DOI: 10.61782/fa.2025.0122

FORUM ACUSTICUM
EURONOISE 2025

AN APPROACH FOR EVALUATING TIME-VARYING ANNOYANCE CAUSED BY RAILWAY NOISE

Jonas Egeler[1] **Christine Huth**[1] **Anton Schlesinger**[1] **Manfred Liepert**[1]
Christoph Ende[2] **Daniel Johannes Meyer**[2] **Thomas Koch**[2]
Jens Bartnitzek[3] **Laura Höhle**[3]
Benjamin Schlüter[4] **Ralf Böhme**[4]
[1] Moehler + Partner Ingenieure GmbH, Augsburg, Germany
[2] Fraunhofer Heinrich Hertz Institute, Berlin, Germany
[3] A+S Consult GmbH, Dresden, Germany
[4] Deutsche Eisenbahn Service AG, Putlitz, Germany

ABSTRACT

We propose a new approach for predicting railway noise annoyance. Using an experimental design to record time-varying annoyance combined with a machine learning model, we developed a psychoacoustic metric that aligns more closely with human annoyance perception than the traditional A-weighted SPL. This work is part of the EAV-Infra project, which explores the application of building information modelling, auralization, visualization and psychoacoustics in infrastructure planning (www.eav-infra.de, [1]).

To evaluate railway noise annoyance, we recorded audio of 335 train passings. From this dataset, 50 recordings were selected and presented to 22 participants in a listening test. Participants rated time-varying annoyance using a mechanical slider.

Statistical evaluation shows highest inter-personal variation in responses for medium annoyance levels and lower variation at the extremes. Particularly annoying events like rattling, squeaking, or flatspots are clearly visible in the median response. A CRNN was trained with mel-log spectra of the passings to predict the median time-varying annoyance ratings, achieving an RMSE of 5.93, well within the inter-personal interquartile range.

**Corresponding author*: *jonas.egeler@mopa.de.*

Keywords: *psychoacoustics, railway noise, annoyance, eav-infra*

1. INTRODUCTION

Railway noise is a significant environmental issue that affects millions of people living near railway lines. Understanding the psychoacoustic aspects of railway noise annoyance is crucial for developing effective noise mitigation strategies and improving the quality of life for affected communities. In the following, we provide a short overview of the current research on railway noise annoyance, focusing on key studies and findings in the field.

The railway bonus ("Schienenbonus") in Germany was a regulation that allowed railway noise to be assessed as 5 dB(A) lower than road traffic noise, under the assumption that railway noise was less disturbing. It was introduced in 1990 based on sociological studies from the late 1970s and early 1980s and aimed to promote rail transport by making it easier to meet noise protection standards.

Hugo Fastl's study "Railway bonus for sounds without meaning? – Loudness is main reason for railway bonus" [2] investigates why railway noise is often perceived as less annoying than road traffic noise, a phenomenon known as the "railway bonus". Fastl and his colleagues conducted psychoacoustic experiments to determine if the meaning of the sound contributes to this bonus. They found that even when the meaning of the sound was removed, the railway bonus persisted, suggesting that factors like spectrum and time structure play a

