distance propagation paths, air attenuation and refraction is important to model as well. Some refraction effects may arise from inhomogeneity of the propagation medium caused by wind or temperature variation. Diffraction will be caused by buildings and barriers, and can be highly influential on the resulting response depending on the structure of the scenario under consideration. Regarding a vehicle pass-by, the Doppler effect is prominent even at relatively low speeds.

Finally, for a wavefront to fulfil its journey from a perceptual audio point of view, it needs to pass through our own sound barriers, the human hearing system. Modelling the response of this system can be exceptionally complicated, even impossible considering the latest advances in research. But luckily, it can be divided in parts: diffraction from the head and torso of the body, filtering due to the pinnae and ear canal, scaling air vibrations by the mechanical structure of the middle ear, transferring those mechanical vibrations to electric impulses with the help of the inner ear, and then processing and distribution of these data by neural centres in order for the physical domain to become mental. Most auralisation applications though, require that sound will be introduced to a person and only model the body diffraction and pinnae, using head related transfer functions, HRTFs.

2.2 Examples of existing and developed methods

As mentioned, several methods have been developed for auralising urban environments. Considering linear acoustics where the modelling stages for source strength and sound propagation are separable, auralisation procedures can treat these stages independently. Here, a sample of these methods that inspired this work will be presented.

Source modelling

As the most prominent source at common driving speeds, descriptions of tyre-road interaction, tyre vibrations and rolling resistance, can be found extensively in literature. For auralisation purposes though, a narrower range of material is available. A model for predicting tyre-road noise levels can be found in [19] which is further described in [20].

When sound radiating from the rest of the vehicle is needed, that is the propulsion source (engine, air intake, air exhaust etc.), other methods can be applied. One possibility that has been tested during this project is to use synchronous granular synthesis of recorded engine run-ups. As this was an exploratory work, recordings have been acquired only from three microphone positions around a stationary petrol car with four cubicles: at the back and front of the car and one at its side, distanced 7.5 m away and placed on the ground to avoid interference from ground reflections. Another microphone was positioned at ear-level height (1.5 m), to compensate for interferences from the body of the vehicle.

With the full time signal available, individual grains must be detected. These will be the core elements of the synthesis process, and will consist of samples that resemble one full combustion cycle. Detection of grains from the recording is realised by an iterative procedure that passes a windowed sinusoid test signal through the data, and finds the position of the highest correlation through convolution. For efficiency and to avoid errors, a search threshold should be set alongside the starting point of the convolutions. For the acquired recordings it has been found that ±5 % of a full cycle period is optimal. As a grain is found, it is stored to a library along with windowed tails before and after the detected grain with duration of one stroke of the cycle. The tails are using hanning windows, and are needed to smoothly transit to the next grain. In Figure 2.1 a grain with the left and right windowed tails is shown. The data's next sample after the last grain sample (without the tail) is used as initial for the subsequent grain detection. Once the full recording has been scanned, a signal can be synthesised by arranging the grains randomly, and a stationary, accelerating or decelerating in respect to engine speed signal can be simulated, while sequential placements of the same grain is avoided. The result can then be combined with tyre-road interaction sounds as has been made in [19, 20].

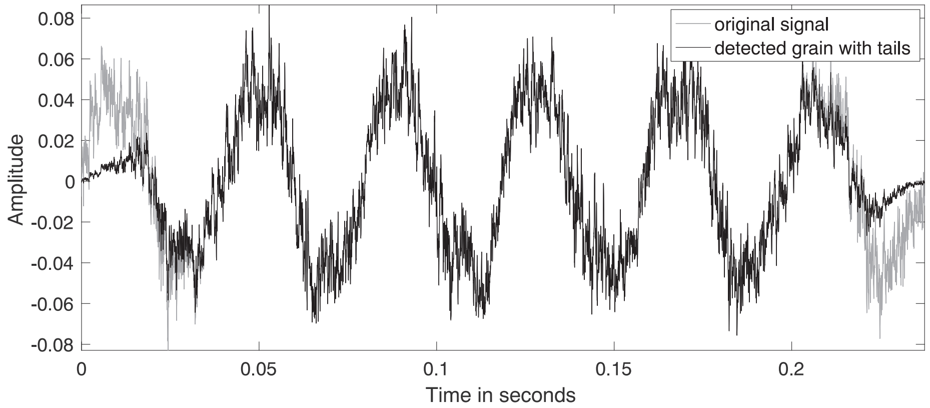


Figure 2.1: Detection of a single grain

Modelling of the propagation path

The proposed method utilises techniques as firstly introduced in [21] and [22], and has been validated with the LISTEN model [15]. This method is described more in detail in [20] and physically describes in the frequency domain, air attenuation, ground effect, spherical spreading, and effects of air turbulence, while it applies Doppler effect in time domain.

Another way to simulate the propagation path has been attempted by the author. The finite-difference time-domain (FDTD) technique, provides realistic simulation of sound propagation, discretising the first order partial differential wave equation in a staggered finite-difference spatial grid (e.g. [23]). The basic idea is that while pressure and velocity points are described on the grid, pressure fluctuations will propagate through the medium governed by the wave equation which can be solved in time with different valid operators. An unstaggered in time technique, using the 4th order Runge-Kutta integration, is chosen due to its stability. Once set, the model can simulate complex environments with any composition of barriers, boundaries, materials and fluctuations. While the model yielded successful results, the approach has been put on hold as the computational effort needed for the scope of the work is too high. Hybrid methods though exist that address computational limitations (e.g. [24]). A summary of computational methods and their suitability for urban sound planning can be found in [13].

