inter.noise
HAMBURG 2016

# Background traffic noise synthesis

Georgios ZACHOS[1]; Jens FORSSÉN[2]; Wolfgang KROPP[3]; Laura ESTÉVEZ-MAURIZ [4]
[1,2] Division of Applied Acoustics, Chalmers University of Technology, SE-41296 Gothenburg, Sweden

## ABSTRACT

When planning the development of urban areas, it is important to assess the future acoustic environment. Currently, this evaluation is achieved with the help of acoustic indicators, but they do not suffice for a holistic description of the perceived sound environment. New indicators can be extracted through listening tests and analysis of different acoustic scenarios. However, generating such scenarios using auralisation models for outdoors sound propagation is often computationally highly demanding. Here, a simplified auralisation model is described, focusing on background traffic noise simulation on flat city scenarios. For computational efficiency, the proposed method partly relies on physical models for air attenuation, ground effects and spherical spreading. The doppler effect and the contribution of individual vehicle pass-bys are achieved with the help of modulation transfer functions, and spatial imagery is realised by both non-corellated phase spectra and modulation transfer functions. Power profiles from measurements are used to model rolling noise. The proposed model is assessed through listening tests against the LISTEN demonstrator on its perceived speed and distance from the listener. The perceived speed is matching better to the LISTEN between 70 kmph and 90 kmph, while above 300 m and up to 900 m from the source, the distance is more correctly guessed from the subjects.

Keywords: urban sound planning, auralisation, background traffic noise

## 1. INTRODUCTION

During, or preferably before, an urban development procedure of a residential area, the urban planner will realise that the main source of undesirable sound in the area, is, with a few exceptions, road traffic. The effects due to prolonged exposure of traffic noise are investigated and documented in such extent (1–3) that it is sensible to plan avoiding it. While local traffic contributes greatly to this noise, background traffic from distant roads is also a factor that adds an additional significant noise layer and worths to be concerned.

There are currently methods such as Nord2000 (4) that are used for noise evaluation of road traffic, but their output is limited to sound pressure level averages over frequency. The need for additional indicators which better describe reaction to traffic noise demands the right tools to invent them, and one would be auralising the scenarios in question, since an approach of using recordings is unfeasible considering the needed versatility.

Auralisation of traffic noise is considered within the LISTEN project (5) by auralising single car pass-bys. The LISTEN demonstrator uses parameter based models of the sound source and physical models of the propagation so it can be computationally heavy. The proposed model mainly uses profiles and envelopes to form the noise of distance traffic. The tonal part of individual car engines is not explicitly

[1] georgios.zachos@chalmers.se
[2] jens.forssen@chalmers.se
[3] wolfgang.kropp@chalmers.se
[4] laura.estevez@chalmers.se

modelled as the main component in the investigated scenarios (high speed roads) is dominated by road-tyre interaction.

The model is tested against the LISTEN demonstrator with preliminary subjective evaluation tests. For simplicity and better interpretation of the data, the tests only include flat-city scenarios. By pair wise comparison, in which the subjects are called to change parameters of the model to match to the LISTEN output.

## 2. METHOD

The description of the auralisation model will be presented below covering different stages of the procedure. As the noise resulting from the interaction between tyre and road dominates comparing to the car engine noise in high velocities, which the latter are mostly considered in distant traffic scenarios for urban sound planning, a noise signal is chosen as a source. Initially, the system is fed with pink noise, and passes through shaping profiles on the frequency domain in third-octave bands.

### 2.1 Spatial imagery

In most situations, distant traffic noise will cover a large angular range for a receiver, resulting in a wide spatial auditory image. To achieve this effect, the concept of decorellation is used. Total decorrelation of a signal between the left and right ear of a receiver will result in wide spatial imagery, thus a non localised noise source (6). Sequentially, to achieve a stereo effect for this model, the input noise's phase spectrum is randomised for each channel on the frequency domain, before reverting back to the time domain. As the noisy signal is passed through a short time Fourier transform (STFT), it will result in a periodic signal which can be used concatenated with itself, given that the windowing of the signal is short enough to avoid perceivably repetitive patterns. Here, a window of 2048 samples is chosen, with 75 % overlap.