more significant role. The study concluded that loudness, rather than the meaning of the sound, is the main reason for the railway bonus. This research helps in understanding how different characteristics of noise influence human perception and annoyance.
Lercher et al. focused on the impact of railway noise in sensitive areas, particularly during nighttime [3]. The research involved classical and binaural sound measurements in an alpine valley, revealing that freight trains are significantly louder than passenger trains. The study found that traditional noise assessments often underestimate the annoyance and sleep disturbance caused by railway noise, especially in complex acoustic environments. It was concluded that the application of a railway bonus is inappropriate in such conditions, as it does not accurately reflect the true impact of railway noise. This work highlights the need for more nuanced noise assessment methods that consider psychoacoustic factors.
On January 1, 2015, because newer studies indicated that the original assumptions were outdated and no longer justified, the railway bonus was abolished in Germany.
The impact of low-frequency content, impulsive events, flatspots, and squealing on noise annoyance is significant and multifaceted. Low-frequency noise can penetrate buildings and other structures more easily than higher-frequency noise, leading to greater indoor annoyance and sleep disturbance. Impulsive events, such as the sudden noise from train brakes, can startle individuals and cause a higher level of annoyance compared to continuous noise. Flatspots on train wheels, which occur due to uneven wear, can create repetitive thumping sounds that are particularly irritating to listeners [4]. Squealing noises, often produced by train wheels on curves, are high-pitched and can be extremely disturbing, contributing to increased annoyance and stress [5]. These factors highlight the complexity of railway noise and the need for comprehensive assessment methods that consider both the physical and psychological impacts of noise exposure. Addressing these specific noise characteristics is crucial for developing effective noise mitigation strategies and improving the quality of life for those living near railway lines.
Recent studies have further explored the relationship between psychoacoustic factors and railway noise annoyance. For instance, Kasess et al. [6] conducted a study on the relation between psychoacoustical factors and annoyance under different noise reduction conditions for railway noise. They investigated various conditions of pass-bys of cargo and passenger trains, including free field and spectral mitigations caused by noise barriers and rail dampers. The study found that loudness-level-based models were more effective in describing annoyance compared to models based on the A-weighted sound pressure level. Understanding the influence of signal characteristics on railway noise annoyance allows for the creation of more accurate assessment methods that truly reflect the impact of railway noise on affected communities, ultimately improving the quality of life for those living near railway lines. Given that the time-varying perspective is currently under-represented in existing research, we will place particular emphasis on this aspect in our study.

2. STIMULI

Figure 1. Measurement point close to the highspeed-track 5919 nearby Bad Staffelstein/Germany

Within the framework of the EAV-Infra initiative, an extensive measurement campaign was performed across seven sites in Germany, in order to capture audio of train pass-by events in diverse sound propagation scenarios, such as open fields, noise barriers, and multiple buildings. The campaign aimed to gather sound source data for auralization, evaluate the impact of noise barriers and rail dampers, and record various train passings with a binaural microphone for psychoacoustic analyses and development of the new annoyance metric. In figure 1 one of the measurement positions is shown. More detailed information on the campaign has been published in [7].
A HEAD acoustics BSU III.2 artificial head was utilized to record the binaural signals. At four of the recording sites, six additional measurement microphones (Microtech Gefell MV210 + MK255) were positioned in an

AB setup to derive transfer functions. These were used to virtually relocate the artificial head from its original position, 12.5 meters from the track without noise mitigation measures, to the remaining five microphone positions. Various train types at different speeds were included, and further variability was created by filtering the recordings with the AB transfer functions for noise mitigation measures and complex sound propagation scenarios. This process augmented the initial dataset of 111 binaural passings to a total of 335 binaural signals.
For the listening test on railway psychoacoustic annoyance, the dataset was narrowed down to keep the required time at an acceptable limit. A subset of 50 train passings was selected to ensure equal distribution in terms of annoyance, speed, passing time, and train type. The statistical characteristics of this dataset are presented in figure 2.

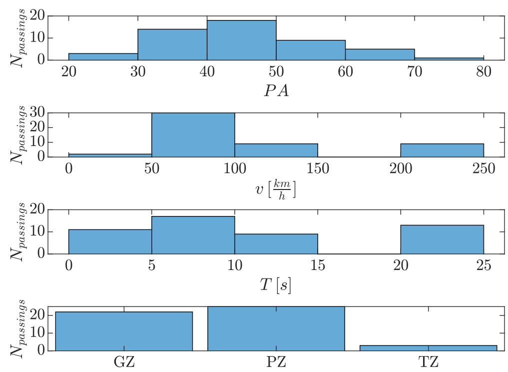

Figure 2. Distribution of psychoacoustic annoyance PA, speed v, passing time T and train type (GZ: cargo train, PZ: passenger train, TZ: individual locomotive) in the dataset comprising a total of $N = 50$ passings.

3. LISTENING TEST

The annoyance caused by train passings of varying lengths is expected to fluctuate over time, influenced by distinct events such as squeaking, impulses, or periodic flat spot occurrences. To understand the subjectively perceived time-varying annoyance of a wide range of train passings with a stimulus length up to 25 seconds, the line length method ("Methode der Linienlänge" [8]) was utilized. Additionally, the method of magnitude estimation [9]) was employed to obtain overall, non-time-dependent annoyance ratings.