2.3 Method used

For the development of the method, a simplified scenario has been considered. It is a flat city case without barriers and obstacles (except ground). The listener is set to the same level as the road with ear level at 1.5 m above the ground. Traffic is accounted to be a single lane of uniform speed for each sound sample. Flow distribution, the headway time between two pass-by events, draws random values following a gamma distribution, based on [25], using a random number generation method according to [26]. Following, the methodology will be described, starting from the source that feeds the model, also found in Papers I and II.

Source noise and Doppler effect

As can be seen in Paper I, for the source model, third octave band car engine power profiles from the LISTEN demonstrator [15] are used. These are parameters extracted from recorded data, but any other input can by put here instead, if provided. The driving speeds at which the data is measured range from 30 to 110 km/h with steps of 20 km/h. After being normalised with respect to loudness [27], the engine power output is translated to third-octave band tenth order filter coefficients for the final listening tests, as described in paper II, of which the response is shown in Figure 2.2.

In order to auralise the static in time power profiles, white noise is fed to the filters. Doppler effect is not simulated in time domain on the propagation stage as usually, due to the form of the model, but a part of it is dealt with here. From a single power profile, two more are created by transferring energy amounts between frequency bands depending on speed, to simulate the Doppler frequency shift of a movement with constant speed as seen in Figure 2.3. The ratio between the Doppler and source frequencies needs to be found, and expressed as a shift in third-octave bands:

$$\Delta f^{\pm}_{1/3} = \frac{f_{Doppler}}{f_{source}} \frac{1}{2^{1/3}} = \frac{v}{v_s} \frac{1}{2^{1/3}} \tag{2.1}$$

where v is the speed of sound in the propagating medium, and v_s the source speed in direction to listener. For programming purposes, the fractional part of $\Delta f^{\pm}_{1/3}$ is considered as the amount of energy that needs to be transferred to or from each frequency band (depending on the direction of the source). The integer part plus one ($\lfloor \Delta f_{1/3} \rfloor + 1$) then, is the number of third octave bands away from the original,

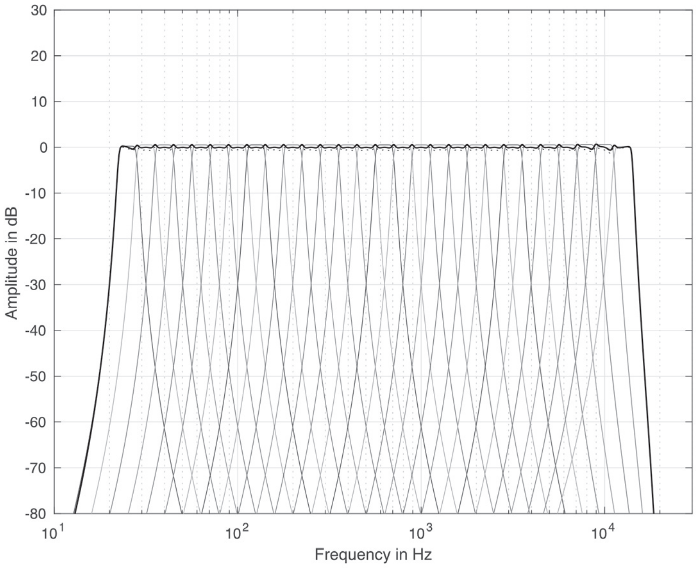


Figure 2.2: Filter response of filter banks

that is needed to be transferred to or from. Depending on the source speed, the shift might be more than one third octave.

While in reality the shift is variable during a pass-by, this simplification considers that the virtual listener is positioned in head-on direction from the approaching and receding vehicles. The aim of this is to simulate an accumulated Doppler effect from a distant traffic which extends on the two sides of the listener, without explicitly modelling each pass-by.

Propagation effects

Up to now, the resulting stimuli will be an enveloped noise, shifted in frequency for the left and right channel, and propagation effects can introduce their way in, namely air attenuation, ground effect, and air turbulence.

This and the following section, processes signals with a time-frequency representation, to set possible to work in frequency domain without needing the full length of the signal a priori of the calculations. A well known method is used, the Short-Time Fourier Transformation (STFT) [28]. For fast implementations of the Descritised Fourier Transform (DFT) [29], the Fast Fourier Transform (FFT) algorithm is used. Given that the signals used are finite dimensional and discrete, the formula for the FFT implementation can be expressed as:

$$X(k) = \sum_{m=O}^{N_{FFT}-1} x(m) e^{-2\pi ikm/N_{FFT}} \tag{2.2}$$

and the inverse, in order to return to the time domain is defined as:

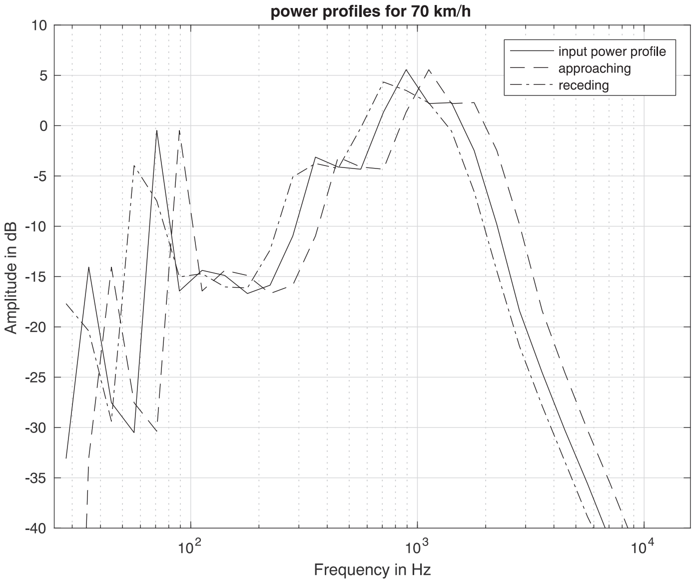


Figure 2.3: Power profiles shifted in frequency to emulate a cumulated Doppler shift

$$x(m) = \frac{1}{\sqrt{N_{FFT}}} \sum_{k=0}^{N_{FFT}-1} X(k) e^{2\pi i k m / N_{FFT}} \tag{2.3}$$

where k denotes frequency, and m time indexes. As the signals processed are not periodic, windowing is needed to avoid aliasing caused from the edges of the sample. A periodic hamming window wth $N_{FFT} = 2048$ samples is used. For the STFT procedure, short pieces of the audio stream with same number of samples as the window function need to be processed. To avoid amplitude variations and for a good reconstruction in time domain, a 75 % overlap of the windows is chosen, meaning that the "hop" size (the period which an FFT is performed) is 512 samples. Window size has an impact on the resolution in frequency domain, as well as on the time domain resolution, when the signal is reconstructed. In the current work, impact like events will not occur, allowing for larger window lengths.

For air attenuation, the standard method [30] is used in third-octave bands. Ground effect and air turbulence is applied as in [20]. This method is also used for the LISTEN demonstrator, so the same environmental and ground parameters have been chosen for validating puproses, that is temperature of 24°C, static atmospheric pressure of 101.325 hPa, and asphalt for the ground with flow resistivity $2 \cdot 10^7$ Nsm^{-4} [31].

Modulation transfer functions

With propagation effects included, the output will resemble closely the targeted traffic, but only in terms of averaged frequency content. To create the perception

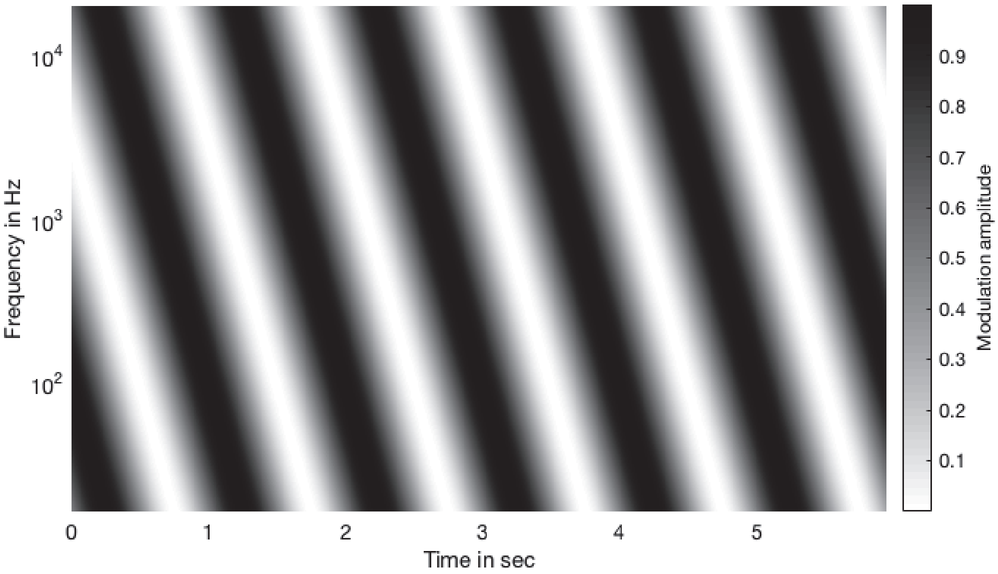


Figure 2.4: MTF pattern ($A = 1$, $\omega = 1$ cycles/second, $\Omega = 0.1$ cycles/octave)

of fluctuation due to individual pass-by events on the audio stream, modulation transfer functions (MTFs) are applied. MTFs are created due to the effectiveness of dynamically rippled spectra, in giving rise to responses at the auditory cortex [32], and have been used in speech intelligibility research [33]. They are orthogonal and described by Equation (2.4) where $x = \log_2(f/f_c)$, with f frequency in Hz and f_c the frequency where a minimum occurs. A describes the amplitude of the ripples, Ω the ripple velocity in cycles/octave, ω the velocity in cycles/second, and ϕ adjusts phase. A is set to 2.5 dB resembling the directivity pattern of rolling noise [34]. A visual representation is shown in Figure 2.4.

$$MTF(x) = A \cdot \sin\left(2\pi \cdot (\omega \cdot t + \Omega \cdot x) + \phi\right) \tag{2.4}$$