### 2.2 Source speed profile

This stage serves as the source model of the system. The third-octave band power profiles of LISTEN recorded data is used. The velocities of the profiles range from 30 kmph to 110 kmph for every 20 kmph and are used to filter the input in the frequency domain with 28-band 6th order filters. For in-between values linear interpolation on each octave band is used. The filter coefficients use a damping algorithm to allow smooth transitions between different speed profiles. The power profiles are loudness equalised according to (7). Furthermore, for the left and right channel of the listener, the power profiles have been shifted respectively up and down in frequency, in order to simulate the doppler effect of traffic with constant speed. This implementation allows using any kind of noisy source profile as an input.

### 2.3 Air attenuation and ground effect

A standard method is used for air attenuation (8), with parameters matching the ones of the LISTEN project (5) for direct comparison. Similarly the ground effect, i.e. the signal reflected from the ground to the listener, is modelled as in (5). As the scenarios considered here involve only distant traffic, both air attenuation and the ground effect are considered to remain constant in time during a single pass-by. Working in third-octave bands, the output of the filters pass through this stage. Air fluctuations are not considered here and amplitude correction corresponding to spherical spreading is performed after the reconstruction of the signal in time domain.

## 2.4 Individual pass-bys

Even on distant traffic scenarios, in addition to the accumulation of car noises, individual pass-bys can still be perceived especially when a line of sight from the observer to the source exists. Instead of simulating single pass-bys and arranging them in time, another approach has been followed.

Dynamically rippled spectra have shown to be largely effective in evoking responses in the auditory cortex (9). Their application on speech intelligibility research by (10), has resulted in the spectro-temporal transfer functions (MTF). MTFs are considered fit to measure detection thresholds on hearing, as the neural auditory system has shown higher intelligibility skills when these are present. As such, they will be used as a synthesis tool for a plausible auralisation of car pass-bys. MTFs are orthogonal and are described as:

$$MTF(x) = A \cdot \sin\left(2\pi \cdot (\omega \cdot t + \Omega \cdot x) + \phi\right) \tag{1}$$

where $x = \log_2(f/f_c)$, with $f$ frequency in Hz and $f_c$ is considered the frequency where a minimum occurs. The amplitude of the ripples is defined by $A$, the ripple velocity $\Omega$ is described in cycles/octave, and $\omega$ is their velocity in cycles/second; $\phi$ is the phase of the function.

To simulate single pass-bys, MTFs are shaping the noise source in the frequency domain before applying propagation effects. As the density of the traffic decreases, a single pass-by is more distinct, and as such the amplitude of the MTF will increase to give a more pronounced pass. The modulation amplitude factor $A$ of the MTF is set to 2.5 dB, as the directivity pattern of the tyre-road interaction noise demonstrates a 2.5 dB drop from the line of direction of motion of the vehicle to 90 degrees around it (11). A dip of the functions appears on the closest point between the receiver and a car on its route alongside a line.

The scenarios tested here, only include constant traffic flow with constant speed. In order to not avoid energy fluctuations to the receiver during the pass-bys, a pass-by overlap frequency (POF) equal to 1 is set for the listening tests. According to the distance of the receiver, and velocity of the pass-bys, $POF = 1$ sets the timing between them as such that when an approaching vehicle is at -3 dB below its peak value (midpoint, closest to the receiver), the leaving vehicle is also at -3 dB.

### Doppler

For an accumulated doppler effect of a one way lane, the frequency spectrum of the power profiles of each channel (left and right) is shifted. This will give the effect of a single lane road where there is a constant speed profile from the left to the right when looking to the road. On a realistic case of a one-way drive however, due to the doppler effect the frequency of a passing car will be rising when approaching the listener and falling upon distancing, weighting the effect accordingly on the left and right ear. To include this to the proposed model, the fact that MTFs do not only provide a modulation of the spectra in time but also in frequency, is used. Depending on the speed direction of the ripples or alternatively, as it is used here, the increment sign of the phase, they travel upwards or downwards on the frequency axis. This creates the perception of a constant ascension or descension of the signal's pitch the same way it is experienced with Shepard tones (12), and thereby may be able to simulate the doppler effect for vehicles passing both left to right and right to left.