Figure 3. Experimental setup with mechanical slider, visual feedback and questionnaire

A selection of 50 train passings was presented to 22 participants (9 female, 13 male, aged 22 to 55) using calibrated STAX SR-307 headphones. Prior to evaluating annoyance, participants received a detailed introduction to the task, including examples of the sound scenarios to be assessed. The definition of noise annoyance was communicated as an "adverse attitude that people form toward sounds that distract attention from or otherwise interfere with ongoing activities such as speech communication, task performance, recreation, relaxation, and sleep" [10].
Each stimulus was presented three times in a randomized order. Participants were instructed to adjust a 100 mm analog fader to reflect their current perceived annoyance during the playback of each passing. The scale, ranging from 0 to 100, and the current annoyance level set with the fader were displayed in real-time via a VST-plugin. After each passing, participants had 5 seconds to record their overall annoyance rating in a questionnaire.
The stimuli were presented in three blocks, each containing 50 passings. The passings within each block were randomized, and there was a five-minute break between blocks. The experimental setup is illustrated in Figure 3.

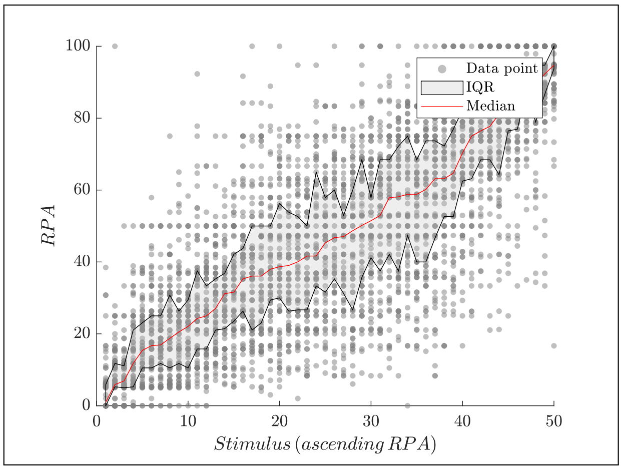

Figure 4. Overall annoyance ratings: grey dots show intra-personal median ratings over the three repetitions per stimulus, red line shows inter-personal median and grey area shows inter-personal interquartile range. Stimuli are sorted by ascending RPA.

4. INVESTIGATION OF TIME-VARYING AND OVERALL ANNOYANCE

Figure 4 gives a detailed overview of the overall annoyance ratings. It can be seen, that almost the entire range of the scale was used. For the least and the most annoying stimuli, the interquartile range is smaller compared to the stimuli of medium annoyance.

In figure 5, the time-varying listening test results are plotted. For each of the 50 pass-bys, the inter-individual median and inter-individual interquartile ranges of both the continuous ratings (red line with red shaded area) and the overall annoyance ratings (purple line with purple shaded area) are shown. It can be seen that the 50 selected stimuli cover a fairly wide range of annoyance. For some recordings (e.g., row 1, column 1), there is a prominent dynamic in the annoyance rating, with some particularly annoying events clearly standing out.

An interesting question is what criteria the subjects used to form their overall assessment of the annoyance of the passing trains. While it was initially assumed that the test subjects were guided by the peaks of the continuous assessment, similar to overall loudness [8], the listening test results for several stimuli (e.g., LDO|SDD50B|GZ|94 in row 1, column 2 or LDO|SSDD25A|GZ|70 in row 2, column 1) contradict this assumption. It can be seen that the overall assessment trends rather towards the median value of the individual timesteps.

For a better comparison of the continuous and overall assessments, the median of the individual time steps of the continuous assessment was calculated (red dotted line). For approximately one-third of the stimuli, the overall judgement results are below the median of the continuous assessment, for another third, the results are approximately the same, and for the last third, they are above the median of the continuous assessment. Therefore, the evaluation of overall annoyance seems to be based not only on the maximum annoyance reached during the continuous assessment.

To determine the actual reasons for the difference between the continuous and overall assessments, a statistical analysis of the 50 signals was conducted. The difference ΔRPA between the overall annoyance and the median of the individual time steps of the continuous assessment was compared with various other parameters. Figure 6 shows the relationship between ΔRPA and the mean pass-by speed, pass-by duration, and the psychoacoustic quantities loudness, sharpness, fluctuation strength, and roughness. While speed and stimulus length apparently have no influence on the difference between overall and continuous annoyance, loudness and fluctuation strength result in the highest correlations, with $r = 0.60$ and $r = 0.57$. For quiet and non-fluctuating stimuli, the overall annoyance is below the median of the continuous assessment, but for loud and strongly fluctuating stimuli, the overall assessment is above the median of the continuous assessment.