As MTFs attempt to resemble individual pass-by events, they are used to shape each channel before applying propagation effects. For more direct control of the functions, ω is set to zero so the ripples remain stationary with in time, while the phase is either increased or decreased in order to move the ripples. Usage of these spatio-temporal functions, serves a double purpose. Depending on the position of a virtual vehicle on the road, the phase of the MTF is updated, such that it will result in ripples travelling upwards on the frequency axis on the channel that vehicles are approaching, and downwards on the channel that vehicles are travelling away. This resembles to pass-by fluctuations both due to changes in the the interference pattern in frequency and amplitude modulations. It also creates the perception of constantly rising (or falling for the opposite channel) of the noisy signal's pitch, like it is experienced with Shepard tones [35]. This then, may serve as well as an approximate simulation of the Doppler effect in combination with the shifted power

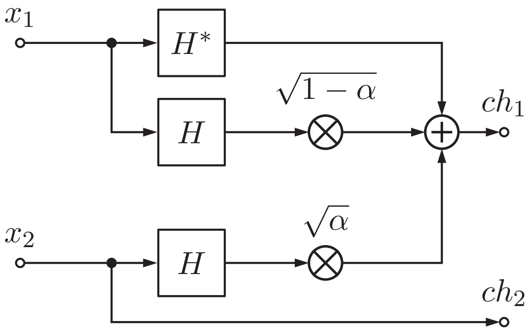


Figure 2.5: Signal flow of the coherence control between left and right channels

profiles (Section 2.3 and Figure 2.3).

Additionally, if $\phi = 2n\pi$ where n is a natural number, the MTFs of the two channels are identical, and while $|\phi|$ is increasing the two functions are traveling in opposite directions. With a proper selection of the parameters f_c and Ω, a motion in the auditory scene is perceived, travelling on the horizontal plane, being in the centre when $\phi = 0$. This motion follows the change of phase of the MTF which, as mentioned before, can be thought as the position of a virtual vehicle. As the reasons behind this are unknown to the author, optimal values for these parameters have been found through screening tests as $f_c = 1.5$ kHz, and $\Omega = 0.3$ cycles/octave.

Stereo image

In spite of the movement perceived due to MTFs, the stereo image of the output is still perceived mainly in the centre. This is to be expected, since for both sounds the same generated noise is used. On the contrary, when using different random number realisations, the stereo image will be as wide as possible [36]. For the case of background traffic noise of this model, two channels are decorrelated in respect to the distribution of vehicles. As described in paper II, the coherence α between the two channels is controlled according to this distribution, and is used as a mixing parameter between the two noisy channels. This means that $\alpha = 1$ (perfectly coherent signals) when a virtual car is in front, and decreases when there is a distribution of vehicles on the left, right, or both sides of the listener, inversely related to their distance form the centre. The signals are passed through the low-pass filter H with a cut-off frequency of 100 Hz (where H^* is the inverse, high-pass, filter), as decorellated signals in low frequencies will result in a confused perception of the spatial image. The flow for this process is shown in Figure 2.5, where x_1 and x_2 are the noise generators (left and right channel), and ch_1, ch_2 are the outputs of the process.

Finally, the signal is brought back to time domain, and amplitude reduction due to spherical spreading is applied.

3 Validation

The method proposed was preliminarily validated with a small number of subjects as described in paper I, and, as seen in paper II, a proper validation procedure, with small changes in the model was conducted. The second validation procedure will be discussed here.

3.1 Subjective tests

For testing, a real-time (or more technically correct, interaction-time) demonstrator is built. Through this demonstrator, dynamic change of most of parameters of the model is possible. The test is divided into two parts. During the first, the subjects are called to perform an A/B comparison of mixed output from the LISTEN demonstrator and the proposed model, on traffic scenarios of four different vehicle speeds, 50 - 110 km/h every 20 km/h, and five different distances, 100 m - 900 m every 200 m, of the listener from the road. In total twenty different scenarios are presented. Prior to the A/B comparison, the subjects are asked to find the optimal positions of two sliders (Figure 3.1), representing the speed and the distance of the model, aimed to match the parameters with the ones of the LISTEN output. The subjects were not informed on the functionality of the sliders, and could freely switch between A and B sounds. When the parameters are chosen so the auralisation is perceived most similar to the LISTEN output, a similarity rating from 0 to 10 is required to continue to the next randomised in sequence scenario. The goal of this design aims to test the similarity of the two models, while the perceived vehicle velocity and traffic distance of the simplified method is assessed. This task required 15-25 minutes, depending on each subject.