### Pseudo-stereo effect

The MTFs are set to travel along the frequency axis with an opposite direction on each channel, namely upwards on the left, and downwards on the right for left-right passage and conversely. By choosing the right parameters $f_c$ and $\Omega$ of the MTFs, a perceptual motion will result in the resulting field, either coming from the left and continuing to the right on the virtual stereo image or conversely. Here, $f_c$ has been chosen to be the frequency where the two MTFs begin with $\phi = 0$, and thus are identical. Furthermore, when $\phi = 2n\pi$, an individual pass-by is considered to be positioned midway, that is the closest point of the road to a listener. A psychoacoustic explanation of this effect is yet unknown to the author and as such,

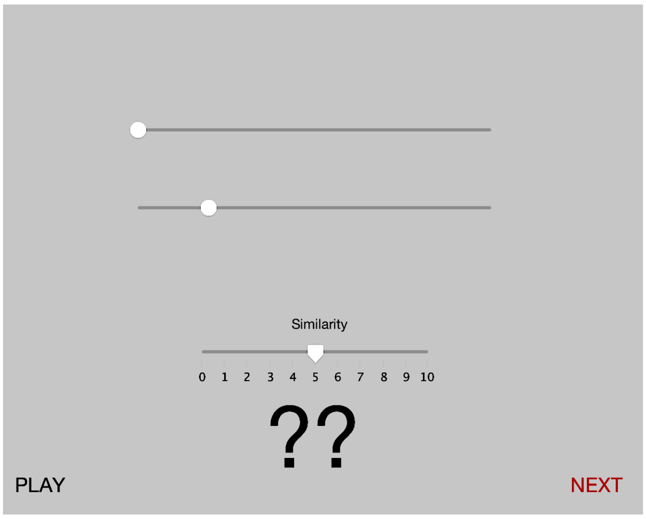


Figure 1 – Graphical interface built for the tests. It provides interaction time control of the parameters speed and distance of the model, as well as a switch between the modelled sounds and the LISTEN mix

screening tests with expert subjects were performed to determine the values to achieve the best working of the effect. It has been found that for $f_c = 1.5$ kHz, $\Omega = 0.3$ cycles/octave, and the phase increasing on the left channel and decreasing on the right, the virtual source appears to be moving from left to right, and vice versa for right to left when the changes in the phase of the modulations are inverted.

## 3. Subjective preliminary tests

In order to assess the validity of the proposed model, it is tested subjectively against the output of the LISTEN demonstrator. Pair matching with a mixed output from the LISTEN demonstrator using vector based amplitude panning for source localisation (13) is asked from the participants. For each pare, pass-bys of a single speed from 50 kmph to 110 kmph and constant uniform flow is chosen. They could freely switch between the LISTEN mix and the output from the synthesis model, performed on real/user interaction time. The parameters speed and distance of the model are randomised after each comparison, and the subject's task is to try and match these values to the ones of LISTEN through a graphical interface as shown in figure 1. The symbol noted as "??" on the interface serves as a switch between the model output (the button has the "??" label) and the LISTEN mix (the button is labelled as "!!"). After a decision is made, it is asked to assess how similar are the sounds after the parameters have been matched. The participants were not informed that the given stimuli represent traffic noise or what the parameters effect. The test is considered preliminary as only 10 participants contributed.

### Results

The box plot of similarity ratings in Figure 2 show how similar the model is perceived comparing to the LISTEN mixed output. The results are grouped by vehicle velocity and the shading represents variations on distance from the traffic. The median similarity is betwen 40 % to 60 %, where in some cases the upper quartile (75 % of the answers) reaches 70 %. The ratings reach 80 % at 70 kmph and 110 kmph for 25 % of the listeners (upper quartile to upper whisker). Outliers are observed at 100 m distance for all the speed profiles, and some at 300 m for 90 kmph.

Figure 3 presents a box plot of the the speed parameter matching to the LISTEN mix, grouped by vehicle velocity. While background traffic noise of 110 kmph could not be matched correctly, the matched speed parameters follow an ascending curve that follows the speed profiles from the reference sounds. Especially for the LISTEN sounds at 70 kmph, the matched median parameters are between 60 kmph and 70 kmph. According to middle quartile results (50 % of the answers) the speed matching can be