5. PREDICTION OF RPA AND COMPARISON WITH OTHER ACOUSTIC MEASURES

To generalize the findings of the listening test and predict $RPA(t)$ for new audio material, a machine learning approach using the mel-log spectrogram representation of the passings in combination with a Convolutional Recurrent Neural Network (CRNN) was investigated. Details on the model architecture and training process have been published in [11].

In figure 7, it is demonstrated for one exemplary pass-by how the trained CRNN is able to predict the listening test results for unseen data. The mismatch between the target signal $RPA(t)$ and the predicted signal $\widehat{RPA}(t)$ is well within the inter-individual interquartile range of the listening test and the temporal dynamics are reproduced quite well. The overall prediction error (RMSE between the target and predicted signals of the unseen test dataset) is 5.93.

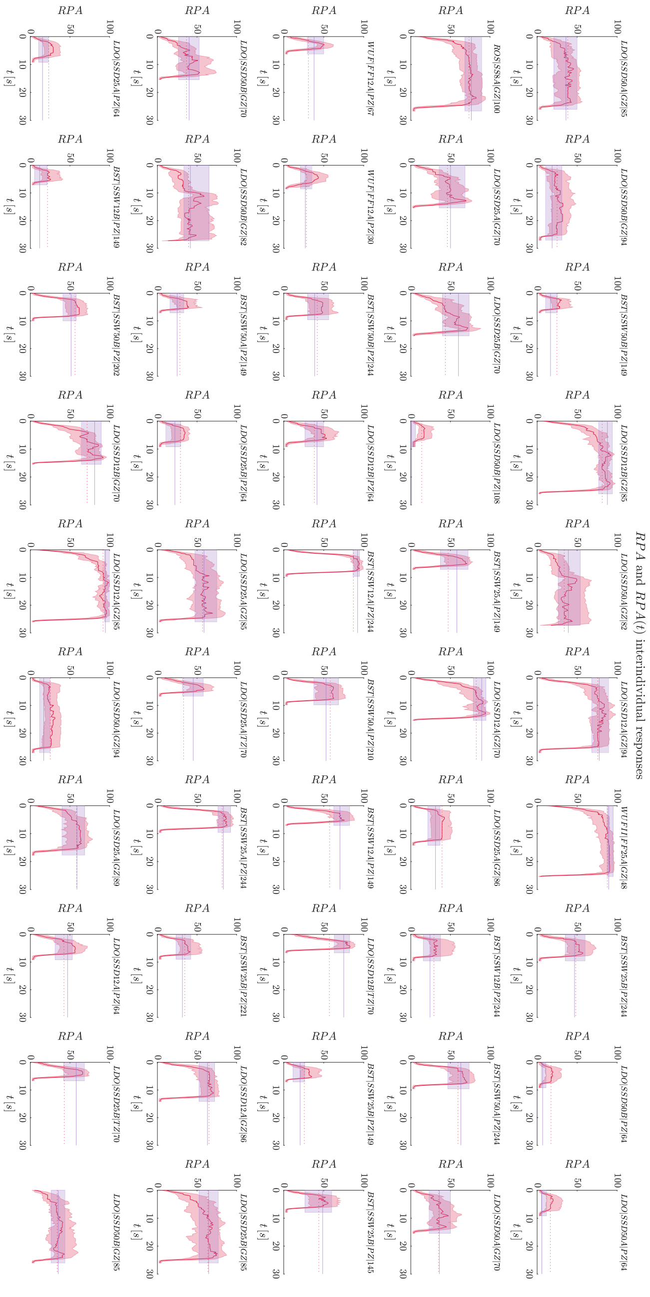

Figure 5. Time-varying $RPA(t)$ (red) and overall RPA (purple) in comparison. Bold line: inter-individual median, shaded area: interquartile range. Red dotted line: median over all timesteps in $RPA(t)$ signal.