The second part is designed to be faster (5-10 minutes) to avoid fatigue of the participants. This time, on top of both the simplified model, and the LISTEN mix resembling background traffic, nearby traffic pass-by events from LISTEN are added,

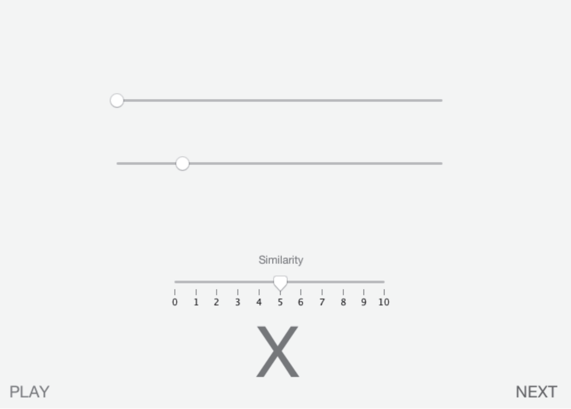


Figure 3.1: Graphical user interface used for the listening test, part 1

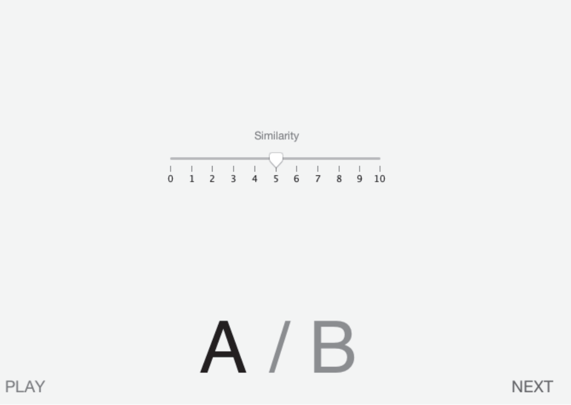


Figure 3.2: Graphical user interface used for the listening test, part 2

at 15 m away from the listener. The task in hand now is to only assess similarity of the background traffic events, while the previously open parameters, are now fixed to values determined by the physical quantities that they represent (Figure 3.2). Masking effects are assumed to lower the ability of the subjects to judge similarity of the two cases, and comparing with the previous results it will be possible to assess wether the model could work on more realistic situations. There were 25 participants, of which 3 exceeded, or did not reach, the predicted time frame completing the tasks, so their answers were discarded.

3.2 Results

To check distributional characteristics of the results, acquired data is presented as box plots. To understand box plots, they can be divided in 4 groups, or else quartile groups. The lower part, a line expanding from the lower limit of the box (the lower quartile) to the lower whisker, represents the twenty-five percent of the answers. Similarly, the upper group that expands from the upper quartile (upper limit of the box) to the upper whisker, shows the twenty-five percent of the answers that fall above it. The box itself can be divided in two other parts by a line denoting the median (middle quartile) of the data. Above that line, 25 % of the answers fall between the median and the upper quartile, and below, 25 % fall between the lower and the middle quartile. Dots outside that range, correspond to data points that have values either larger or less, than 3/2 times of the upper plus lower quartiles. They are grouped by distance in Figures 3.4 and 3.5, and grouped by speed in Figures 3.6 and 3.7. Depending on the grouping, different shadings represent variations either on distance or on speed. The median of the answers of the first part of the subjective test, ranges between 5 and 7, with an overall mean of 6 (of a discrete scale from 0 to 10), while the upper quartile (75 % of the answers) are up to 9. For 75 % of the listeners (upper quartile) ratings reach 8 for most of the speed scenarios. Outliers are observed mainly at 700 m and 900 m away from the traffic (see Figure 3.4). On the second part, where local car pass-by events are present, the median of similarity ratings ranges from 6 to 9, with an overall mean of 8, and the upper quartile (75 % of the answers) can be found on 10.

A clear trend cannot be seen here, but given the distribution of the ratings between different scenarios, it is sensible to perform confidence interval analysis over their overall mean. The 95 % confidence intervals, for the two tests are shown in Table 3.1. The intervals have also been calculated for the results in groups with the same speed profile (see Figure 3.3). A hypothesis test has been performed, where the null hypothesis is that the true mean of the similarity ratings is the same. The null hypothesis is rejected with 0.001 significance (i.e. 99.9 % level of confidence). Parameter matching on traffic speed and distance from the first part of the test, as can be seen on Figures 3.8 and 3.9, present a wide range of answers with no apparent trend.

Test part	Mean	95 % C.I.	C.I. range
1 - w/o nearby traffic	6	6.0 - 6.4	0.4
2 - w/ nearby traffic	8	6.6 - 7.1	0.5