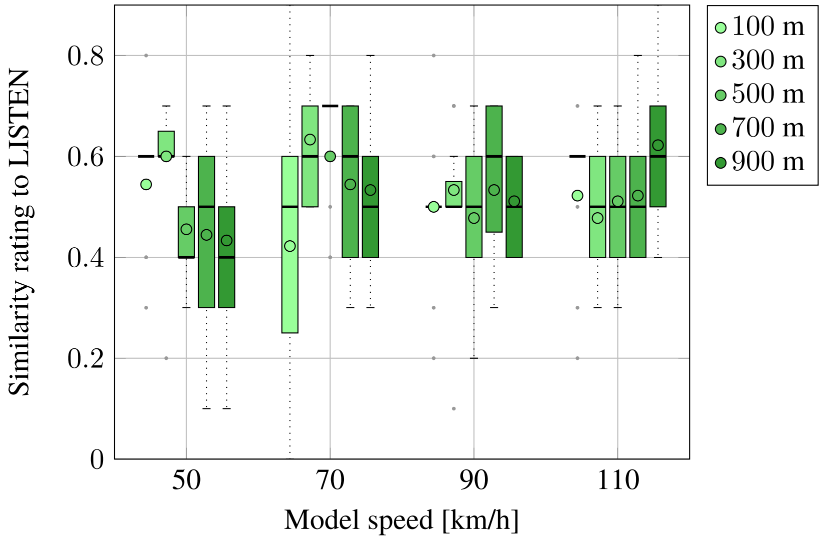


Figure 2 – Similarity ratings; from light to dark, distances from 100m to 900m every 200m

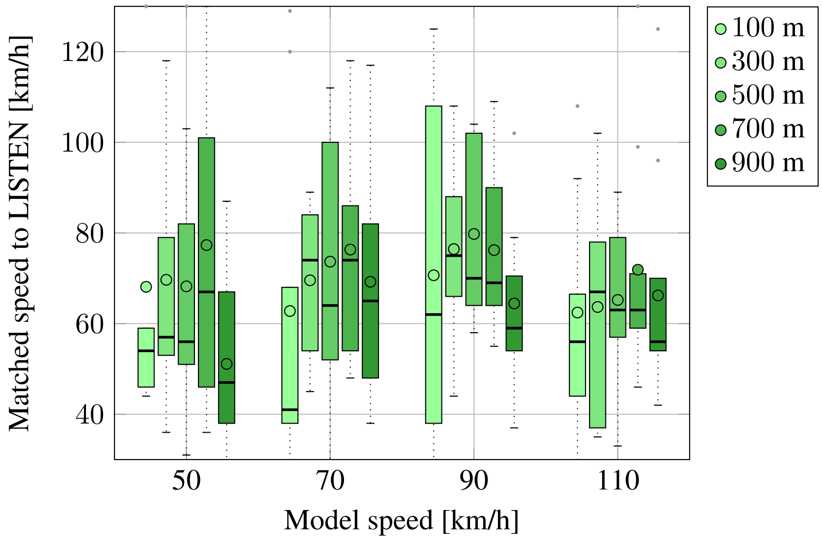


Figure 3 – Speed match results; from light to dark, distances from 100m to 900m every 200m

considered acceptabe. For all cases, closest distance (100 m) resulted in answers that either deviated a lot from the presented scenario as seen at 70 kmph, or were too widespread as for the case of 90 kmph where 50 % of the answers cover a range from 40 kmph to almost 112 kmph. Most of the outliers are found at 110 kmph.

Distance matching grouped according to different distance scenarios is shown in Figure 4 where the shading represents different speed profiles. Except at 100 m away from the source, the perceived distance of traffic noise from the proposed model follow the distance of the sounds generated from the LISTEN demonstrator. At 100 m only 50 kmph is matched correctly. The scenarios in which 70 kmph is tested, show a wider spread in the answers and mismatch at 500 m. There are a outliers mostly on 50 kmph and 900 m away.

## 4. Discussion

Due to computational obstacles of simulating dynamic traffic noise, an efficient simplified model for auralising background traffic noise has been proposed. Source modelling and propagation affects are calculated in real time and applied with overlapping short Fourier transforms and as such can be run and