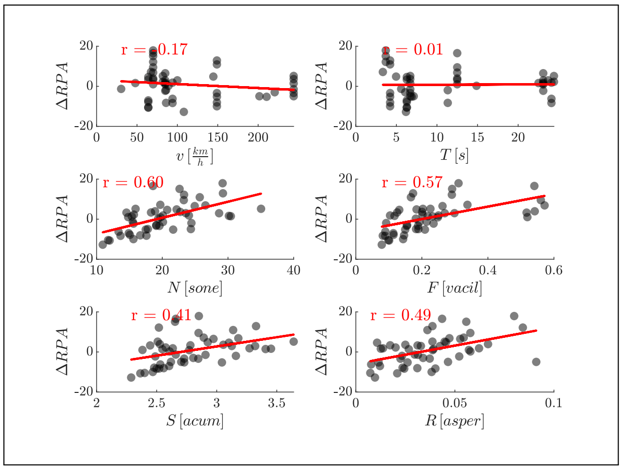

Figure 6. ΔRPA is the differene between overall RPA and the median over all $RPA(t)$ timesteps. The influence of various parameters on ΔRPA is shown.

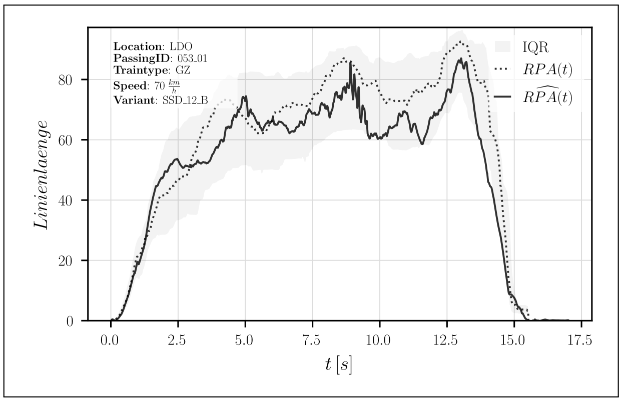

Figure 7. Comparison of target signal $RPA(t)$ and predicted signal $\widehat{RPA}(t)$ for one exemplary pass-by of the test dataset.

In figure 8, $L_{AF}(t)$ and $\widehat{RPA}(t)$ are plotted against time for two train pass-bys: one cargo train and one passenger train. These pass-bys were deliberately chosen because they exhibit similar L_{AF} levels throughout the entire signal, yet the perceived annoyance of the cargo train passing is significantly higher. This comparison demonstrates that $L_{AF}(t)$ does not capture all aspects of railway noise annoyance.

In figure 9, $PA(t)$ and $\widehat{RPA}(t)$ are plotted against time for the same two pass-bys as in the previous example. It can be seen that the $PA(t)$ indicator is able to capture some of the temporal peculiarities of the cargo train pass-by, particularly the flatspot event at half-time. However, it does not effectively capture the overall annoyance level, which, according to the listening test, is significantly higher for the cargo train pass-by compared to the passenger train pass-by.

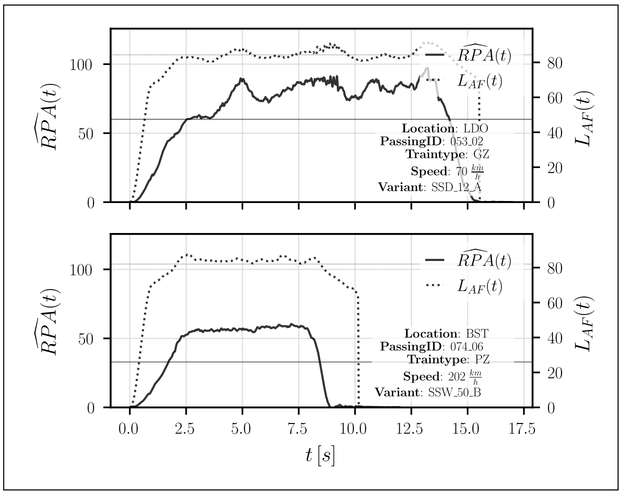

Figure 8. Comparison of an annoying cargo train (GZ) passing and a less annoying passenger train (PZ) passing: $\widehat{RPA}(t)$ vs. $L_{AF}(t)$.