Table 3.1: 95 % confidence intervals of similarity ratings against LISTEN

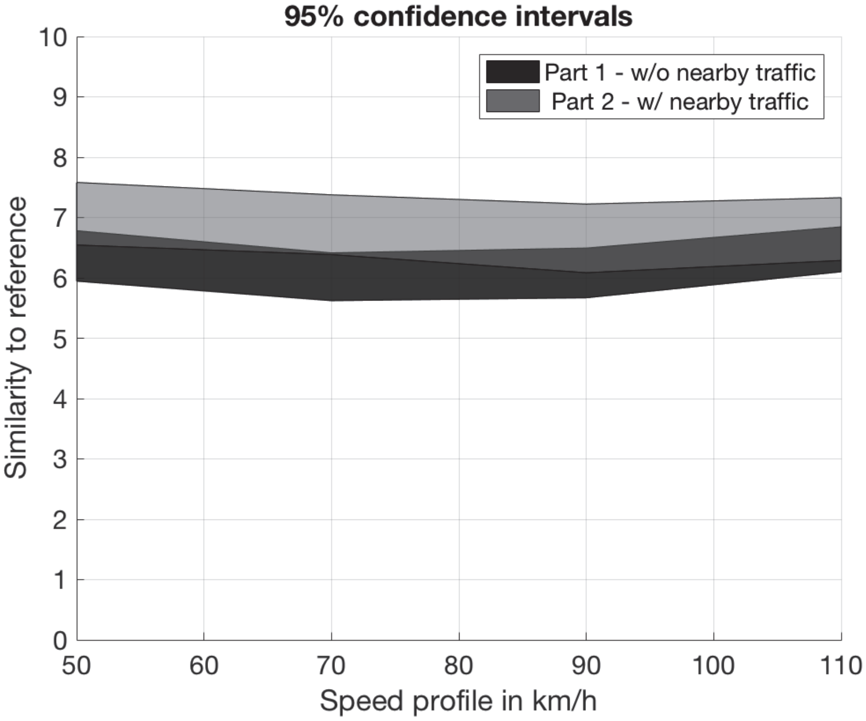
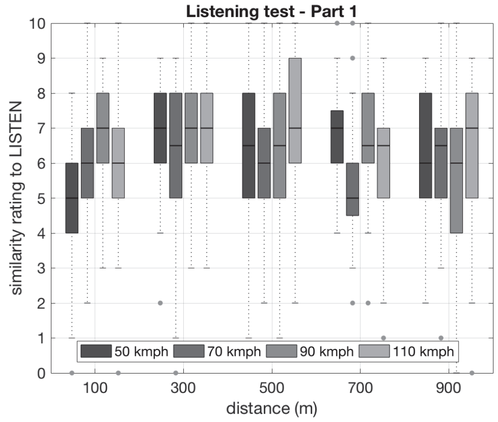


Figure 3.3: 95 % Confidence intervals of similarity ratings



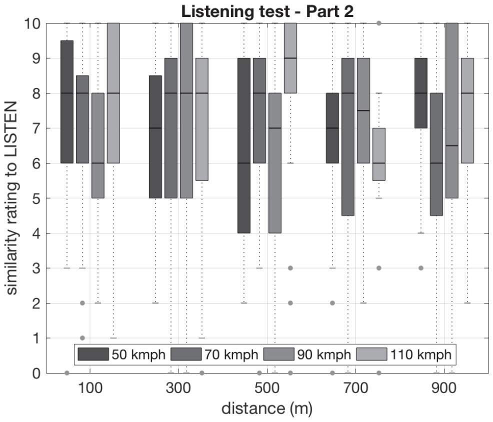
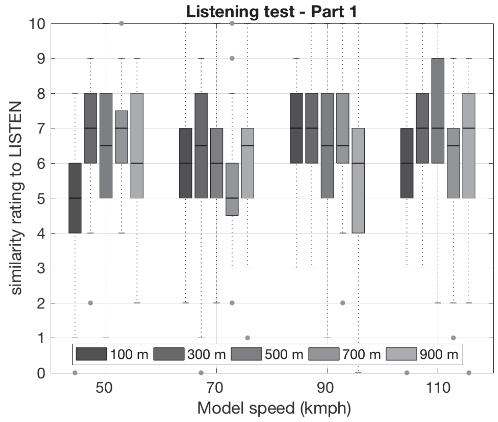


Figure 3.4: Similarity ratings grouped by distance (part 1)

Figure 3.5: Similarity ratings grouped by distance (part 2)



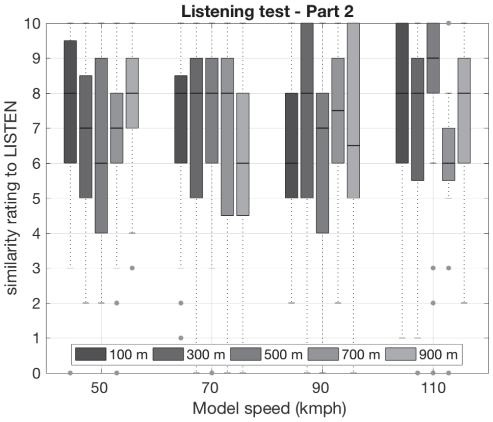
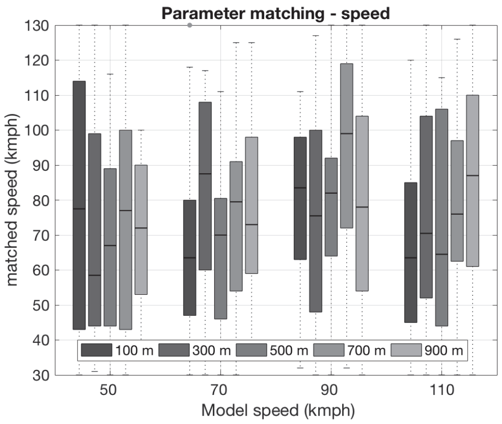


Figure 3.6: Similarity ratings grouped by speed (part 1)

Figure 3.7: Similarity ratings grouped by speed (part 2)



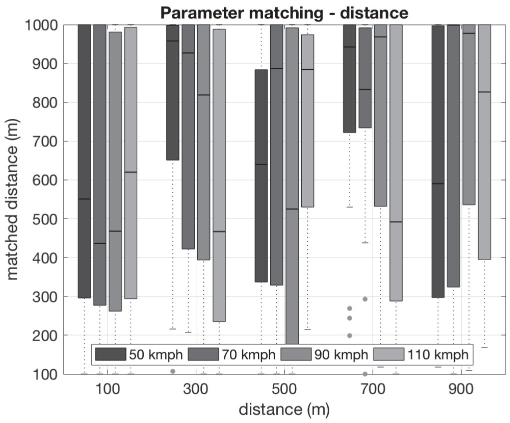


Figure 3.8: Parameter matching - Speed

Figure 3.9: Parameter matching - Distance

4 Conclusions

During the development of the simplified background traffic model, two reports summarise the process and the validation procedure. In Paper I, the concept and details for building the model are described, as well as a preliminary validation procedure using subjective listening tests. Data from the tests gave promising results in order to continue for a validation test with sufficient number of subjects, as described in Paper II. In this paper, further details and alterations of the model are given. The model is validated against a reference model, the LISTEN auralisation tool, which has earlier been successfully validated. The acquired results are presented in Paper II, and discussed on the following section.