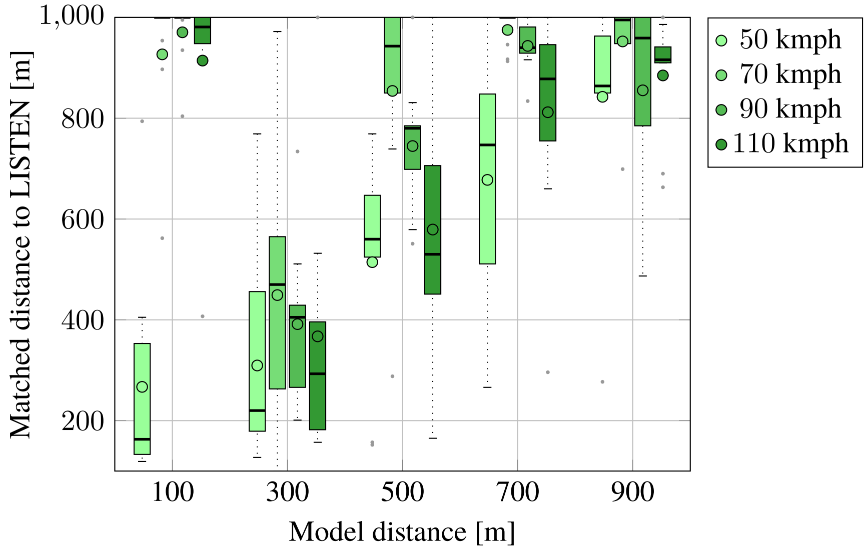


Figure 4 – Distance match results; from light to dark, velocities from 50 kmph to 110 kmph every 20 kmph

modified on real time. The model has been tested against the output of the LISTEN demonstrator through preliminary listening tests. The participants were asked to match two of the parameters of the model as close as they could to the validated LISTEN output, the pass-by car velocity and distance. By consulting the results of the tests, the model can be considered acceptable on 70 kmph for distances from 300 m and above. Shorter distances and higher speed profiles were more difficult to be matched at the LISTEN output, while 110 kmph could not be matched at all. Similarity ratings of 75 % would be preferred instead of 40 % to 70 %, but further listening tests with more samples should be conducted.

The difficulty of matching correctly on parameters in some cases, could be an effect of the proposed model as the source is modelled with filtered noise, in contrary with the LISTEN output in which also tonal components are explicitly added to the source model. For low velocities up to 50 kmph, at least on the provided data, the tonal components of the source contribute at the same levels than the road-tyre interaction noise and can be more prominent. For 70 kmph the rolling noise exceeds the tonal characteristics of the engine, and for higher speeds the tyre noise dominates even more. Low velocities in the model might require tonal components to be perceived correctly. On higher speeds, where the tonal characteristics of the source shift upwards in frequency towards the tyre noise peak, additional tones could be a requirement to avoid masking and make the tones more prominent.

An interesting effect has also been observed during development of the model, namely the pseudo-stereo moving source effect. When two modulation transfer functions used to shape noise are combined one for each channel/ear and with opposite velocities between the channels, a motion modulation on the listener's spatial field appears which follows different directions according to $f_c$ and $\Omega$. These parameters have been determined empirically through screening tests, but through a psychoacoustics approach it might yield more information on the auditory system's response on MTFs.

## Future work

As few subjects participated for the assessment of the model, it should be conducted again with more subjects to distinguish any tendencies. An upgraded model synthesising tonal components by either additive or granular synthesis would be able to run in real time as well. Further investigation of combining modulation transfer functions on both ears of a listener, might result in useful tools for the position of moving elements in the virtual acoustic field.

Upon successful validation of the model, the next step would be to test more realistic scenarios including both foreground and background traffic noise. Adding sound obstructive constructions like sound barriers or buildings, the model can be assessed for background noise traffic suitability. Testing it also with the presence of other urban traffic elements, it will be interesting to distinguish the threshold levels of the latter, above which the background traffic model is indistinguishable from a validated model like the LISTEN one or on site recordings.

## ACKNOWLEDGMENT

The authors would like to thank Lars Hansson for his help on the development on a programming and technical level of the real time demonstrator of the model, that was also used at the listening tests.

## References

(1) Birgitta Berglund, Thomas Lindvall, Dietrich H Schwela, et al. Guidelines for community noise. In *Guidelines for community noise*. OMS, 1999.

(2) Lars Barregard, Ellen Bonde, and Evy Öhrström. Risk of hypertension from exposure to road traffic noise in a population-based sample. *Occupational and environmental medicine*, 66(6):410–415, 2009.

(3) Wolfgang Babisch, Bernd Beule, Marianne Schust, Norbert Kersten, and Hartmut Ising. Traffic noise and risk of myocardial infarction. *Epidemiology*, 16(1):33–40, 2005.

(4) Jorgen Kragh, Birger Plovsing, SÅ Storeheier, Gunnar Taraldsen, and HG 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.

(5) 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.

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

(7) ISO. 532: Acoustics – method for calculating loudness level. 1975.

(8) 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.

(9) 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.

(10) 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.

(11) R Nota, R Barelds, and D van Maercke. Technical report har32tr-040922-dgmr20 harmonoise wp 3 engineering method for road traffic and railway noise after validation and fine-tuning. *Harmonoise project report D*, 18, 2005.

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

(13) Ville Pulkki. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, 45(6):456–466, 1997.