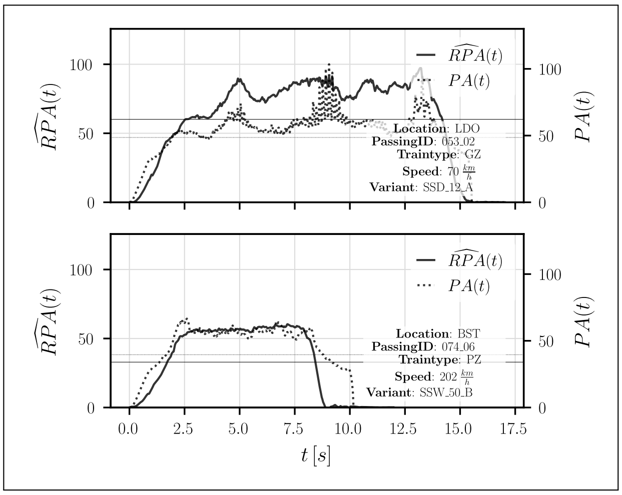

Figure 9. Comparison of an annoying cargo train (GZ) passing and a less annoying passenger train (PZ) passing: $\widehat{RPA}(t)$ vs. $PA(t)$.

6. CONCLUSION

Our study highlights the complexity of railway noise annoyance and the limitations of the traditional $L_{AF}(t)$ in capturing the full spectrum of perceived annoyance. The use of the CRNN approach for the metric design, demonstrates promising results in predicting time-varying annoyance levels for unseen audio material. The findings also underscore the importance of considering psychoacoustic factors like loudness and fluctuation strength, which significantly influence the overall annoyance rating in comparison to the time-dependent rating.
Future research should focus on not just individual pass-bys but also on complex traffic situations and propagation scenarios. This will help in understanding the cumulative impact of multiple noise sources and their interactions, leading to more comprehensive noise mitigation strategies.

7. PROJECT CONSORTIUM AND FUNDING

The research project EAV-Infra is funded as a strategic project by the Federal Ministry of Economics and Climate Action. The project is scheduled from March 2022 to July 2025.

8. REFERENCES

[1] C. Ende, T. Koch, T. Kron, L. Höhle, J. Bartnitzek, A. Schlesinger, J. Egeler, C. Huth, R. Böhme, and B. Schlüter, “Auralization and visualization for infrastructure planning in the joint project ”eav-infra”,” in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 268, pp. 4569–4580, 2023.

[2] H. Fastl, M. Fruhmann, and S. Ache, “Railway bonus for sounds without meaning?,” *AG Technische Akustik, MMK, TU-Muenchen*, 2003.

[3] P. Lercher, D. Kühner, H. Lin, and A. Fiebig, “Psychoacoustic assessment of railway noise in sensitive areas and times: is a rail bonus still appropriate?,” in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, (Innsbruck, Austria), pp. 5900–5907, 2013.

[4] C. Huth, J. Egeler, and M. Liepert, “Studie zur quantifizierung von flachstellen auf dem deutschen schienennetz,” tech. rep., Moehler + Partner Ingenieure AG, 2023. Herausgeber und Programmverantwortung: Bundesministerium für Digitales und Verkehr.

[5] T. Maly, C. Kasess, C. Kirisits, H. Waubke, F. Biebl, and N. Ostermann, “Psychoakustische analyse auffälliger bogengeräusche paab,” tech. rep., Technische Universität Wien, 2023. Herausgeber und Programmverantwortung: Bundesministerium für Klimaschutz.

[6] C. H. Kasess, T. Maly, P. Majdak, and H. Waubke, “The relation between psychoacoustical factors and annoyance under different noise reduction conditions for railway noise,” *The Journal of the Acoustical Society of America*, vol. 141, no. 5, pp. 3151–3163, 2017.

[7] J. Egeler, C. Huth, A. Schlesinger, C. Ende, T. Koch, J. Bartnizek, L. Hoehle, and B. Schlueter, “Railway psychoacoustic annoyance - an indicator for planning infrastructure projects,” in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, pp. 7247–7258, 2024.

[8] I. Stemplinger, *Beurteilung, Messung und Prognose der Globalen Lautheit von Geräuschimmissionen*. PhD thesis, Technische Universität München, 1999.

[9] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models*. Berlin, Heidelberg: Springer-Verlag, 3rd ed., 2007.

[10] S. Fidell, “Noise-induced annoyance,” in *Handbook of Noise and Vibration Control* (M. Crocker, ed.), ch. 25, pp. 316–319, Hoboken, NJ: John Wiley and Sons, 2007.

[11] J. Egeler and C. Huth, “Railway psychoacoustic annoyance: Custom time-varying psychoacoustic metric design with crnns,” in *Fortschritte der Akustik - DAGA 2025*, (Copenhagen, Denmark), 2025.