4.1 Paper II – validation results

The subjective tests are designed in two parts. In the first, participants were given the task to match two for them unknown parameters, traffic speed and distance, to the output stimuli of the reference model, and then to assess the similarity between the two sounds. In the second part of the test, similarity is assessed as well, including foreground traffic noise from the reference model, in order to test the simplified model under a more realistic scenario. Here, the parameters are fixed to the physical quantities they represent. While parameter matching resulted in a broad spectrum of answers with low coherence, comparing similarity ratings of the two parts, improvement can be seen on the overall perception of the model. At a first inspection of the answers for similarity with the reference, the proposed model itself (without any additional stimuli), gives satisfactory results as the 75 % are reading 9 out of 10. When all answers are considered, the mean 6. Adding pass-by events that occur in closer proximity, the upper quartile (75 %) reaches to 10, and the mean is 8. As the mean values between these two tests are close, confidence intervals are calculated to give better insight of the results. The computed 95 % confidence intervals of the overall mean show that the distribution of similarity rating means do not overlap, with 6.0-6.4 for background traffic (first part), and 6.6 - 7.1 when adding local traffic (second part), while their range remains similar. This indicates that the ability of judging similarity of the two models, is shifted and not stretched between the two traffic scenarios, which brings the proposed method closer to validation results of the reference model when used in conjunction with other elements consisting the sonic environment. It can further be noticed in Figure 3.3, where confidence intervals of the similarity means are grouped by speed profile, that there is 95 % possibility that the mean value distributions of the two parts do not overlap for 70 km/h, while there is less overlapping at lower than at higher speeds. It should be expected then, that the model is more suitable to be used with other auditory elements (e.g. foreground traffic) on lower speeds. Lastly, to make certain that the mean of the similarity ratings do not interfere between the two test parts, a hypothesis test is performed, where the latter argument is correct and highly significant with $P < .001$.

4.2 Overall conclusions

The proposed model was created with an simplified basis in mind to achieve a computationally efficient method for background traffic noise synthesis, as a tool for auralising a part of an urban environment. As explained in Paper I, distant traffic is considered as shaped noise based on recorded data, without explicitly calculating every vehicle pass by. Modulation transfer functions are used for modulations in time and frequency that resemble the pass-by events, while ground effect, distance, and air attenuation are modelled explicitly. Doppler effect is achieved in two steps, by shifting the power spectrum input for each channel (left and right), and by the modulation transfer function ripples travelling upwards and downwards in frequency, in the respective channel. The noise generators of the two channels are cross-mixed, to achieve controlled coherence between the channels, with a cut-off frequency of 100 Hz. This serves as a stereo image controller of the model, which varies in accordance to traffic distribution. The model is perceived more realistic when nearby pass-by events near the listener occur, but more tests would be preferable in order to further distinguish types of auditory elements that should be modelled in more detail.

4.3 Future work

Enhancing the proposed model should not necessarily mean that every single car in it is explicitly described. Having additional tonal components as well as including time varying filters simulating vehicle directivity, could result in a clearer perception of the distance and speed of traffic. An additional attempt would be to mix detailed and simplified modelling, where only the two or three vehicles nearest to the listener can be modelled in detail, depending on their distribution in space, in reference to the listener. This might result in further spectral masking which gives a virtual perception that every element of the traffic can be heard.

Bibliography

[1] Andrew Whiten and Carel P Van Schaik. The evolution of animal 'cultures' and social intelligence. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 362(1480):603–620, 2007.

[2] Bernard Gauthiez. The history of urban morphology. *Urban Morphology*, 8(2):71–89, 2004.

[3] Sonia Alves, Laura Estévez Mauriz, Francesco Aletta, Gemma M. Echevarria-Sanchez, and Virginia Puyana Romero. Towards the integration of urban sound planning in urban development processes: the study of four test sites within the SONORUS project. *Noise Mapping*, 2(1):57–85, 2015.

[4] World Health Organization (WHO). Burden of disease from environmental noise Quantification of healthy life years lost in Europe. *Geneva, Switzerland*, 2011.

[5] World Health Organization (WHO). Night noise guidelines for Europe. Technical report, 2009.

[6] Daniel Shepherd, Kim Dirks, David Welch, David McBride, and Jason Landon. The covariance between air pollution annoyance and noise annoyance, and its relationship with health-related quality of life. *International Journal of Environmental Research and Public Health*, 13(8), 2016.

[7] Djamel Ouis. Annoyance From Road Traffic Noise: a Review. *Journal of Environmental Psychology*, 21(1):101–120, 2001.

[8] Aslak Fyhri and Ronny Klæboe. Direct, indirect influences of income on road traffic noise annoyance. *Journal of Environmental Psychology*, 26(1):27–37, 2006.

[9] Jan Schnupp, Israel Nelken, and Andrew King. *Auditory neuroscience : making sense of sound.* MIT Press, 2011.

[10] Aslak Fyhri and Ronny Klæboe. Road traffic noise, sensitivity, annoyance and self-reported health-A structural equation model exercise. *Environment International*, 35(1):91–97, 2009.

[11] Daniel Västfjäll. Influences of current mood and noise sensitivity on judgments of noise annoyance. *The Journal of psychology*, 136(4):357–370, 2002.

[12] Gabriel Moser and Monique Robin. Environmental annoyances: An urban-specific threat to quality of life? *Revue Europeene de Psychologie Appliquee*, 56(1):35–41, 2006.

[13] Maarten Hornikx. Ten questions concerning computational urban acoustics. *Building and Environment*, 106:409–421, 2016.

[14] Wolfgang Kropp, Jens Forssén, Laura Estévez-Mauriz, and (Eds.). *Urban Sound Planning – the SONORUS project.* Chalmers University of Technology, Division of Applied Acoustics, 2016.

[15] Jens Forssén, Tomasz Kaczmarek, Peter Lundén, Mats E Nilsson, and Jesper Alvarsson. Auralization of traffic noise within the LISTEN project: Preliminary results for passenger car pass-by. In *Euronoise 2009.* Institute of Acoustics, 2009.

[16] Michael Vorländer. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality.* 2008.

[17] Ulf Sandberg and Jerzy A. Ejsmont. *Tyre/road noise reference book.* INFORMEX, 2002.

[18] Keith Attenborough. Review of Ground Effects on Outdoor Sound Propagation from Continuous Broadband Sources. *Applied Acoustics*, 24:289–319, 1988.

[19] Thomas Beckenbauer, Philippe Klein, Jean-François Hamet, and Wolfgang Kropp. Tyre/road noise prediction: A comparison between the SPERoN and HyRoNE models - Part 1. *The Journal of the Acoustical Society of America*, 123(5):3388–3388, may 2008.

[20] Jens Forssén, Alice Hoffmann, and Wolfgang Kropp. Auralisation model for the perceptual evaluation of tyre–road noise. *Applied Acoustics*, 2017.

[21] Tomasz Kaczmarek. Road-vehicle simulation for psychoacoustic studies. In *International Congress on Acoustics*, Madrid, 2007.

[22] Julien Maillard and Jacques Martin. A simulation and restitution technique for the perceptive evaluation of road traffic noise. In *Forum Acusticum*, Budapest, 2005.

[23] David Keith Wilson and Lanbo Liu. Finite-Difference, Time-Domain Simulation of Sound Propagation in a Dynamic Atmosphere. Technical report, Hanover, N.H. : US Army Corps of Engineers, Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory, 2004.

[24] Timothy Van Renterghem, Erik M Salomons, and Dick Botteldooren. Efficient FDTD-PE Model for Sound Propagation in Situations with Complex Obstacles and Wind Profiles. 91:671–679, 2005.

[25] Rodney J. Vaughan. The Distribution of Traffic Volumes. *Transportation Science*, 4(1):97–110, feb 1970.

[26] George Marsaglia and Wai Wan Tsang. A simple method for generating gamma variables. *ACM Transactions on Mathematical Software*, 26(3):363–372, sep 2000.

[27] ISO. 532: Acoustics – Method for calculating loudness level. 1975.

[28] Karlheinz Gröchenig. *Foundations of time-frequency analysis.* Springer Science & Business Media, 2013.

[29] Alan V Oppenheim and Ronald W Schafer. Discrete-time signal processing. *New Jersey, Printice Hall Inc*, 1989.

[30] DIN ISO. 9613-1: 1993. Acoustics. Attenuation of sound during propagation outdoors. Part 1: Calculation of the absorption of sound by the atmosphere. *International Organization for Standardization, Geneva*, 1993.

[31] Jorgen Kragh, Birger Plovsing, S Å Storeheier, Gunnar Taraldsen, and H G Jonasson. Nordic Environmental Noise Prediction Methods, Nord2000 Summary Report. General Nordic Sound Propagation Model and Applications in Source-Related Prediction Methods. *DELTA Acoustics & Vibration, Report AV*, 1719(01), 2001.

[32] Nina Kowalski, Didier A Depireux, and Shihab A Shamma. Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *Journal of neurophysiology*, 76(5):3503–3523, 1996.

[33] Taishih Chi, Yujie Gao, Matthew C Guyton, Powen Ru, and Shihab Shamma. Spectro-temporal modulation transfer functions and speech intelligibility. *The Journal of the Acoustical Society of America*, 106(5):2719–2732, 1999.

[34] R. Nota, R. Barelds, and D. van Maercke. Engineering method for road traffic and railway noise after validation and fine-tuning. *Harmonoise project report D*, 18, 2005.

[35] Roger N Shepard. Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, 36(12):2346–2353, 1964.

[36] Gary S Kendall. The decorrelation of audio signals and its impact on spatial imagery. *Computer Music Journal*, pages 71–87, 1995.